Reprint of

Protein Folding Handbook

Edited by Johannes Buchner and Thomas Kiefhaber

39
Engineering Proteins for Stability and Efficient Folding

Bernhard Schimmele and Andreas Plückthun

39.1
Introduction

The industrial, biotechnological, and medical applications of proteins are often limited by an insufficient protein stability or related problems. Such applications commonly require that proteins be produced on a large scale and remain stable enough to fulfill their functions for a reasonable length of time, often under harsh conditions. However, natural proteins are typically only marginally stable, and it is thus a major challenge for protein engineers to optimize stability and folding efficiency. The approaches that have been successfully employed to achieve this goal are rational design, semi-rational strategies based on sequence comparisons, and the methods of directed protein evolution. Of course, these methods are not mutually exclusive and can be combined to solve practical problems. All studies employing these methods have revealed important rules for protein engineering and at the same time shed light on the principles and mechanisms responsible for the folding and stability of proteins. Recent advances in stability engineering have demonstrated that merely small changes in a given protein sequence can have profound effects on its biophysical properties. The major challenge is therefore to correctly identify and remedy these shortcomings. It is the goal of this chapter to summarize the biophysical principles and technological approaches useful in improving the biophysical properties of proteins through sequence modification.

Considering the enormous array of technologies involved in this endeavor, ranging from computer algorithms to selection technologies, it is not possible to give detailed experimental protocols in this chapter; instead, we will guide the reader to the cited literature.

39.2
Kinetic and Thermodynamic Aspects of Natural Proteins

39.2.1
The Stability of Natural Proteins

Evolution does not per se provide proteins with high stability. In fact, stability is just one of many evolutionary constraints on proteins. Proteins have to fold to a

Protein Folding Handbook. Part II. Edited by J. Buchner and T. Kiefhaber

ISBN: 3-527-30784-2

defined structure with adequate yield in a reasonable time and then have to be just stable enough to perform their function over a certain period. There is no evolutionary incentive to make a protein any "better" than what is needed to fulfill its cellular functions. In contrast, the use of a protein in a formulation at high concentration, its prolonged activity at 37 °C, and its large-scale expression and crystallization, just to name a few conditions, may put far higher demands on the protein than its natural environment. Thus, the natural sequence may not be able to provide these properties, but a mutant sequence may.

Proteins exist and have evolved in order to fulfill a given function, and evolution drives the structural properties of a protein mainly towards increased functionality [1]. In fact, most proteins are only marginally stable, with $\Delta G_{folding}$ in the range of -20 to -60 kJ mol^{-1}. It is still a matter of debate whether this marginal stability is actually a "design feature," e.g., to allow degradation at a certain rate, whether it is caused by the selection pressure towards higher functionality that may not be compatible with high stability, or whether it is just a side effect of the lack of selection for high stability. Function is often linked to higher structural flexibility in certain regions of a protein. Lower stability as a result of this higher local structural flexibility might therefore simply represent an adaptation to increased functionality [2]. If this were generally true, however, stability engineering would fail in most cases, as it would not be able to reconcile stability with preserved protein function. An alternative, more optimistic view for the protein engineer is that marginal stability can be interpreted as a result of genetic drift [3]. In other words, lower stability is not intended; but it simply does not matter, provided that function is maintained. Random mutations occurring during evolution are more likely to destabilize the structure of a given parental protein sequence than to stabilize it or be neutral. However, as long as this stability decrease is not sufficient to render the protein nonfunctional, these destabilizing mutations are likely to accumulate in the sequence. This tendency has also been referred to as "sequence entropy" [4]. As a consequence, stability engineering could be interpreted as the art of identifying these unfavorable mutations in order to reestablish a more stable sequence.

The concept discussed above also sets the basis for the consensus approach to stability engineering, which is discussed in Section 39.3.1 Based on the physical principles of protein folding and a structure-based analysis of the interactions between amino acid residues, rational design can give hints as to which residues need to be altered to achieve a desired effect, and this is discussed in Section 39.3.2. The third focus will be set on the methods of directed evolution, which mimic the mechanisms of Darwinian evolution to evolve proteins with enhanced folding and stability properties (see Section 39.4).

39.2.2 Different Kinds of "Stability"

Before discussing in detail different strategies for rendering a protein more stable, it is worth taking a closer look at some basic features associated with protein stability and folding properties. The term "stability" itself is rather vague, and its precise

meaning has often been adapted depending on the problem being addressed. This leads to different definitions of "stability." We will now briefly analyze the definitions and differences between thermodynamic, kinetic, and thermal stability, as well as folding efficiency, which is also sometimes discussed in this context. Even though these properties are interconnected, they are not equivalent. This has important implications for protein engineering, as it is difficult to predict in advance how a given mutation will influence each of these properties. These influences will in fact be different for any protein under investigation according to the free energies of the native and unfolded states and the folding intermediates, as well as the folding pathway and the respective rate constants.

39.2.2.1 Thermodynamic Stability

Thermodynamics describes the global unfolding behavior of a protein. The corresponding *thermodynamic stability* ΔG describes the differences between the free energies of the native (N) and the unfolded (U) states. Importantly, ΔG is an equilibrium property for a reaction involving two or more states. The simplest model of an unfolding reaction is the equilibrium of a two-state unfolding reaction:

$$\mathrm{N} \underset{k_{\mathrm{folding}}}{\overset{k_{\mathrm{unfolding}}}{\rightleftharpoons}} \mathrm{U}$$

where $k_{\mathrm{unfolding}}$ and k_{folding} are the rate constants of the respective unfolding and folding reactions. As a consequence, values for ΔG can be deduced only if the described process is fully reversible and no intermediate states are populated to a significant extent, unless they are explicitly known and measured. ΔG describes to what degree the two states are populated at a given temperature according to

$$\Delta G_{\mathrm{unfolding}} = -RT \ln K \tag{1}$$

where K is the equilibrium constant of the unfolding process. The treatment of experimental data is much simplified in such a model, as stability is affected only by the free energies of the folded and unfolded states. Because ΔG also provides an overall measure of all energetic contributions of interactions occurring upon protein folding, it is a convenient quantity for comparing the energetic effects of single amino acid replacements in a given folded structure. Because of the great scientific interest in deducing the principles of protein folding on a quantitative basis, many studies have been carried out with carefully selected model systems in which these strict requirements that allow the determination of ΔG are fulfilled. Moreover, ΔG represents an intrinsic quantity, at a given temperature in a given buffer, that is independent of the experimental setup and therefore reproducible in any case. The major problem in such studies, however, is confirmation that the observed transition can in fact be described as a fully reversible process, and denaturant-induced transition data alone are often not sufficient to allow reliable judgment. Only if data derived from measurements on the basis of different spectral probes (or better yet, additionally by differential scanning calorimetry, yielding the model-independent enthalpy change of unfolding) agree with each other are

the deduced ΔG values likely to be correct. Nevertheless, many useful conclusions can be drawn about the effects of mutations even if these strict requirements cannot be completely fulfilled in all cases.

Another fact complicates the use of ΔG as a measure for protein stability. Mostly as a result of the fact that the hydrophobic effect is the major driving force of protein folding [5], ΔG itself is a characteristic, curved function of temperature. It is defined as

$$\Delta G(T) = G^U - G^N = \Delta H(T) - T\Delta S(T) = -RT \ln K \tag{2}$$

emphasizing especially that the enthalpy change ΔH is itself a function of temperature. This change of ΔH with T can be described by the heat capacity change

$$\Delta C_p = \left(\frac{\partial \Delta H(T)}{\partial T}\right)_p \tag{3}$$

Although often ignored, the temperature dependence of ΔG should thus not be left out of the account if one is dealing with the stability engineering of proteins. The large heat capacity change upon the transfer of nonpolar solutes to water, which is the basis of the hydrophobic effect, results in a curved function of ΔG versus temperature. By using the definition of ΔC_p, ΔG can be approximated as a function of T, the melting temperature T_m, the enthalpy change at T_m $\Delta H(T_m)$, and ΔC_p.

$$\Delta G(T) = \left(1 - \frac{T}{T_m}\right)\Delta H(T_m) + (T - T_m)\Delta C_p - T\Delta C_p \ln \frac{T}{T_m} \tag{4}$$

If $\Delta G(T)$ is plotted as a function of T, the curve increases at low temperatures and decreases at high temperatures (Figure 39.1). The temperature at which folded and unfolded states are equally populated, and thus $\Delta G = 0$, is called the melting temperature T_m.

The respective curvature of ΔG versus temperature is strongly dependent on the change of heat capacity ΔC_p upon unfolding. A mutation may change ΔC_p, $\Delta H(T_m)$, and T_m in any combination, thereby altering the shape and the position of this curve. Higher thermodynamic stability at a given temperature ($\Delta G(T)$) can thus be achieved, e.g., by an upshift of this curve with constant maximum. Right-shifting of the curve will decrease ΔG at lower temperatures but increase it at higher temperatures, which goes along with an increase of T_m. Flattening of the curve due to a lower change of heat capacity upon unfolding may cause lower thermodynamic stability at most temperatures, even though T_m is increased.

These considerations contain important implications for the analysis of engineered mutants. First, a measured decrease of ΔG at a certain temperature does not necessarily mean that at higher temperatures the thermodynamic stability might not have been increased. Second, by determining the change of ΔG, no conclusions can be drawn about a potential change of the melting temperature T_m. While they are typically related, a low ΔG at a given temperature does not necessar-

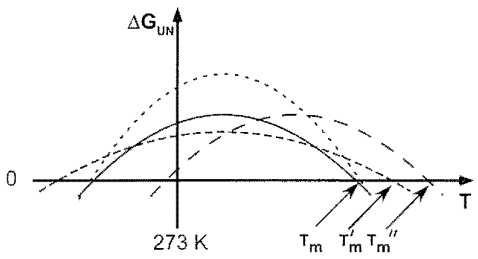

Fig. 39.1. The complex relationship between the melting temperature and the free energy of folding. Schematic representations of the free energy difference between the folded and unfolded states of a protein, ΔG_N, as a function of temperature, T. A typical protein is shown in curve 1 (————). The shape of this curve changes if a mutation affects ΔC_p, $\Delta H(T_m)$, and/or T_m of the protein. In the examples shown, not just one but several of these parameters are changed. Higher thermodynamic stability at a given temperature can phenomenologically be the result of a curve upshift (·····). Right-shifting the curved function (— — — —) results in a right-shift of the maximum of the ΔG_N function as well as in a shift of the melting temperature T_m towards higher temperatures (T_m'). If the ΔG_N function is flattened because of a small ΔC_p (– – – –), a higher melting temperature (T_m'') can result, even though the thermodynamic stability is actually decreased over a broad temperature range. All depicted curves represent extreme cases, and typically a combination of the described alterations will occur upon mutating a protein.

ily mean that T_m will be decreased. T_m and ΔG must therefore be regarded as different properties. It should also be recalled that thermodynamic parameters, while easily reproduced, give no information about how long it will take until equilibrium is reached or, in the case of a nonreversible reaction, until a certain fraction of proteins is inactivated. Thus, thermodynamic stability does not necessarily provide information about whether a protein will meet the stability requirements for an intended application.

T_m itself often serves as a measure of protein stability, more precisely of *thermal stability*. In the literature, the expression "thermal stability" is again used with different meanings. In the described case, it represents the melting temperature for a reversible process. However, a complete thermodynamic analysis of the vast majority of proteins is not possible, because either intermediate folding states are populated or folding in the absence of denaturants that solubilize the unfolded state is not a fully reversible process. As will be discussed in the following section, thermal stability can nevertheless serve as a very practical means of describing protein stability, defined either as the transition temperature of an irreversible process or as the half-life of a protein under a given set of conditions.

39.2.2.2 Kinetic Stability

In most cases, outside the biophysical research lab, protein unfolding is an irreversible reaction. Initially, the aggregation of unfolded molecules or of folding intermediates prevent the back-reaction of folding, and this is followed, after pro-

longed times at high temperature, by chemical inactivation or, in impure samples, proteolysis. In a simple model, kinetic inactivation can be described as

$$N \underset{k_{folding}}{\overset{k_{unfolding}}{\rightleftharpoons}} U \xrightarrow{k_{agg}} X$$

As discussed in the Introduction, one evolutionary constraint on proteins is that their three-dimensional structure remains viable for a certain period of time. In the case of non-equilibrium conditions and irreversible reactions, reaction kinetics becomes the important parameter. The folded state can resist high temperatures for a considerably long time if either the rate constant of inactivation or the rate constant of unfolding is sufficiently low. The rate constants are determined by the free energy of activation, e.g., the difference in energy between the folded state and the transition state of unfolding. Proteins can thus be kinetically stabilized by increasing this activation barrier.

This kinetic stability is different in some crucial aspects from the thermodynamic stability mentioned above. Engineering of proteins for enhanced stability has to deal with both aspects. The best mutations for enhancing kinetic stability will not necessarily be the best mutations for enhancing thermodynamic stability, and not every mutation that increases thermodynamic stability will automatically have a positive effect on protein half-life.

Kinetic stabilization is a common theme in nature, and there are several indications that many proteins from thermophilic organisms are indeed stabilized kinetically rather than thermodynamically [6, 7]. Another important example is that of proteins from the coats of viruses and phages that have to protect their genetic material under very adverse conditions [8]. In extreme cases the native state of a protein can even be less stable than its denatured state, but the native fold can still be kinetically trapped, and large kinetic unfolding barriers can provide the protein with an extremely long half-life [9].

There are different ways of describing and determining kinetic stability. Even if the reaction proceeds in an irreversible way, a practical "melting curve" can still be determined, and the observed transition is cooperative. The midpoint of this transition can serve as a practical means to compare the thermal resistance of different protein variants. Alternatively, one can use the half-life of the protein at a given temperature as an empirical means of stability. However, one has to be aware that this will not reflect equilibrium conditions. The observed values are actually kinetic values and thus are not independent of the exact experimental conditions, such as the protein concentration, the heating rate, etc. A more thorough analysis of kinetic stabilization must include kinetic measurements to determine the respective unfolding rates at different temperatures [10, 11].

For medical applications, engineered proteins usually have to fulfill defined stability requirements such as the absence of aggregation in the formulation used, long-term stability, and prolonged activity at 37 °C. The mere analysis of thermodynamic and kinetic parameters under defined reaction conditions in vitro is not always a reliable indicator of protein behavior under in vivo conditions. Mechanisms other than the intrinsic properties of the protein, such as proteolysis or ag-

gregation with other proteins, can affect half-life. For practical utility, the half-life can therefore also be determined by measuring the percentage of molecules that retain their function after incubation under the respective conditions, e.g., in human serum at 37 °C for several days [12].

39.2.2.3 Folding Efficiency

A different issue of kinetics is related not to unfolding but to the folding of the protein in vivo. More precisely, the question is, which percentage of a protein will actually fold to the native state, as opposed to going to misfolded states, soluble aggregates, or inclusion bodies? Even though the folding efficiency in vivo is not directly related to stability, correlations can often be observed [13]. Additionally, the efficiency of protein folding in vivo is usually the predominant factor influencing the expression yield and is therefore also crucial for large-scale production of functionally intact proteins. Importantly, many mammalian proteins with a high potential for medical applications, especially those secreted or expressed on the cell surface, can rely on the complex folding machinery of the eukaryotic cell to reach their final native state, and the secretory quality-control system of eukaryotic cells allows discrimination of proteins by their folding behavior [14]. Moreover, they usually do not need to be expressed in high amounts in their native physiological context. The resulting lack of selection pressure on their efficiency of folding during evolution is likely to be one of the causes for the difficulties often observed when attempting their overexpression. Unfortunately, these tend to be the proteins of greatest pharmacological utility.

Despite the fact that thermodynamic stability underlies folding efficiency, the kinetic partitioning into productive folding or aggregation is influenced by many different factors. For illustrative purposes, we can again describe this by a very simple scheme:

$$U \xrightarrow{k_{int}} I \begin{cases} \xrightarrow{k_{folding}} N \\ \xrightarrow{k_{agg}} Agg \end{cases}$$

Folding intermediates are often the source of aggregation, and the overall folding efficiency will therefore depend mainly on the nature of these intermediates. This includes the free energy of the folding intermediates themselves, as well as their half-lives and any efficient pathways to aggregation. In vivo, the situation becomes much more complex, as additional parameters such as interactions with cellular components and chaperones or degradation by host cell proteases come into play. The final output of a properly folded protein will therefore depend on all of these kinetic competitions [15]. Protein expression in the bacterial cytoplasm in many cases shows correlations between soluble expression yield and thermodynamic stability of the protein [16, 17]. Additional complications can arise from transport steps. For example, the expression of proteins in the periplasm of *E. coli* is dependent on the prior transport of the polypeptide chain through the inner membrane, and the folding yield is subsequently influenced by the folding and aggregation

reaction in the periplasmic space as well as by interactions with periplasmic factors such as chaperones and proteases. In some cases, mutations that show positive effects on in vivo folding yield have no influence on the overall stability of the protein [18]. Conversely, mutations that strongly increase thermodynamic stability sometimes result in lower folding yields [19]. Nevertheless, many mutations act synergistically on both properties, because they are likely to reduce the free energy of the folded state as well as the free energy of folding intermediates and thereby lower the energetic activation barrier to folding [13].

39.3 The Engineering Approach

39.3.1 Consensus Strategies

39.3.1.1 Principles

As discussed in the preceding sections, marginal protein stability is likely to be a side effect of "sequence entropy" occurring during natural evolution, because the major driving force of evolution is positive selection towards an enhanced functional property, while stability has to be maintained at only a minimum level to secure function. Mutations are likely to occur in a random fashion during this process; the probability that a mutation will have a stabilizing effect on the protein is very low, whereas the probability that the mutation will have destabilizing effects is very high. However, as long as the remaining amino acid sequence is still able to fold into a given structure and the overall domain stability does not fall below a certain threshold, the resulting protein sequences will not be eliminated during the course of evolution [20]. Destabilizing mutations are therefore often selectively neutral and thus accumulate in a given parental sequence. The same should also be true for folding efficiencies. Most proteins are not needed at high concentrations or may even become harmful to the organism in such a case. Similar to stability, folding yield is selectively neutral, provided that the minimum level for cellular function is maintained.

This sets the basis for a semi-rational approach to protein stabilization, which is called the consensus approach [21] and is based on sequence statistics. Because mutations occur randomly, the distribution of amino acids at a given position in a set of homologous proteins can be described, in a very crude approximation, by Boltzmann's law. The consensus approach assumes that at a specific position in a sequence alignment of homologous proteins, the contribution of the respective consensus amino acid to the stability of the protein is on average higher than the contribution of any non-consensus amino acid. Replacement of all non-consensus amino acids in a sequence by the respective consensus amino acid should therefore increase the overall stability of the protein. Obvious advantages of the consensus approach are that it is comparatively simple and is not strictly dependent on structural information at high resolution.

The prerequisite for building a non-biased consensus is the availability of sequences homologous to the protein under investigation. The number of sequences should be large enough to make the sequence statistics reliable and to exclude bias in the resulting consensus sequences. Figure 39.2 shows an alignment of homologous sequences of single repeat modules, the smallest structural entity of a class of proteins known as leucine-rich repeat (LRR) proteins. Because the length of these modules varies among the different classes of LRR proteins – influencing their topology – only repeat modules of a length of 24 amino acid residues have been used for this alignment. The probability of each amino acid occurring at a given position is calculated to derive a consensus sequence, representing the most frequently occurring amino acid residue at each position. The distribution of residues at each position can provide information on structurally forbidden residues and allows weighing the consensus with respect to variability. In most cases, the consensus will contain the residues important for defining the structure of the proteins. In the case of an enzyme family, it will also include the "functional" residues, i.e., those of the active site. In the case of binding proteins, such as antibodies or repeat proteins, the "functional" residues (those involved in binding) are not conserved but are different for each individual molecule, which has to adapt to its target.

Although at first sight no sophisticated structural analysis seems to be required, this is true in only the simplest of cases, where a single family of related sequences can be represented by a single consensus. Frequently, multiple families have emerged that use mutually incompatible solutions of packing. A good example is that of antibodies for which subgroup-specific consensus building has been very fruitful [22]. Averaging over all families would simply yield the consensus of the most-represented family and, if they are equally represented, may result in mutually incompatible residues. Thus, structural analysis can be very helpful in deciding whether an "averaging" of different sequence families is permissible or not. Because of this problem of interacting residues, a simple averaging may lead to incompatible pairs; therefore, these residues should be changed only as groups. The danger of disrupting these interactions by substitutions with consensus residues is especially high in cases where a very broad set of sequences is used for the alignment. To minimize this risk, the sequence statistics can be extended by analysis of covariance in order to derive probabilities that describe the joint occurrence of amino acid residues at two defined positions [23]. As explained above, before deriving a consensus, the aligned sequences can be divided into subclasses, which are likely to contain interacting pairs or groups of residues. Certain variations of residues involved in salt bridges, distinct hydrogen-bonding patterns, or packing of the hydrophobic core are characterized by complementary changes between these subclasses, with mutations to a certain residue at one position being compensated by a mutation to a complementary residue at another position. The subclasses can be built either by being based on sequence homology alone or by including structural information if available. Even though the definition of these subclasses is always dependent on the homology cutoff set by the investigator, a simple dendrogram analysis can be used to group the complete set of sequences into distinct families. Additionally, by building the consensus sequence of each family separately, fol-

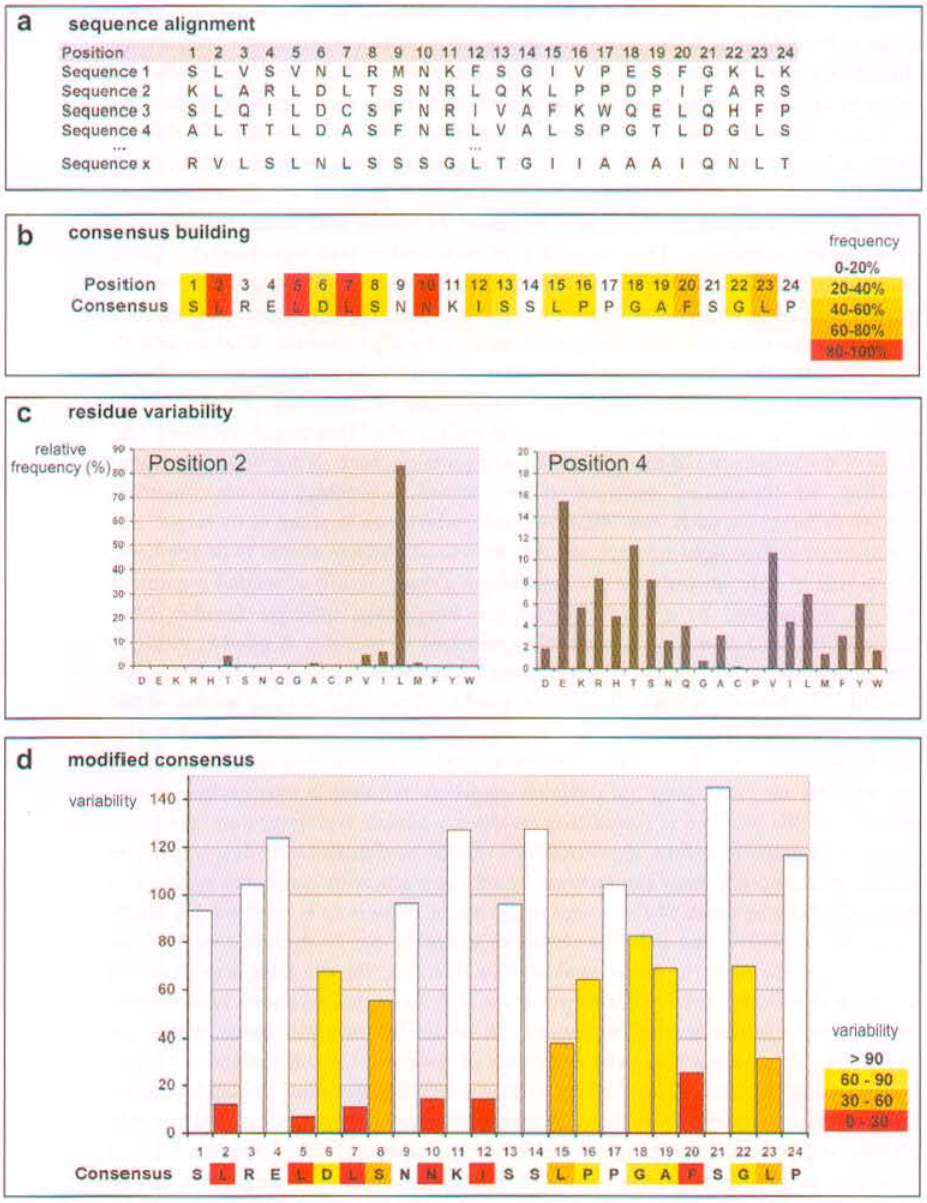

Fig. 39.2. Example of the analysis for deriving and analyzing a consensus sequence. (a) From a sequence alignment of 3077 sequences of 24 amino acid LRR motifs of the Pfam database (http://www.sanger.ac.uk/Software/Pfam/), the relative frequencies of amino acid residues at each position are calculated. (b) Based on the relative frequencies calculated from the alignment, a consensus sequence can be derived representing the most frequently

lowed by a comparison of these consensus sequences, distinct structural features of each group can be recognized in some cases.

The consensus concept has been applied successfully to a large variety of proteins and structural motifs to date. Important lessons for an effective application of the consensus approach can be learned from these studies, and we will therefore give a few examples to briefly discuss some of the advantages and limitations of the method.

39.3.1.2 Examples

There is always the concern that the stabilizing and destabilizing effects of introduced mutations will counterbalance each other and that the overall change in protein stability will be small. Steipe et al. [20] applied the method for the first time on immunoglobulin V_L domains and predicted 10 potentially stabilizing mutations. Six mutations were indeed stabilizing, three had no effect, and one was destabilizing. When applied to GroEL mini-chaperones, 34 predicted amino acid replacements were individually checked, out of which 13 were stabilizing, five showed no effect, and 16 were destabilizing [24]. Lehmann et al. [25] extended the approach to an entire protein. In a first set of experiments, 13 homologous sequences of a fungal phytase were used to build the consensus, and the resulting consensus enzyme showed an increase of 15–22 degrees in unfolding temperature and an increase of the temperature optimum for catalysis of 16–26 degrees compared with each of its parents. In a second set of experiments, additional mutations were predicted by simply adding more sequences to the alignment. By checking the effects of the individual mutations on thermal stability and combining mutations with positive effects, the unfolding temperature could be increased by an additional 21 degrees to 90 °C. No loss of catalytic activity of the enzyme was observed in any case. This work showed that the number of sequences used for the alignment is indeed an

occurring residue at each position. The color code is given on the right hand side of the panel. Residues occurring in more than 80% of all sequences at the respective position are colored in red. Based on these values, consensus sequences with a given homology threshold can be derived (not shown), i.e., with a higher threshold, more positions will be "undefined." (c) Preferences at a given position are reflected not only by the absolute frequency but also by the total number of different amino acid types occurring at each position (residue variability). As an example, the relative frequencies of each amino acid are shown for the highly conserved position 2 and the highly variable position 4. By plotting the relative frequencies of amino acids at a particular position, preferences for certain amino acid types as well as "forbidden" residues can be identified. At position 2, a strong preference for leucine can be observed, and the occurrence of other residues is restricted mostly to hydrophobic side chains. At position 4, all residues are "allowed" except for proline, which is "disallowed" due to secondary structure propensity reasons. (d) By normalizing the relative frequency of the consensus amino acid with the number of "allowed" residues at a given position, a modified consensus sequence can be deduced. The variability V is calculated according to $V = 100 \times N/F$, where N is the total number of different residue types occurring at each position and F is the frequency of the most frequent residue in percentage.

important factor because it might help to optimize ambiguous positions. In addition, the observed large increase in thermal stability could not be attributed to the effect of one single amino acid substitution but rather to the synergistic effects of many replacements.

Even though these and other studies show clearly that the consensus approach allows one to predict stabilizing mutations with a rather high success rate in a rapid way, the effect of each predicted mutation carries some uncertainty, and it is possible that some may contribute destabilizing effects that can counterbalance the stabilizing ones. However, the effects of stabilizing mutations were often found to be additive. Therefore, instead of examining each mutation individually, it is often useful to combine groups of "rather certain" mutations and others that are more speculative.

The application of the consensus concept to families of repeat proteins [23, 26–28] represents, in some respects, a special case due to a number of favorable features. It can nevertheless illustrate the importance of some principles of the approach. The non-globular fold of repeat proteins consists of repeated motifs of 20–40 amino acids. Several results indicate that consensus repeat proteins are indeed much more stable than natural repeat proteins [29, 30]. Repeat proteins might represent an extreme case in which the principles of the consensus approach become very apparent. The structural entities (the repeat modules) are small and thus each protein contributes several repeats to the databases; the number of available sequences consequently becomes very large compared with other proteins. Therefore, a consensus sequence for a single repeat module can easily be assigned and ambiguous positions will occur with lower frequency in the statistical output (Figure 39.2). In addition, interactions that are present within or between several repeats will add to the free energy of folding multiple times, while problem spots would equally be potentiated. Therefore, effects on stability are likely to be consecutively added by introducing additional modules to the array. Even though more-detailed analyses are still needed, initial results pinpoint some of the structural reasons for the stability gains observed upon building a consensus in each of these studies. The regular arrangement of structural motifs gives rise to a more regular H-bonding pattern with a higher number of inter- and intra-repeat H bonds [30]. Loop insertions in natural repeat proteins that are likely to result in more flexible local regions are removed, thereby eliminating local centers of unfolding. In the left-handed helical and disallowed regions of the Ramachandran plot, glycine residues are always present in the consensus proteins, while they are avoided in other places by this design, where their flexibility is not needed and may be harmful to stability.

39.3.2
Structure-based Engineering

Structure-based engineering relies on a detailed analysis of 3D structures, followed by site-directed mutagenesis. We avoid the term "rational" engineering, as it would elicit expectations of perfect predictability and implicitly suggest that all other ap-

proaches are free of logical reasoning. In structure-based engineering, positions have to be identified at which suboptimal amino acids in the original sequences lead to a loss of stability. Subsequently it needs to be specified which amino acids should be introduced as a replacement. The *ab initio* prediction of protein structure, however, is still not a feasible task due to the multitude of potential interactions within the protein and between protein and solvent, which leads to an extremely high number of possible conformations and intermediates of similar energy [31]. Hence, a prerequisite for structure-based engineering is, next to experimentally determined structures, usually the existence of a large experimental dataset within a group of structural homologues that can be used as a basis for predictions. High-resolution structures are necessary to allow the estimation of possible conformational, energetic, and steric influences upon replacement of particular amino acid side chains and can thus help to avoid unfavorable strain in the resulting mutants. Because of the present efforts in the field of structural genomics, these structure-based approaches are likely to become even more important in the future. With this structural information in hand, the goal of designing more stable variants is then to pinpoint particular regions and positions associated with possible stability defects and to subsequently find a better solution to the problem. In contrast, semi-rational approaches like the previously discussed consensus concept are rather crude methods for introducing stabilizing features. Structural and energetic analysis can be used to reexamine the changes proposed by the consensus approach and to fine-tune the system to reintroduce structural features that might have been lost in the averaging process.

Proteins of hyperthermophilic organisms have been of special interest for examining the structural mechanisms of thermostabilization and have been contemplated as guides for the engineering of "problem" proteins for better properties. From a phenomenological point of view, the basis of increased thermostability is frequently set by a flattened ΔG-versus-T curve that is due to a smaller change of heat capacity upon unfolding (Figure 39.1) or by a kinetic stabilization that is due to a strong decrease in the rate of unfolding. The crucial question is, however, what the molecular differences are that give these proteins their favorable properties. Genome-wide comparisons between hyperthermophilic and mesophilic organisms with respect to amino acid composition did not yield any obvious common rules of how these effects are achieved [32]. Therefore, hope was placed on the increasing number of pairwise high-resolution structure comparisons of thermophilic proteins with their mesophilic counterparts. While they provided a more differentiated picture, a "global" rule still could not be derived. A few highly specific mutations are often enough to provide considerably stabilizing effects, but the additive effect of many small contributions, none of them dramatic by itself, may be the usual case. Moreover, rather than relying on one universal strategy, nature utilizes a variety of strategies for the thermal adaptation of proteins [33]. In fact, the list of stabilizing structural features in hyperthermophilic proteins reflects the diverse principles of protein stability and folding that protein engineers try to exploit and that will be discussed in this section. High-resolution structures of thermophilic proteins can thus provide a detailed view of how nature implements these principles

to create proteins of higher stability [34]. However, the lack of a unifying "rule" and the multitude of strategies nature uses provide an important lesson for protein engineers. Depending on the protein under investigation, the strategy of choice can be different, and even for a given protein there may be more than one optimal solution to the problem. Before choosing from the available set of strategies, the focus should therefore be on identifying potential "weak points" responsible for stability defects in a given protein structure.

We will now discuss some structural features associated with protein stability as well as strategies for altering these features towards more favorable biophysical properties. The given list of course lays no claim to completeness but should point out some important principles. Any replacement in a given sequence may have multiple effects on protein stability, and destabilizing effects can often outweigh the stabilizing ones. An assessment of potential destabilizing effects is therefore crucial. Wherever possible, references are made to the different forms of "stability" discussed in the Introduction. In the case study provided at the end of this section, an example will be given to demonstrate how consensus approaches and structural analysis can be combined to yield useful results.

39.3.2.1 Entropic Stabilization

An obvious strategy for increasing the free-energy difference between the folded and unfolded states, and thus the thermodynamic stability, is to decrease the entropy of the unfolded state. The underlying concept is to decrease the flexibility of the polypeptide chain, usually by introducing an additional intrachain linkage. Such entropic stabilization has become a common strategy for protein engineers. The prerequisite for success is that the mutations rendering the unfolded protein less flexible do not introduce unfavorable strain in the folded three-dimensional structure or result in any steric incompatibilities [35]. We now discuss several ways to achieve an entropic stabilization.

Introduction of Disulfide Bridges The introduction of additional disulfide bridges is a straightforward way of establishing an intrachain linkage to reduce the entropy of the unfolded state [36, 37]. The magnitude of the entropic effect is thought, as a crude approximation, to be proportional to the logarithm of the number of residues between the two bridged cysteines [38]. The spatial distance between the residues to be replaced with cysteines has to be evaluated with care in the model of the folded protein in order to prevent perturbations of the native structure upon formation of the disulfide bridge. It should be noted, however, that the energetic effect of additional disulfide bridges is not only entropic but also of a far more complex nature, giving rise to entropic as well as enthalpic contributions to the change in the free energy of folding [39, 40]. For example, an additional decrease in the free energy of folding can result from the reduced solvation energy of the unfolded state [40]. In contrast, a reduced solvation energy of the folded state would have the opposite effect, while residual structure in the denatured state would again push the equilibrium to the side of the folded protein. Because the disulfide bond itself is hydrophobic in nature, it is often engineered into the inte-

rior of the protein. This is not an easy task, as it can negatively affect core packing. Even though there are several examples of successful protein stabilization by introducing artificial disulfide bridges [41, 42], the complex energetic effects can also cause a destabilization of the protein [43]. Furthermore, the introduction of additional cysteines often results in a rather drastic decrease in folding efficiency, because incorrect and intermolecular disulfide formation can remove large portions of expressed protein by aggregation. Since disulfide formation does not occur in the cytoplasm, secretion to the bacterial periplasm or to the eukaryotic ER is required for functional expression, usually associated with lower yields than for the production of cytoplasmic proteins. Alternatively, if the protein is produced by refolding, redox conditions have to be adjusted, which can be difficult if the native protein also contains free cysteines.

Circularization An alternative approach with the same underlying concept is the circularization of proteins by fixing the loose N- and C-termini via a peptide bond. In addition to the entropic effect, the fixing of the loose ends can prevent local unfolding events occurring at the termini and thereby kinetically stabilize the native structure. With the discovery of inteins, which mediate protein-splicing reactions, a tool that allows the directed formation of peptide bonds between ends fused to different parts of the intein became available. Intein-mediated protein ligation has been used to covalently link the termini of β-lactamase, a protein that is especially amenable to this strategy due to the close proximity of its N- and C-termini [44]. In accordance with polymer theory, the thermal stability of the protein was enhanced by about 6 degrees, from 45 °C to 51 °C. For circularized DHFR [45] an increased half-life at elevated temperature was observed. The close proximity of the termini is of course a prerequisite for this procedure, and the stabilizing effect is likely to become marginal if the loose ends are linked via long unstructured loops. Similar to the situation upon introduction of artificial disulfide bridges, destabilizing enthalpic effects may negate the favorable entropic contribution [46]. In addition, low protein ligation efficiencies and difficulties in separating circular from linear forms of the protein often cause additional technical challenges. It remains to be seen whether this technology is robust enough for biotechnological or biomedical applications.

Shortening Solvent-exposed Loops Short, solvent-exposed loops are rather fixed in the native state, but a comparably large number of additional conformations become accessible in the unfolded state, while long loops have a large number of conformations also available in the native state. Thus, the shortening of loops should in principle lead to a relative decrease in the loss of conformational entropy upon folding. Conversely, increasing the loop lengths by insertion of glycine residues into the loops of the four-helix bundle protein Rop has indeed resulted in a strong and continuous decrease in thermodynamic stability [47]. In addition, loop shortening can have the effect of abolishing hot spots of local unfolding events and may result in kinetic stabilization. Even though it has become obvious that loop shortening or tying down of loops by external interactions is a common theme in

thermostable proteins of thermophilic organisms [48, 49], the strategy is often hard to realize for a given protein target, as the danger of introducing additional strain in the native state is high, and solvent-exposed loops are often important with respect to function.

Reduction of Chain Entropies By considering the conformational entropies of amino acid side chains, another strategy for decreasing the entropy of the unfolded state becomes apparent. Because of the five-membered-ring nature of the proline side chain, it not only restricts the possible conformations of the preceding residue but also can adopt only a few conformations itself. It therefore has the lowest conformational entropy of all amino acids [50]. In contrast, glycine, which has no side chain, has the highest conformational entropy. Substitutions of non-glycine residues with proline or the replacement of glycines by other residues should therefore reduce the entropy of the unfolded state.

Positions that allow substitutions with proline are, however, very rare. Because proline is poorly compatible with α-helices and incompatible with β-strands, the position of a new proline must not be part of these secondary structure elements. At most positions in the native structure, the respective torsion angles will be incompatible, and the mutations are thus very much restricted to loop and turn regions. Again, care should be taken not to remove any favorable interactions of the replaced amino acid side chain [51]. In order to examine in advance whether the respective site is permissive for a substitution with proline, the dihedral angles of the site can be checked and should lie in the range of ϕ/ψ -50 to $-80/120$ to 180 or, alternatively, -50 to $-70/-10$ to -50.

Similar restrictions apply for the replacement of glycines with any other residue. In many cases this will create steric overlaps, and such negative structural crowding effects can outweigh the positive energetic benefits.

39.3.2.2 Hydrophobic Core Packing

Exposed residues are often directly involved in ligand or substrate binding and therefore often play a functional role. In contrast, the residues of a protein's interior usually play mostly a structural role, and the associated hydrophobic effect is thought to be the main driving force of protein folding and thermodynamic stability.

In known structures the core residues fill almost the entire interior space, provide many favorable van der Waals interactions, and maximize hydrophobic stabilization by exclusion of the solvent. In principle, an increase in thermodynamic stability of 4–8 kJ mol^{-1} can be achieved for each additional methylene group buried [52]. Paradoxically, the importance of the hydrophobic effect for folding and stability of proteins simultaneously limits its applicability for protein engineering. Because the hydrophobic core is already densely packed in almost all native proteins, most changes here will create over-packing or packing defects, causing an overall destabilization rather than an improvement in stability. In addition, even subtle changes of core residues can lead to a rearrangement of external residues and thereby alter the functional properties of the protein [53]. Care should also be taken to avoid the introduction of conformational strain by the mutation of core

residues, as destabilizing effects from a strained conformation can sometimes compensate the energetic gain of an increase in buried hydrophobic volume [54]. Improvement of core packing must therefore be based on analyses made from high-resolution structural information in combination with sequence comparisons. This allows one to specifically look for cavities in the core that indicate imperfect packing. If the hydrophobic surface area around the cavity is large, additional van der Waals interactions can be provided by the introduction of sterically fitting alkyl or aryl groups from hydrophobic side chains, thereby decreasing the size of the cavity [55, 56].

39.3.2.3 Charge Interactions

Oppositely charged amino acid residues, if appropriately positioned, have the potential to form salt bridges, whereas like-charged residues lead to repulsions. The magnitude of the effect of charge-charge interactions on overall protein stability is still a matter of discussion [57]. In the case of ionic interactions between side chains buried in the hydrophobic core, the high energetic cost of transferring charged ions from aqueous solution to the low-dielectric interior of the protein also has to be taken into account. If a single charge were buried in a protein, which would be extremely rare in a natural protein, the design of an ion pair would be very attractive. However, if a hydrophobic pair were to be replaced by an ion pair, the resulting energy would have to be higher than the loss of the previous pair plus the cost of burying the charge. Nevertheless, the high contribution of buried salt bridges to the overall stability of the native protein structure underlines their potential for introducing additional stability [58]. The optimal spatial arrangement of the interacting side chains and their respective charges is, however, crucial. Moreover, buried charged side chains are often not only part of interacting charge pairs but also part of complex charge clusters built from many side chains, which are able to magnify the effect.

Similar rules apply for the interactions of charged residues on the protein surface. Only the perfect arrangement of charges seems to be able to make up for the desolvation penalty that has to be paid upon formation of a salt bridge. The effect on thermal stability, however, can be drastic. Increasing the free energy of unfolding by the changing of charges includes maximizing the number of salt bridges and, equally important, the removal of repulsive interactions [59], which are not uncommon in natural proteins. Predictions on a structural basis can be difficult due to the often higher flexibility of side chains on the protein surface, but simple models that allow predictions about potential stabilizing and destabilizing surface charges can be used [60]. The key is to consider not only nearest neighbors but also a whole network of charges that have to optimally interact and avoid repulsions. Because the surface charge distribution can have a huge impact on stability, but is defined by many residues at different positions, these residues are also a valuable target to be combined with selection techniques as discussed in Section 39.4.

A special case of electrostatic interaction is the "helix dipole." By reducing the net partial charges at the helical ends through placement of side-chains, which pro-

ductively interact with the helix dipole, the helical structure is stabilized [61]. Introduction of negatively charged residues at the N-terminal end and positively charged residues at the C-terminal end leads to this stabilizing effect. The provided stability gain is, however, marginal (less than 4 kJ mol^{-1}).

39.3.2.4 Hydrogen Bonding

There was no initial reason to believe that intrachain hydrogen bonds in the native state would be more energetic than those of the unfolded chain to water [5]. By including terms of entropy change of the solvent and additional van der Waals interactions upon polar group burial, however, the positive contribution of hydrogen bonding to protein stability has become generally accepted [58, 62]. Despite this ongoing discussion, hydrogen-bonding patterns are a highly valuable target for the stability engineering of proteins. Because engineering deals with improving folded proteins, the major concern has to be how to satisfy the existing hydrogen-bonding network in a structural context. The basis for this endeavor is structural information of high resolution, and the most lucrative goal is to identify potential residues that represent buried but unsatisfied donors of hydrogen bonds. Site-directed mutagenesis of a nearby residue to provide a hydrogen-bonding acceptor can cause a stability gain in the range of 2–10 kJ mol^{-1}, depending on the geometry and other compensating effects [63, 64].

A special structural context that can provide significant additional stabilization by either hydrogen bonds or ionic interactions is the anchoring of relatively loose structural elements like loop structures or the N- and C-termini, thereby tightening "hot spots" of local unfolding [49, 65].

39.3.2.5 Disallowed Phi-Psi Angles

The stereochemistry of the polypeptide backbone can be defined by the dihedral angles ϕ and ψ, and any individual residue in a structure is defined by a single set of ϕ, ψ values. For conformational analysis of protein structures, the Ramachandran plot representing the dihedral angle space is an excellent starting point [66]. The ϕ, ψ values of amino acid residues in protein structures usually reside in three preferred or "allowed" regions of the Ramachandran plot, called right-handed helical, extended, and left-handed helical. The right-handed and extended conformations correspond to α-helix and β-strand secondary structures, respectively, and the vast majority of non-glycine residues lie within these two regions. The left-handed region corresponds to structural features at the termini of secondary structure elements and describes regions involved in the reversal of the polypeptide chain. There is a high preference in this region for glycine residues, as the β-carbons of non-glycine residues can sterically interact with the polypeptide backbone, resulting in unfavorable energies. In some cases, the substitution of non-glycine residues by glycines in the left-handed helical region increases thermodynamic stability (up to 8 kJ mol^{-1} in RNase H) [67]. Although the strict introduction of glycines in such positions is an important point to consider for de novo protein design, it does not represent a general rule for stabilizing the native states of proteins. In

some cases, the energy penalty for the accommodation of unfavorable strain can be offset by lost unfavorable or new favorable local interactions, such as hydrogen bonding or hydrophobic interactions. Some replacements of this kind can therefore even lead to destabilization rather than stabilization [68].

The same rules apply in principle to the disallowed regions of the Ramachandran plot. Steric clashes result in a high energetic cost in the folded structure, especially for non-glycine residues in all disallowed regions. Stabilizing mutations to glycine have been introduced with energy gains of up to 18 kJ mol^{-1} [69]. It has been noted, however, that certain non-glycine residues also have propensities to occur in the disallowed regions – such as Asn and Ala in the type II′ turn region – and the energetic cost of their occurrences is often low [70]. Other residues found in the disallowed regions are small polar residues [71] that compensate for the energy cost by making additional hydrogen bonds. Unfavorable conformations often occur in very short loops [72], where the rest of the structure may constrain the loop efficiently, and interactions with the solvent may also offset the energy costs.

Conformational analysis by the Ramachandran plot therefore provides a convenient and fast way to assess possible conformational strain in the tertiary structure associated with particular target residues. However, a close inspection of possible side chain interactions is required, and the analysis should be extended to identify potential compensating features of the residues to be replaced.

39.3.2.6 Local Secondary Structure Propensities

Effects on the overall stability of a protein can also be influenced by the respective secondary structure propensities of amino acid residues for the α-helical and β-sheet conformations. However, these effects are usually marginal. Nevertheless, even at the expense of other favorable trends, such as avoiding the exposure of hydrophobic side chains to the solvent, a given residue is often favored at a certain position due to its secondary structure propensity [73]. If a particular secondary structure element does not form efficiently, many interactions between this element and other parts of the protein can be lost. The energetic consequences and the influences on folding efficiency are, however, hard to predict at this point.

39.3.2.7 Exposed Hydrophobic Side Chains

The removal of exposed hydrophobic side chains increases the polar surface of a protein. Such mutations are not likely to affect thermodynamic stability but presumably do affect the folding efficiency by influencing the rate of aggregation of intermediates. Interestingly, they also do not affect the solubility per se (the amount of native protein that can be dissolved in buffer) and seem to act mostly on folding intermediates [74]. Because of lateral interactions, e.g., between neighboring loops, partially exposed hydrophobic amino acids can even increase stability. It must be kept in mind however, that not only the aggregation pathway, but also the stabilities of folding intermediates are an important parameter for kinetic stability, which can possibly be influenced by the existence of exposed hydrophobic side chains. Moreover, hydrophobic cavities at the protein surface are often in-

volved in specific binding functions of the protein, and the removal of these "functional hot spots" has to be avoided. Therefore, the hydrophobicity of the protein surface must be carefully balanced.

39.3.2.8 Inter-domain Interactions

In proteins consisting of more than one domain, additional principles come into play. The overall stability not only reflects the intrinsic stabilities of the single domains but is also influenced by the stability of the interface of the domains in cases where they are interacting with each other. Stabilization of this interface can be achieved mainly by increasing and optimizing the hydrophobic surface area of the interface. Because hydrophobic side chains at the interface are usually exposed during folding and transient opening of the domain interface, a tradeoff between interface stability and folding efficiency is often observed. Additional dramatic effects on stability can be observed in cases where the two domains exhibit very different intrinsic stabilities [75]. In such a scenario, the unfolding of the protein and the loss of function are strongly related to the unfolding of the less stable domain [76]. An important aspect of kinetic stabilization can be observed in two-domain proteins, where one domain can slow down the unfolding, and therefore the aggregation, of the other. Thus, the native state becomes kinetically stabilized in the assembly and a stable domain interface reduces the extent of its transient openings and, thereby, the resulting exposure of hydrophobic patches that would favor aggregation. Alternatively, covalent cross-links (e.g., disulfide bonds) between the interfaces of multimeric proteins can be introduced. For example, the introduction of disulfide bonds between the interfaces of *Lactobacillus* thymidylate synthase not only increases their thermal stability but also leads to reversible thermal unfolding [77].

39.3.3 Case Study: Combining Consensus Design and Rational Engineering to Yield Antibodies with Favorable Biophysical Properties

The following example illustrates how consensus approaches, rational design, and experimental data can be combined in a synergistic fashion to iteratively optimize biophysical properties.

The smallest form of an antibody able to bind the antigen in the same manner as the whole IgG consists of two domains, the variable domain of the heavy chain (V_H) and the variable domain of the light chain (V_L). Both domains interact with each other via an interfacial region of highly hydrophobic character. In single-chain Fv (scFv) antibody fragments, the two domains are covalently linked by a flexible linker region of typically 15–20 amino acids (Figure 39.3a). The binding site for the antigen usually involves three loop regions in each of the domains, named complementarity-determining regions (CDRs). Antibodies that are based on human antibody sequences possess great potential for many medical applications, either directly as an antibody fragment or after reconstructing an IgG [78]. Antibody fragments can be expressed in a convenient manner in *E. coli*, thereby providing

rapid access to these proteins [79]. Nevertheless, there are often drastic differences between individual antibodies concerning their expression yield and their stabilities. Ideally, the recombinant antibodies would all provide favorable biophysical properties.

The starting point for such recombinant antibodies is a library. One fully synthetic library of this kind, the human combinatorial antibody library (HuCAL), was designed based on the consensus concept [22]. Based on human antibody germ-line sequences, several sequence families were created. Importantly, instead of averaging over all possible human antibody sequences, the consensus sequences were built for each family separately, resulting in seven consensus frameworks for V_H (V_H1a, V_H1b, V_H2, V_H3, V_H4, V_H5, and V_H6) and seven frameworks for V_L ($V_\kappa1$, $V_\kappa2$, $V_\kappa3$, $V_\kappa4$, $V_\lambda1$, $V_\lambda2$, and $V_\lambda3$). This diversity is important, as the use of different frameworks allows a variety of non-CDR contacts to the target, thereby greatly increasing the range of targets being recognized. By using this strategy, each human V_H and V_L subfamily that is frequently used during an immune response is represented by one consensus framework, and thus the immune response is closely mimicked. The consensus building was further restricted to the framework regions, while the CDRs were diversified in a manner guided by structure. Thereby, functional diversity is maintained in optimized sets of frameworks.

Instead of fully relying on the statistical output of the consensus building, structural modeling was employed in order to decrease the risk of disrupting interactions between certain amino acid residues that might be in contact with each other in the three-dimensional structure. Because many structures of antibody domains are available, the modeled structures of each framework family could be compared with the respective natural template structures. The models were checked according to several principles, which have been outlined in the preceding section. At this point the models were mostly checked for whether the interactions in natural antibody domains were correctly recreated. The consensus sequence models were inspected for any unfavorable strain in the structures, represented by unfavorable regions of the Ramachandran plot or any obvious cavities in the hydrophobic core. Moreover, the sequences were checked for the existence of residues already known to be involved in conserved intra-domain interaction patterns – such as conserved charge clusters and hydrogen-bonding patterns – as well as for exposed hydrophobic side chains known to decrease expression levels. Nevertheless, differences between the natural subclasses became apparent in the models.

Empirical results had suggested that certain natural framework types display more favorable stabilities and expression levels than others. These differences are already existing in the original natural human germ-line sequences. The intrinsic differences in terms of biophysical properties for each subtype were then experimentally explored in a systematic fashion, first on single domains, then on scFv fragments [13]. The experiments confirmed the observed trends, showing that V_H3 displays the highest thermodynamic stability and soluble expression level, when expressed as an individual domain, among all V_H subtypes. In contrast, V_H2, V_H4, and V_H6 display the least favorable properties in terms of stability, folding yield, and the tendency to aggregate. For the V_L domains, members of the V_κ

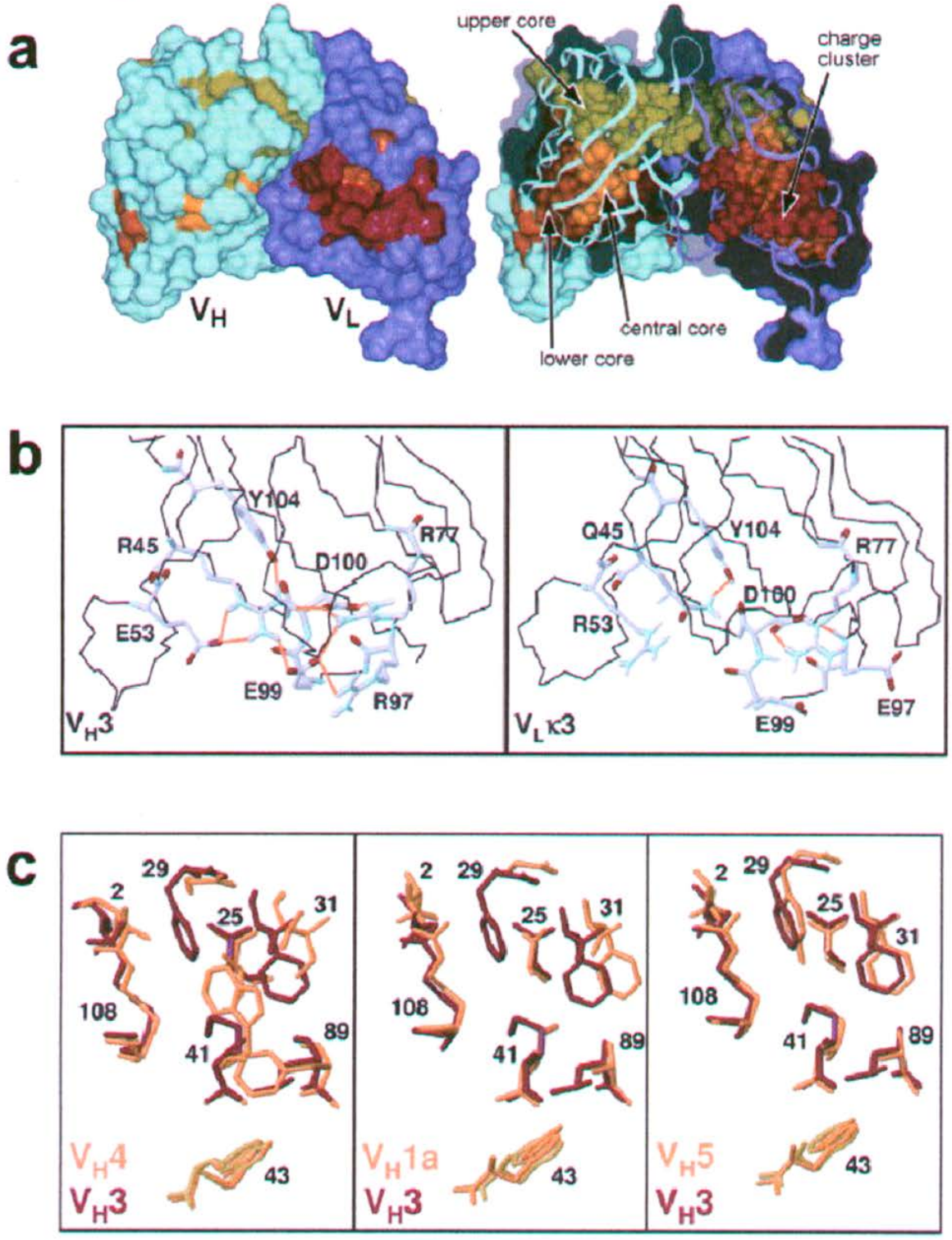

Fig. 39.3. (a) Structure of an antibody Fv fragment consisting of the variable heavy chain (V_H) and the variable light chain (V_L). Each domain V_H and V_L is characterized by three hydrophobic core regions (upper [green], central [yellow], and lower [orange] core) and a charge cluster at the base of each domain (red). Even though the residues defining these

subtypes showed slightly higher stabilities and expression yields than the V_λ subtypes, but the behavior was much more homogenous overall. In order to trace back these differences to the structural level, the model structures of the different subtypes were compared with each other. Several structural features can be invoked to explain the extraordinary stability of V_H3 domains in comparison with the even-numbered V_H subtypes.

First, differences in the hydrogen-bonding networks have an influence on the thermodynamic stabilities. Long-range interactions involving several residues are concentrated in a charge cluster at the base of V_H domains to establish a complex interaction network. In V_H3 domains the ionic and hydrogen-bonding interactions within this charge cluster are very well satisfied, whereas fewer interactions are observed at the analogous positions in other subtypes (Figure 39.3b). Corresponding to the subclass, different hydrogen-bonding networks are formed in the charge cluster of V_H domains. Some of these networks are less extended and contain a smaller number of interactions than in V_H3.

Additionally, based on the residues at three different positions in the first β-strand of the V_H domains, the domains can be classified into four different structural subtypes with respect to their conformations in the first framework region [80]. Mutations bringing together incompatible residues and thus "mixing" subtypes have previously been shown to have a large unfavorable effect on the stability of the whole scFv [81].

Also, clear differences can be observed for hydrophobic core packing of the family subtypes. The upper core region of V_H3 is densely packed, whereas cavities can be identified in V_H4, V_H1a, and V_H5 on the basis of structural alignments (Figure 39.3c). In the lower core, two of the stable domains have an aromatic residue, while the others do not.

Finally, a comparison of the Ramachandran plots showed additional non-glycine residues with positive ϕ-angles and residues with higher secondary structure pro-

regions are conserved within the same germline family, these sequence motifs differ between different germ-line families. (b) Arrangement of the residues defining the charge cluster of V_H3 and $V_\kappa3$ domains [13]. Importantly, the charge cluster consists of a network of buried charges and hydrogen bonds rather than pairwise interactions between individual residues. (c) Furthermore, subtype-dependent packing differences occur. As an example, the residues defining the upper hydrophobic core region of different human V_H subtypes are shown as structural superpositions. Structural alignments are shown of V_H4 (PDB entry 1DHV), V_H1a (1DHA), and V_H5 (1DHW) [22], with the most stable framework, V_H3 (1DHU) (left panel). While the upper core residues of V_H3 are densely packed, cavities occur for the other subtypes. In the least stable subtypes, V_H4 and V_H6, the bulky aromatic residues Phe29 and Phe31 are replaced by smaller residues, and the created space is only partly filled up by compensating residues Trp41 and Val25. The loss of the phenyl ring by replacement with Gly in V_H1a (middle panel) as well as the substitution of Leu89 by Ala are not compensated for by larger side chains at other positions, thereby creating hydrophobic cavities. The same is true for position 89 in V_H5 (right panel). Adapted from Ewert et al. [13].

pensity at certain positions for the even-numbered subtypes compared with the odd-numbered ones.

The immediate question was, therefore, whether the results of this structural trouble-shooting could be used to project favorable properties of V_H3 domains onto the less stable subtypes and thereby add another step to the optimization of antibody sequences while maintaining the structural diversity of the immune system. Instead of using the stable V_H3 framework exclusively, resulting in a loss of diversity, some point mutations might be enough to correct some of the shortcomings of the less stable domains.

Based on these comparisons, the mutation of six residues in a scFv containing a V_H6 framework led to an overall increase in thermodynamic stability of 20 kJ mol^{-1} and a fourfold increase in soluble expression yield [73], indeed bringing this framework to the level of V_H3. The effects of the single mutations on stability were almost fully additive, while the effects of folding efficiency (soluble expression yield) were only qualitatively additive. The most dramatic effects on thermodynamic stability were obtained by mutations removing an unsatisfied H-bond donor in the hydrophobic core and introducing glycines at positions with positive ϕ-angles. Individually, all mutations, except the one in the hydrophobic core, led to slight increases in soluble expression yield. Interestingly, one mutation that removed a hydrophilic, solvent-exposed glutamine residue by a replacement with hydrophobic valine on the basis of higher secondary structure propensity also significantly increased the expression yield, possibly since valine secures this stretch to be in β-sheet structure.

In antibodies, disulfide bond engineering has also been investigated. Optimized disulfide bonds engineered between V_H and V_L indeed significantly increased the half-life of an Fv fragment at 37 °C [82]. This strategy was subsequently extended to interchain disulfides in generic framework positions [83]. Even more stable proteins can be obtained by combining the single-chain Fv approach with the engineered disulfide [84, 85]. It should be pointed out, however, that the additional inter-domain disulfide significantly reduces the yield of folded protein when produced in the bacterial periplasm, and such proteins have to be prepared by in vitro refolding. Therefore, the additional disulfides are not ideal for antibody libraries; instead, optimized frameworks have shown the greatest promise for combining diversity with stability.

In summary, the consensus concept provides a convenient tool for proceeding with large steps in sequence space and a high probability of accumulating features that are favorable for higher stability. In many cases, structural analysis can serve as a trouble-shooting tool to identify shortcomings that might have been created by the consensus-building process or, as in the case of the antibodies described here, that are inherent to the natural sequence. In addition, it serves to rank the potential mutations identified by the consensus approach, keeping the mutational load on the target molecule to a minimum. Experimental data are important not only to validate these results but also to give important hints for future design approaches. Many entries in the table of experimental antibody stability data linked to mutations have come from directed evolution experiments. By combining

semi-rational and rational approaches with experimental data, optimization of biophysical properties can be achieved in an iterative fashion. This interplay of experiment and structural analysis can therefore be an effective way to probe the vast sequence space in a systematic manner in order to find the valleys of free energy.

39.4 The Selection and Evolution Approach

39.4.1 Principles

The previous section may have led to the impression that mutations enhancing the biophysical properties of proteins can rapidly be identified. In the highlighted case of antibody domains, only the wealth of structural data and empirical measurements available allowed predictions with a high probability of success. Some of the empirical data used successfully for structure-based engineering have come, paradoxically, from directed evolution experiments (see Section 39.3.3). Despite the rapidly increasing amount of structural data and the better understanding of folding mechanisms of proteins, the effects of introduced mutations still cannot be predicted with a high degree of accuracy in most cases. The main reason is that even slight alterations in the primary sequence can lead to profound conformational changes in tertiary or quaternary structure; consequently, structural predictions have to be very accurate and usually must be backed up by empirical data.

Because rational engineering uses site-directed mutagenesis followed by biophysical investigation to probe the effects of specific amino acid substitutions, it becomes very labor-intensive if many mutants have to be checked individually and if no additional hints are available. Additionally, as soon as small synergistic effects of several mutations need to be checked, the combinatorial explosion rapidly exceeds the sample number that can be handled efficiently. More importantly, the restriction to certain target residues automatically excludes alternative solutions to a given problem that may not have been obvious by the initial analysis. For example, affinity improvement of a protein to its binding partner is often achieved more efficiently by slight spatial adjustments of residues directly involved in binding, rather than by substitution of these residues. This kind of spatial adjustment can be caused by mutation of residues whose location in the native structure is further away from the actual binding site (so-called "second-sphere mutations") [86]. Today, it is almost impossible to predict this kind of mutation by rational means.

When it comes to stability engineering of proteins, the problems associated with rational design procedures are even intensified. First, because the energetic contributions of single-site mutations are usually small, the need to sample multiple mutants – in which synergistic effects of several mutations are combined – is more acute. Second, the factors and principles responsible for the overall stability of a native protein are still far from being completely understood, and a multitude of different forces and interactions contribute to it. A complete analysis will have to

consider not only the interaction network of the native protein but also the effects on the denatured state. Additionally, potential aggregation pathways will have to be considered. Even simple substitutions like the ones discussed in the previous sections often make contributions to the entropy as well as the enthalpy of the folded and unfolded states, including the solvent in either state. Third, because rational approaches always rely on the current theoretical knowledge, new principles underlying protein stability will rarely be uncovered. In any case, rational engineering requires a clear definition of the problem by the investigator in order to find a solution. Ironically, in the field of protein engineering, the exact definition of the problem is often the problem itself.

Thus, to overcome these limitations, an experimental setup is needed that allows the creation of a vast number of variants of a given protein and that subsequently can identify "superior" molecules that best fulfill a desired property. Nature samples the vast sequence space by the strategy of Darwinian evolution, a cyclic iteration of randomization and selection. Nature thereby adapts proteins to fulfill a function under the given environmental conditions. Recent developments in molecular biology have made it possible to mimic Darwinian evolution in a reasonable time in vitro. Not only has this "evolutionary approach" become the most powerful method to date for engineering proteins towards a desired property, but it also provides new insights into the mechanisms and principles that are responsible for this property. However, as has been illustrated in the case study on antibodies, such experiments can be used not only to solve a particular problem but also to gather information about which residues tend to become enriched in particular positions. This again provides a database for rational engineering.

Many different selection and evolution strategies have been developed in recent years, but all of them have several features in common that reflect the principles of Darwinian evolution. The starting point is the generation of a genetic library of mutants derived from the wild-type sequence of the protein under examination. Several methods exist to create sequence diversity. In error-prone PCR, the error rate of polymerases is increased by performing the PCR reaction in the presence of deoxynucleotide analogues or in the presence of other metal ions. By using bacterial mutator strains, which are characterized by deficient DNA repair systems, random mutations are introduced during DNA amplification in the bacterial cell, albeit usually at a lower frequency [87]. In DNA shuffling, which mimics the natural process of sexual recombination, genes are randomly fragmented by nuclease digestion and reassembled by a PCR reaction in which homologous fragments act as primers for each other [88] (Figure 39.4). The staggered extension process [89] is another possibility to obtain recombined genes in vitro. Here, the polymerase-catalyzed extension of template sequences is extremely abbreviated, and repeated cycles of denaturation and extension lead to several template switches – thereby recombining elements from different genes – before the extension finally yields full-length products. Additionally, techniques are available to focus the diversity to certain regions on the whole gene, such as the use of degenerate primers in PCR or the "doping" of a shuffling reaction with degenerate primers. Depending on the problem and the sequence under investigation, each of these methods has its

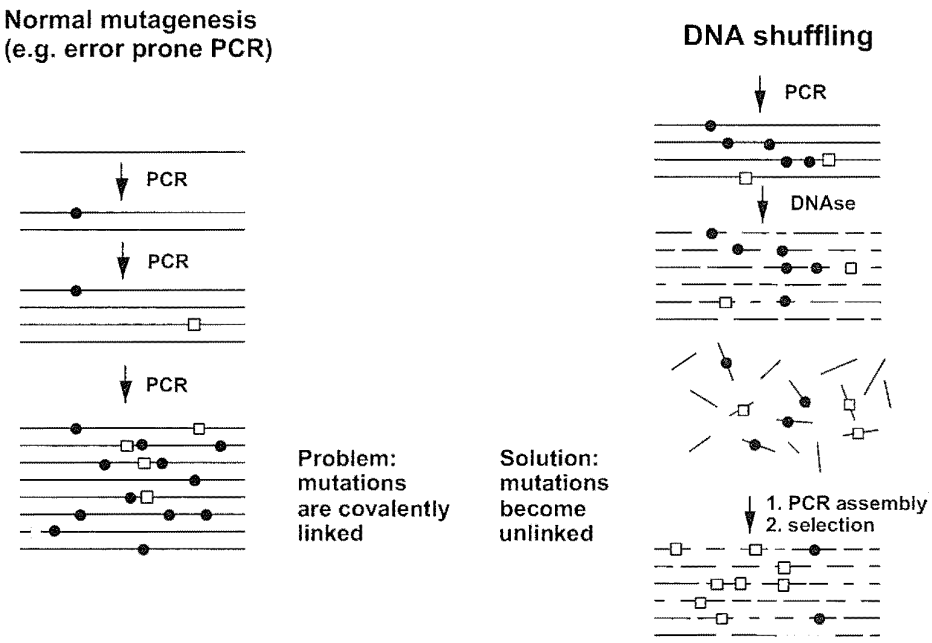

Fig. 39.4. Methods to create genetic diversity by biochemical means. On the left, error-prone PCR (see text) is depicted schematically. Two types of mutations are shown: favorable ones (open squares) and unfavorable ones (closed circles). In successive cycles of PCR, more mutations of each are introduced, and usually molecules will contain some of either type. Thus, the beneficial effect of the favorable mutations can be completely obscured by the presence of unfavorable ones, if the error rate is too high. Therefore, this method is most successful if it is not used at too high an error rate. On the right, DNA shuffling according to Stemmer [88] is shown. A short DNAse digestion breaks up the DNA into small pieces, and PCR is used to reassemble the gene. Thereby, mutations are "crossed" and genes with largely favorable mutations can be obtained that can be enriched by selection. Nevertheless, successful evolution experiments have been carried out with either method.

advantages and disadvantages. In any case, the generated library should obey two major criteria: it must be diverse enough to contain individual sequences with beneficial mutations, and it should be of high enough quality to reduce the experimental "noise" (such as sequences with stop codons or frameshifts) in the subsequent selection experiment [90].

A primary prerequisite of any selection system is the coupling of the genotype (the gene sequence) and phenotype (the respective protein displaying the properties) of any individual library member. Briefly, this can be done by two strategies. Either the gene and the protein need to be compartmentalized in cells or artificial compartments, such that the "improved" phenotype stays connected to the altered gene, or the protein has to be physically coupled to the gene, such that they can be isolated as a particle containing both gene and protein (Figure 39.5). By selecting

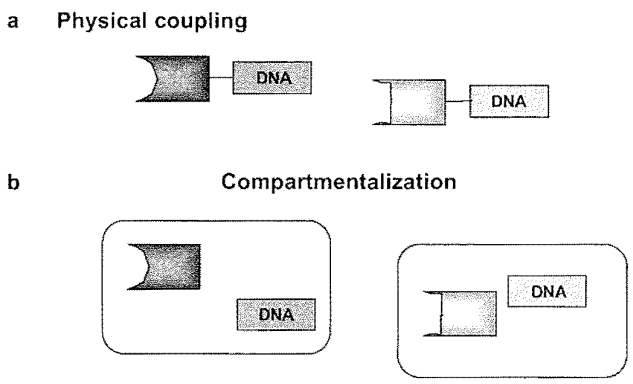

Fig. 39.5. Two principal strategies to link phenotype and genotype are depicted. The genetic material must be connected in a unique way to the protein, which defines the phenotype, such that the gene encoding the valuable mutation can be selectively amplified. A collection of two mutant proteins and mutant genes is shown (light gray and dark gray). (a) This connection can be realized through a direct link (e.g., in phage display or ribosome display), as shown on the top. In this case, all assemblies can freely diffuse in the same volume and the selection must filter out those proteins with the desired function, e.g., by binding to a ligand. (b) Alternatively, gene and protein must be in the same compartment. In nature, this is realized in cells, and microbial cells can be manipulated so that each takes up one variant of a mutant collection. The key is to identify the improved phenotype. This can be done by screening (individual assays on cells) or selection (giving cells with the desired property of the protein a growth advantage). Rather than natural cells, water-in-oil emulsions can be used to create artificial compartments of small water droplets in an oil phase. Usually, the emulsion must be broken and the proteins must be selected by binding to identify the one with the desired phenotype. For details, see text.

the proteins displaying the desired properties, the linked gene sequence can be inherited and amplified subsequently. The various selection technologies differ mainly in the way this physical linkage or compartmentalization is achieved. This will be discussed in more detail in the following section. The library must then be screened or subjected to selection for a certain function, and a defined selection pressure can be applied to direct the selection towards a molecular quantity of interest.

It is necessary at this point to discriminate between "screening" and "selection." Screening methods examine individual members of a library for a given property (e.g., catalytic activity or solubility). For certain properties, screening is often the only way to go. The number of mutants to be screened in a reasonable time depends on the versatility of the screening method. Despite constant progress in automation and miniaturization, even the best screens to date usually do not allow assessment of more than 10^6 variants in a reasonable time. In contrast, selection methods force the single library members to compete with each other, and members that best fulfill the specified criteria are enriched. Often, the selection is based

on the binding of particular variants to an immobilized ligand. Note that the binding is only a "surrogate quantity" of the real property to be improved. The basis for a successful enrichment is an efficient counter-selection of variants that do not possess properties fulfilling these criteria. Especially during the selection of proteins for higher stability, an efficient counter-selection, in addition to the correct choice of the applied selection pressure, is one of the major experimental challenges [91].

The basic rule of screening and selection technology describes the importance of assigning the correct selection pressure. This has been succinctly phrased: "You only get what you screen for!" [92]. An analogous statement can be made for selection. Even though the rule sounds plausible, it is often difficult to translate this statement into an experimental setup, because an additional complication is introduced by the fact that an explicit selection pressure towards just one property is impossible to realize. Depending on the selection procedure used, more than one molecular property will have an influence on the enrichment process, including, for example, affinity or catalytic activity, thermodynamic stability, folding efficiency, and toxicity of the respective molecule. The outcome of the selection experiment will therefore always be a "compromise solution" with respect to the weighting of many properties in a given experimental setup. Assigning the right selection pressure therefore means biasing the selection towards a certain property, rather than exclusively altering this property. For these reasons we will discuss the application of the various selection technologies with an emphasis on the available selection pressures for stability engineering. Some of the described methods and examples will be based on mere screening of library members, but the main focus will be on the selection for favorable protein variants.

Choosing the appropriate selection pressure is just one side of the coin. The proper adjustment of its strength is another important factor. If the selection pressure on the system is too low, molecules with the desired properties will be lost in the background noise of the experiment – i.e., they are not enriched. On the other hand, if the selection pressure is too stringent, even the best variants might not pass the "survival threshold."

The genes of variants that survive the applied selection pressures are subsequently amplified and subjected to another so-called "round of selection." This either can be performed in the absence of further mutagenesis to simply enrich the best members of the initial library (which is constant) or, to completely mimic the Darwinian principle, the selected members can be subjected to alternating rounds of randomization or recombination before subsequent selection, leading to an adaptation of the library from round to round. However, as most of the mutations will be non-beneficial, care has to be taken that the mutation rate is low enough to allow successful enrichment of improved members and not to extinguish the whole population. We use the terms "combinatorial selection" for a process in the absence of such mutations and "evolutionary selection" for a process that includes such mutations.

It has been pointed out that the method of DNA family shuffling is in some respects similar to the consensus approach because it combines gene fragments

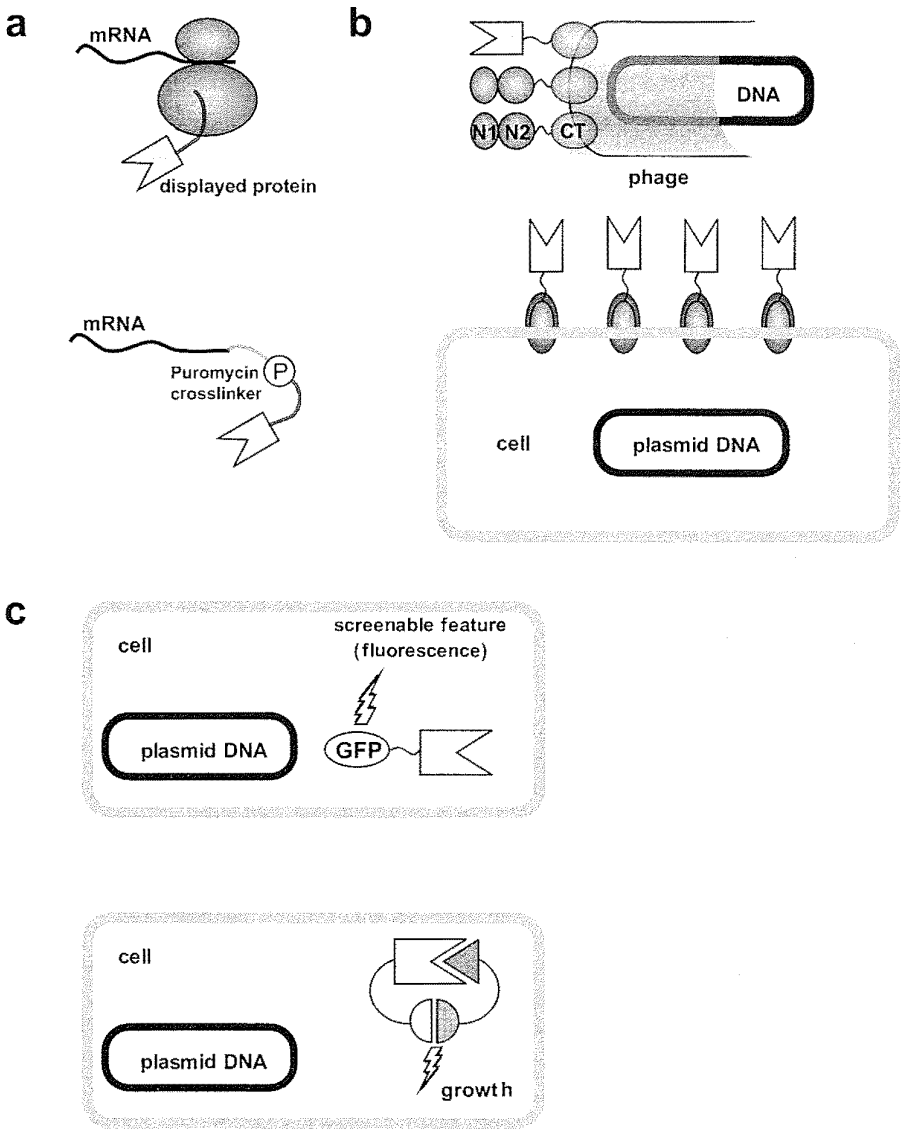

Fig. 39.6. Common display systems used for selection for enhanced biophysical properties. In in vitro display technologies (a), the proteins are produced by in vitro translation, and protein production does not rely on a host organism. The displayed proteins are linked to RNA either by stabilized ternary complexes (ribosome display, upper panel) or by a covalent puromycin cross-linker (mRNA display, lower panel), thereby establishing the genotype-phenotype linkage. (b) In contrast, partial in vitro display systems rely on cells to produce the displayed protein, but selections can be performed in vitro. In phage display (upper panel), the protein is displayed on the surface of filamentous phages, usually fused to

from homologous sources in a random fashion. For mere statistical reasons, the probability of replacing a residue with a consensus residue by gene shuffling is also higher than replacing it with a non-consensus residue [93]. There is, however, a crucial difference: when recombining genes in a random fashion, a selection or screening step is needed subsequently to identify the "fittest" members of the resulting collection. Rather than being based on theoretical assumptions, gene-shuffling methods mimic the natural process of evolution by identifying the members exhibiting a desired property by explicitly subjecting them to selective pressure for this desired property. In practice, however, even with the largest of libraries, a full "re-equilibration" of residues will not take place. In contrast, the consensus approach is based on the explicit assumption that the statistical preferences in a given (often limited or biased) set of sequences indeed reflect the energetic preferences. However, there may be other reasons that certain sequences are prominent.

39.4.2 Screening and Selection Technologies Available for Improving Biophysical Properties

As described above, any selection technology relies on four major steps: (1) generation of a genetic library, (2) establishment of the link between genotype and phenotype upon translation into protein, (3) subsequent screening or selection under defined conditions, and (4) re-amplification of selected members. Depending on how these steps are performed, selection technologies can be subdivided into in vitro methods, partial in vitro methods, and in vivo methods.

While in vivo systems rely on a host organism to express the protein and to carry the respective genetic information, all in vitro display technologies have in common that the protein production and the selection process are performed entirely in vitro, i.e., that the protein is obtained by cell-free translation. In the partial in vitro methods, the genetic information is introduced in cells, where protein production occurs, while the selection process is performed in vitro. All techniques differ in the way the physical linkage between the protein and its genetic information is established. The principles of linking genotype and phenotype are shown in Figure 39.6.

the CT domain of the minor coat protein g3p. The three domains (N1, N2, and CT) of the minor coat protein g3p are depicted. The phage particle also carries the gene encoding the displayed protein. In other partial in vitro technologies, the proteins are displayed on the surface of the expressing host cell itself, which are either bacterial cells (bacterial surface display) or yeast cells (yeast surface display) (lower panel) that also harbor the respective plasmid DNA. (c) In vivo screening or selection systems use the properties of fused reporter proteins for screening (upper panel) or split proteins for intracellular selection (lower panel) in which cellular growth and therefore amplification of the genetic material occurs only if the two protein halves are reconstituted upon interaction of the fused library protein with its target.

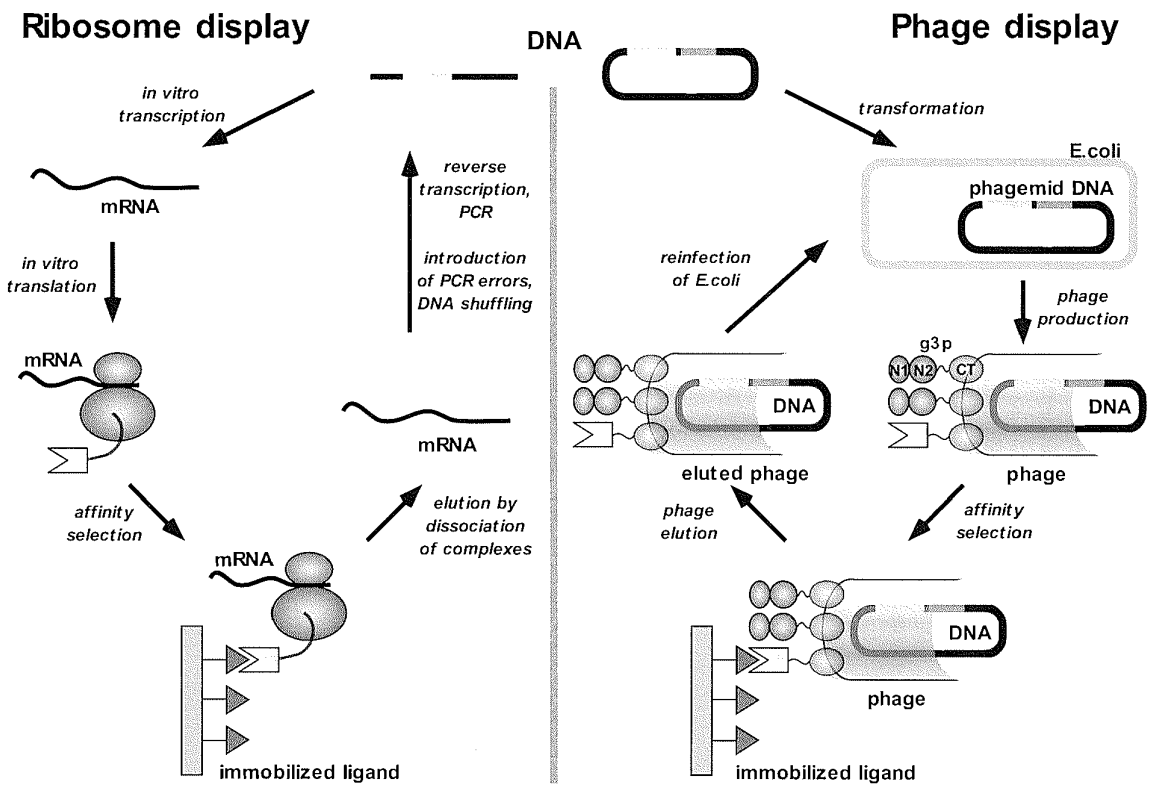
Ribosome display
DNA
Phage display
in vitro transcription
mRNA
in vitro translation
mRNA
affinity selection
mRNA
immobilized ligand
reverse transcription, PCR
introduction of PCR errors, DNA shuffling
mRNA
elution by dissociation of complexes
transformation
E.coli
phagemid DNA
phage production
g3p
N1 N2 CT
DNA
phage
affinity selection
DNA
phage
immobilized ligand
phage elution
DNA
eluted phage
reinfection of E.coli

Fig. 39.7

39.4.2.1 In Vitro Display Technologies

In vitro display technologies started off with the selection of peptides [94] and were then made efficient enough to select for functional proteins [95]. A number of technologies have been developed, which will be described briefly: ribosome display [95, 96], RNA-peptide fusions [97, 98], and water-in-oil emulsions [99, 100]. Water-in-oil emulsions constitute artificial compartments, rather than a physical link, and thereby create a phenotype-genotype coupling and can also use a selection step for interaction with a target.

In ribosome display, the genetic library, usually in the form of a PCR product, is directly used to produce mRNA by in vitro transcription. This mRNA contains a ribosome-binding site for the subsequent translation of the protein and several features that stabilize it against degradation. The encoded protein variants are expressed in a cell-free translation reaction by a stoichiometric number of ribosomes. The essential linkage of the translated proteins to their respective mRNA molecule is achieved by eliminating the stop codon and stabilizing the ribosomal complexes, which prevents the release of the translated protein from the ribosome. In the related method of mRNA-peptide fusions, additional steps are used after this stoichiometric translation to covalently couple a linker between the end of the mRNA and the protein. The ribosome is then removed and the complexes are purified. These complexes are then used to bind to an immobilized target. In all these methods, after the selection process the re-amplification of genetic material can be performed solely by biochemical means, using reverse transcriptase to first synthesize RNA-DNA hybrids, which are then subsequently amplified in a PCR reaction. The resulting DNA then serves as a template for the production of mRNA used for the next round of selection (Figure 39.7).

In vitro display technologies offer several advantages over in vivo selection systems. First, much larger library sizes are accessible because the creation of large molecular diversity on the genetic level does not pose a great challenge, and diversities up to 10^{13} are easily achieved. For in vivo systems, the critical step that limits the size of the displayed library is the transformation efficiency of the respective host organism. Only a small fraction of the initially generated pool will actually enter the cells, and, depending on the host organism, library sizes of about 10^7 variants in yeast or up to 10^{11} variants in bacteria can be achieved. In contrast, for

Fig. 39.7. Selection cycle of ribosome display (left panel) and phage display (right panel). For ribosome display, all necessary steps are performed in vitro, whereas for phage display protein translation, folding, and phage assembly take place in the bacterial host cell. Other selection criteria can be applied to the protein of interest (see Figure 39.8) before selecting for binding to a target, thereby introducing selection pressures for higher stability. It has to be guaranteed, however, that the selection conditions are not too harsh to destroy the linkage of genotype and phenotype. Importantly, the genetic pool of library members in phage display remains constant in the course of several selection rounds, whereas in ribosome display, the genetic pool undergoes continuous modifications during the in vitro amplification steps by the limited accuracy of the polymerases, resulting in a true Darwinian evolution. In addition, this pool can deliberately and conveniently be altered by introducing additional mutations or recombination events such as DNA shuffling.

in vitro selection systems such a transformation step is not necessary and the diversity of the library is defined by the number of different RNA molecules added to the cell-free translation reaction or by the number of functional ribosomes, whichever number is smaller. For a library of 10^{14}, a 10-mL translation reaction is required. A second advantage is that, because translation and protein folding take place in vitro, reagents can be added during protein synthesis, which can either promote protein folding or minimize aggregation of the displayed proteins (e.g., chaperones), or substances that exert certain selection pressures can be added. In vitro translation encompasses an additional advantage: many proteins are not compatible with in vivo selection systems because the wild type or at least some of the mutants are toxic to the host organism, undergo severe degradation, or cannot be expressed functionally in the respective cellular environment. Last but not least, a major advantage results from the amplification process after the selection has been performed. Because all enzymes used to convert and amplify the genetic material possess an intrinsic error rate, the genetic pool is never constant and is continuously modified from round to round. Therefore, even without special measures to increase the error rate, usually some "evolution" is observed. The error rates can be additionally increased by performing error-prone PCR or by recombining favorable mutations by means of DNA shuffling as explained in the previous section.

39.4.2.2 Partial in Vitro Display Technologies

We define partial in vitro display technologies as those methods in which host organisms are employed to carry and amplify the genetic information and to produce the proteins from this genetic library, usually encoded on plasmids. The selection step, however, is performed in vitro. Therefore, selection pressures similar to those in the in vitro methods can be applied. The array of available techniques includes phage display [101], bacterial surface display [102], and surface display on yeast [103].

Phage display is still the most popular selection technique, and due to its robustness, most of the studies that aim for enhanced biophysical properties of proteins have utilized this technique. Originally, phage display was applied as a method for the identification and selection of peptides or proteins binding to a specific target [101]. More recently, several developments paved the way for its application as a tool for studying protein folding, and as a result, it is now used for the selection of proteins with improved biophysical properties.

We will briefly review the main principles of phage display in its most common format. Variants of the protein of interest are fused to the minor coat protein g3p of the filamentous bacteriophage, which consists of three domains (namely, N1, N2, and CT) that are connected by glycine-rich linkers. The fusion protein is usually encoded on a phagemid vector, expressed in the E. coli host and assembled. Typically, the other proteins necessary to produce an intact phage particle are provided by the helper phage. The assembled phages are secreted by the cell, which does not lyse, and the phages can be collected. While the protein is displayed on the exterior of the phage (typically as an N-terminal fusion to either N1 or CT), the gene of interest encoding a library member is packed into the phage particle

upon its assembly in the bacterial host. The phages can then be subjected to the selection process. Phages compete with each other for binding to an immobilized target, using the displayed library members for interaction, and can thereby be captured. Optionally, before capturing, the phages can be subjected to harsh environmental conditions. The captured phages are subsequently amplified by re-infecting bacteria, which then produce phages for a new round of selection (Figure 39.7). In the ideal case, only phages that specifically recognize the target would be amplified. Due to reasons such as nonspecific "sticking" of phages to the surface, the efficiency of selection is in practice at most a 1000- to 10 000-fold enrichment over nonspecific molecules and can even be very much lower.

Several aspects are important when phage display is used as a means of evolving proteins with improved biophysical properties. The correct folding of the fusion protein during phage morphogenesis is an important parameter determining the frequency of incorporation of correctly folded proteins displayed on the phage. Folding intermediates that have a strong aggregation tendency lead to aggregation of a particular library member during the assembly process, and rather than the g3p fusion, the g3p wild-type protein from the helper phage will be incorporated. Despite this inherent selection pressure for the folding efficiency of displayed proteins, the enrichment of improved variants is slow [104], and the utility of phage display to study and improve folding of a given protein was initially not obvious [105]. Combined with a selection for protein functionality, it nevertheless sets the basis for selecting proteins according to their folding properties, as will be discussed in Section 39.4.3.2.

Since the selection procedure is performed in vitro, a large variety of external selection pressures can be applied to the phage particles and be combined with functional selection. If function is omitted as a direct indicator for the native state, other criteria that directly correlate with the native state have to be translated into a selectable feature. This opens the door to select for proteins, which do not bind a ligand, albeit with the caveat that it is not assured that the native state is being selected. Such general selection approaches for physical properties take advantage of the fact that unfolded proteins are more susceptible to proteolytic cleavage than are compactly folded ones [106]. One variation on this theme makes use of the modular nature of the g3p protein, thereby linking phage infectivity directly to the proteolytic susceptibility of the target protein. This system has been called "Proside" (protein stability increase by directed evolution) [107] and will be discussed in Section 39.4.4.4.

39.4.2.3 In Vivo Selection Technologies

In vivo selection technologies such as the yeast two-hybrid system [108, 109] or other split protein complementation assays [110, 111] are valuable tools for studying protein-protein interactions in living cells. They commonly employ reporter systems, in which a covalent link between the protein of interest and another so-called reporter protein is established. Cell growth or a colorimetric reaction is dependent on either a specific protein-protein interaction or the solubility of a critical component. The reporter protein transduces certain properties of the host protein,

most importantly its native fold and its ability to interact and/or its solubility or its resistance to cellular proteases, to a screenable or selectable feature of the fused reporter protein itself [112]. Examples of reporter proteins are transcription factors [108], critical metabolic enzymes [111], or proteins that can easily be assayed [113].

Several facts, however, limit the applicability of in vivo systems for stability engineering. First, as the viability of the host organism has to be guaranteed, external selection pressures cannot easily be applied. Thermophilic organisms represent in some cases an interesting alternative, but conditions allowing selections are hard to establish. Additionally, the host environment as well as the fused reporter proteins possibly perturb the characteristics of the target protein. Furthermore, control experiments must ensure that growth is really dependent on the interaction of interest and that mutations in the host have not "short-circuited" the selection strategy.

39.4.3 Selection for Enhanced Biophysical Properties

As has been explained in the Introduction, desirable biophysical parameters include solubility, stability, and folding efficiency. Even though the various selection methods and conditions usually do not improve one of these properties exclusively, they are often biased towards one of them. We will therefore discuss different ways to exert selection pressure on the system with respect to the property that is likely to be changed.

39.4.3.1 Selection for Solubility

Solubility, correctly defined as the maximal concentration of the native protein that can be kept in solution, is usually not a property to be selected, as most proteins – with the exception of membrane proteins – are sufficiently soluble. The word "solubility" in the context of screening is often inaccurately used to refer to "soluble expression yield," and thus it usually mirrors the efficiency of folding in the cell. Even though soluble expression yield is not necessarily equivalent to the folding properties and even less to the stability of a protein, a correlation can often be observed, and soluble expression yield is comparatively easy to screen for. In some cases, a function for a given protein cannot be assigned, or it is difficult and laborious to screen for, and then this property becomes especially important.

The fluorescence yield of bacterial colonies expressing proteins or protein domains fused to the green fluorescent protein (GFP) correlates over a wide range with the soluble expression yield of the fused domain [114]. Screening for fluorescence intensity is a versatile method, and, combined with directed evolution, has the potential to select variants of proteins that are less aggregation-prone than their progenitors [115, 116]. In a similar setup, reporter proteins can be used as selection markers instead, such as by fusing the protein of interest to chloramphenicol acetyltransferase [117]. Assays exploiting protein-protein interactions can also be used for this purpose and they have the additional advantage that the fused reporter sequence can be much smaller and thereby minimize the risk of a perturb-

ing influence of the reporter itself on the domain of interest [113]. To completely exclude such perturbing influences, reporter systems might possibly be established that do not need any fusion at all and exploit instead the stress response of the expression host cell. As the overexpression of proteins often activates particular stress response genes by the accumulation of insoluble aggregates, the respective gene promoters can be employed to activate reporter genes instead [118].

In special cases, reporter systems can also be combined with other selection methods to select for folding properties. Many intrinsically stable proteins contain permissive sites in loops, into which polypeptide chains of variable length can be inserted without loss of function of the host protein. The higher the number of residues inserted, the larger the entropic cost of ordering these residues will be. Consequently, the overall stability of the host protein should decrease. However, the entropic cost of inserting a folded sequence will be lower than that of an unfolded sequence of the same length. The probability of the host protein reaching the native (functional) state should thus correlate directly with the ability of the inserted sequence to fold into a compact structure. If host proteins with various inserted sequences are displayed on phages and if the host protein allows a selection for "foldedness" by means of binding to a target, folded sequence insertions are enriched. This system has been termed "loop entropy reduction phage display selection" [119]. While attractive in theory, it was found experimentally that mostly sequences that keep the hybrid protein soluble are enriched.

Protein solubility, as used here, can also be interpreted as a lower degree of exposed hydrophobic residues, and folded proteins usually display fewer hydrophobic residues on their surface than do unfolded proteins. Conversely, the exposure of hydrophobic residues is usually a sign of non-native states. Display technologies can thus, for example, use the interaction with hydrophobic surfaces to select against more hydrophobic, i.e., less "folded," proteins [120].

In summary, protein solubility can reflect in many cases the folding properties of proteins and thus be used as a selection criterion. Moreover, soluble expression in vivo is often a major requirement for the large-scale production and the convenient in vitro handling of proteins. In any case, solubility should never be confused with protein stability. Even though both properties might correlate in some cases, the governing principles are often of a different nature.

39.4.3.2 Selection for Protein Display Rates

The in vivo folding efficiency of proteins represents an intrinsic selection pressure when using phage display. Because correct folding is a prerequisite for both incorporation into the phage coat and binding to the target, the subsequent selection of phages that are able to bind to the target is partly influenced by the folding properties of the displayed molecule.

However, this does not automatically imply that the selection is driven towards superior folding properties. As mentioned above, the enrichment factors of proteins with superb folding behavior over the poorly folding members are low [104]. Moreover, even though the binding of a given protein to its target is a very direct way to monitor and screen for its proper folding, the selection criterion "folding"

is not decoupled from the selection criterion "binding affinity." If the enrichment factor for one property (folding) is low, the selection is more likely to be driven towards the other (affinity). As a result, selected members will represent a compromise between folding properties, which only have to be sufficient under the given experimental conditions, and the binding affinity to their target.

By randomizing exclusively the residues that build the hydrophobic core of the IgG binding domain of peptostreptococcal protein L, which are not involved in ligand binding but determine the stability and folding kinetics of this small protein, Gu et al. [121] could unambiguously demonstrate the utility of phage display for studying the stability and folding efficiency of proteins. A more extensive randomization effort with subsequent characterization of the selected variants pointed out some important aspects of selection for folding [122]. First, it showed that folding kinetics is not the critical parameter for the selection but rather the overall stability of the mutants, a tendency which could be confirmed by more advanced selection approaches (see Section 39.4.4.2). Second, the authors demonstrated the utility of selection methods to identify certain residues that are important for the folding mechanism. Third, the selected pools were highly diverse and the thermodynamic stabilities of all variants were lower compared to the wild type. In fact, most of the selected proteins denatured just above room temperature.

These results illustrate the principle that any selection – be it natural or in the test tube – simply continues until the minimum requirements are met, in this case, functionality at the selection temperature. Consequently, this suggests that additional, more stringent selection pressures are necessary to really accomplish a directed evolution for improved properties.

39.4.3.3 Selection on the Basis of Cellular Quality Control

Additional selection pressure can be provided by the host organism itself. For example, the secretory quality-control system of eukaryotic cells discriminates proteins according to their folding behavior in an efficient way. This is based on mechanisms that lead to retention of misfolded proteins in the endoplasmic reticulum (ER), followed by degradation of these proteins. In yeast, the Golgi complex can reroute misfolded proteins, which have escaped ER retention, to the vacuole for degradation, thereby constituting an additional important quality-control pathway [14]. In combination with yeast surface display, these mechanisms can be employed to bias functional selections towards enhanced folding efficiency. Several studies have shown that the surface display rate of proteins strongly correlates with their thermostability and their soluble secretion efficiency [123, 124]. As an example, by making use of elevated temperatures during expression as an additional selection pressure, improved T-cell receptor fragments, whose thermostability exceeded by far the expression temperatures, could be obtained [125]. Even though the low transformation efficiency of yeast restricts the accessible library size, one obvious advantage of the method is the applicability to glycosylated eukaryotic proteins, which generally are not amenable to yeast two-hybrid or phage display methodologies [126].

39.4.4 Selection for Increased Stability

39.4.4.1 General Strategies

The key to all selections for stability is to introduce a threshold that separates the molecule with desired properties from the starting molecules. If the population initially lacks functionality, while only a few members are above the selection threshold, the selection system can distinguish between these slight energetic differences. This is the starting situation if the target protein is initially of very low stability and should be brought to "average" properties.

If, however, the starting protein is already of considerable stability, but should be brought to even higher stability, it can be advantageous to intentionally destabilize it prior to selection in order to find stabilizing mutations that reconstitute its functionality. This principle can be applied to very different kinds of proteins, provided that mutations are known that destabilize the protein to an extent that will subsequently allow the selection for alternative stabilizing mutations (Figure 39.8a). Because the effects of independent stabilizing mutations are often additive, the deliberately introduced destabilizing mutations can be reverted in the context of the additional newly selected ones, thereby rendering the molecule far more stable than the original one (Figure 39.8b).

In principle, all reagents and conditions known to destabilize proteins can decrease the number of library members populating the folded state and can therefore be used to exert increasing selection pressure on the system. Increased temperature and denaturing agents are obvious methods that are useful for selecting proteins of higher thermodynamic stability. The major problem is presented by the compatibility of the conditions with the selection method used.

39.4.4.2 Protein Destabilization

An example of the stabilization of a naturally unstable domain by using phage display was a selection performed with the prodomain of the protease subtilisin BPN′ [127]. At room temperature the prodomain folds into a stable conformation only upon binding to subtilisin. By randomizing positions that are not directly in contact with subtilisin and subsequent selection on subtilisin, a mutant was selected that showed an increase of $\Delta G_{\text{unfolding}}$ by 25 kJ mol^{-1}, from −8 kJ mol^{-1} to 17 kJ mol^{-1}, despite the fact that the library size was comparatively small. Intriguingly, the predominant energetic contribution was mediated by a selected disulfide bond. Previously, a similar strategy had been used to select for thermodynamically favored β-turns of the B1 domain of protein G [128]. Based on considerations described above and on the fact that the replacement of amino acids on the surface of proteins would be predicted to have only moderate effects on stability, the authors reasoned that such a selection could be successful only for proteins of marginal stability. Most of the substitutions would then lead to a positive free energy of folding, and thus the molecule would fail to fold into a functional form at all. In fact, selected turn sequences showed clear sequence preferences only if less stable

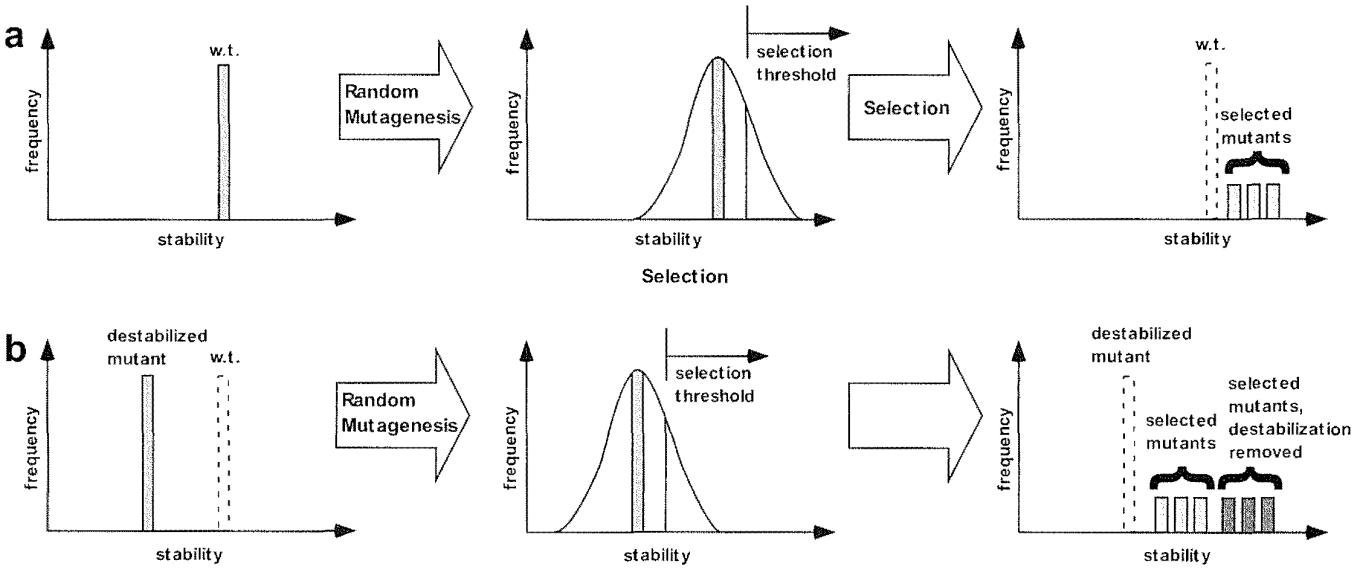

Fig. 39.8. Principle of selection for proteins with improved stability. (a) By random mutagenesis of a given sequence, many mutations are obtained. Some of these mutations are favorable for the stability of the native protein and others are unfavorable. As a result, a diverse pool of proteins with a distribution of different energies is created. Upon exposure of the pool to an external selection pressure, only mutants with stabilities exceeding the selection threshold can be recovered and amplified. However, the selection threshold has to be set high enough to allow an efficient selection of "improved" members. (b) Destabilization of the wild-type sequence prior to selection allows reducing the necessary selection threshold. Thereby, alternative mutations that stabilize the native fold of the protein can be identified and the selection process becomes more efficient. Moreover, additional stability gains can be achieved by removing the deliberately introduced destabilizing mutations after the selection. The initially lost energy is therefore regained, resulting in mutants of higher stability than the wild-type protein. Adapted from Wörn et al. [76].

host proteins were used to accept the turn. In this case, the selected sequences either resembled the wild-type turn or reflected the statistical preferences of turn sequences in the databases and stabilized the protein by 12–20 kJ mol^{-1}, compared to random sequences. Moreover, increased temperature was used as an additional selection pressure during the phage selection.

As discussed in the previous section, the inability of the wild-type target protein to fold is the prerequisite for performing an efficient positive selection for functionality by additional mutations. The internal disulfide bond of immunoglobulin domains significantly contributes to the stability of antibodies [129]. The dramatic loss of free energy of folding upon their removal usually renders antibodies nonfunctional. By first destabilizing a scFv antibody fragment through replacement of cysteines with other residues in both variable domains separately, a completely disulfide-free antibody was identified after several cycles of functional selection and recombination by DNA shuffling [17]. One globally stabilizing mutation was found to compensate for the initial stability loss. This study illustrates some important features of how several parameters exert influence on different stages of the selection process. Interestingly, the stabilizing mutation was already selected during the first rounds, showing that the loss of thermodynamic stability was the primary problem that had to be overcome. Because the stability gain of this mutation was large enough to shift the free energy of folding above the required threshold, all successive rounds did not affect protein stability but led to a fine-tuning with respect to the improvement of folding yield [17]. To illustrate the principle of additivity, reintroduction of the disulfide bridge into the selected variant yielded a scFv antibody of very high stability and superior expression yields [84].

39.4.4.3 Selections Based on Elevated Temperature

In vivo selection systems using thermophilic expression hosts have been employed for the stabilization of enzymes [130]. However, a rather large set of requirements must be met to apply such a selection system. In order to use stability of the enzyme at elevated temperatures as the selection criterion, its enzymatic activity must be vital to the thermophilic organism, and the corresponding gene of the thermophilic host has to be deleted. Randomized versions of a mesophilic enzyme can then be screened by means of metabolic selection. Although the approach is very powerful, the utility of these systems is restricted to special cases.

To evolve enzymes with altered thermal stability, conventional screening for enzymatic activity in vitro after randomization or recombination of the respective gene and subsequent expression of the enzyme is still the most widely used method. An activity screen rapid and sensitive enough to identify slightly improved members from a vast pool of mutants is the key feature of this directed evolution approach [131]. However, individual screens have to be developed for each specific class of enzymes. Even though screening limits the explorable sequence space considerably compared to selection, mesophilic enzymes such as subtilisin E [132] and *p*-nitrobenzyl esterase [133] could be converted into mutants functionally equivalent to thermophilic enzymes by only a few rounds of directed evolution. Similar to observations discussed before, only very few mutations were necessary to in-

crease the melting temperatures by more than 14 degrees. An important feature of screening compared to selection is that the much smaller library size is partially compensated for by directly measuring the quantity of interest: enzymatic activity at elevated temperature. In contrast, most selection systems use a surrogate measure where "false positives" can lead to the phenotype by mechanisms different from the ones desired.

Elevated temperatures not only can be used in screening but also can be combined with selection technologies. While ribosome display is not an option – because low temperatures are essential to keep the ternary complexes of RNA, ribosome, and nascent polypeptide intact and thus ensure the coupling of genotype and phenotype – phage display has proved to be a very suitable method for harsh conditions due to the robustness of the phage particles. Nevertheless, future improvements of in vitro technologies may allow their application under more stringent conditions. As of today, in the case of in vitro technologies, it is vital to destabilize the protein first (see Section 39.4.4.2); then, very significant stability improvements can be selected [134].

The upper temperature limit for selections using filamentous M13 phages is approximately 60 °C [135]. Above this temperature, re-infection titers of the phages are severely decreased, presumably due to the irreversible heat denaturation of phage coat proteins. Phages displaying the protein of interest can be incubated up to this temperature, and proteins still able to function can be selected. It should be noted that the exposure of phages to higher temperature after phage assembly exerts a somewhat different stress on the proteins than in the methods described before. While in the previous examples the functionality of the proteins was influenced mainly by in vivo folding efficiency and the protein stability during the panning procedure, it is the irreversible unfolding reaction at a given temperature that now becomes an additional parameter. The fraction of unfolded molecules will therefore reflect the rate of unfolding and thus kinetic stability.

39.4.4.4 Selections Based on Destabilizing Agents

The use of protein destabilizing agents for biasing the selection pressure towards stability is limited to in vitro and partial in vitro display methods. Like temperature, the concentrations of denaturing agents can be controlled precisely and can be varied from round to round, allowing a gradual increase of stringency.

Even though phages are quite resistant to denaturing agents [136], one should be aware of possible general problems when selections are based on the chemical denaturation properties of the displayed proteins. However, if chemical denaturation is combined with a selection for binding, high concentrations of denaturant can prevent binding to the target, even if the protein is not yet unfolded: as the forces governing ligand binding are very similar to those responsible for protein stability, both are disturbed by chemical denaturants. Conversely, because chemical denaturation is in fact often reversible, removal or dilution of the denaturant prior to ligand binding will often result in refolding on the phage and thus release of selection pressure. Moreover, ionic denaturants such as guanidinium chloride weaken electrostatic interactions and strengthen hydrophobic ones. This might im-

pair the selection of mutations that introduce additional ionic interactions on the protein surface [59] and may favor additional hydrophobic interactions, which is not necessarily desired.

If the protein is known to be stabilized by disulfide bridges, a strategy similar to the one used by Proba et al. [17] described above may be applicable. To identify globally stabilizing mutations, which compensate for the stability loss upon removal of disulfide bridges, selection can be performed in the presence of reducing agents such as DTT. Thus, the disulfide-forming cysteines do not have to be removed in advance. However, one should be aware of the fact that some disulfide bridges in proteins, once formed, are often hard to reduce, especially if they are buried within the protein core. Thus, it is advantageous to add the reducing agent at a time when the protein is not yet folded. Jermutus et al. [134] stabilized a scFv antibody fragment by using ribosome display. In contrast to the phage display method, the synthesis of the targeted protein occurs in vitro, which makes the polypeptide chain accessible to reagents during its synthesis on the ribosome and before folding has taken place. DTT was added during translation of the scFv, and its concentration was continuously increased from round to round. When using ribosome display, the population undergoes slight changes due to mutations occurring during PCR, resulting in an iterative adjustment of the selected variants towards tolerating the increasingly stringent conditions.

39.4.4.5 Selection for Proteolytic Stability

In each of the examples cited above, selection for stability was based on a functional selection. Therefore, only those proteins for which a specific function can be assigned could be targeted; in addition, this function has to be screenable or, better yet, selectable. It would be highly advantageous to completely uncouple function from stability in the selection process to extend the range of problems to which this can be applied. A more general approach would thus be an invaluable tool for engineering any given protein and for selecting stable folds from a pool of de novo designed proteins. For functional selections, an additional problem arises from the fact that after a certain stability threshold is reached, allowing enough proteins to populate the folded state, the selection is likely to run towards improved binding properties instead. On the other hand, certain mutations might be stabilizing but might result in slight structural rearrangements that affect functionality in a negative way. Thus, it will not be possible to recover these mutations [106].

A general approach that has proved to be powerful for stability engineering is based on the concept that compactly folded proteins are much more resistant to proteolysis than are partly folded or unfolded proteins [137]. Several variations on this theme have been applied to phage display selections. In principle, it is only necessary that the phages displaying proteins, which can be cleaved by the protease, can be efficiently separated from the phages displaying protease-resistant proteins. This can be achieved in a physical way by providing the displayed protein with an N-terminal tag sequence allowing capture of only phages with non-cleaved proteins [106]. Alternatively, the ability of phages to re-infect bacteria can be directly linked to the protease resistance of the protein of interest [107, 136]. This

second alternative takes advantage of the modular structure of the protein domains responsible for phage infection. If the displayed domains are inserted between the carboxy-terminal domain of gp3 and the amino-terminal domains N1 and N2, which are required for phage infectivity, proteolytic cleavage of the displayed domain renders the phage noninfectious. This was inspired by the so-called selectively infective phage method (SIP), where it was shown that the g3p domains could be interrupted by additional domains and even an interacting pair [138]. By employing a phage system derived from the SIP technique, which lacks any wild-type g3p, it is assured that infectivity is completely abolished upon proteolytic cleavage [107] (Figure 39.9). Alternatively, phages can be engineered to make the remaining wild-type g3p itself susceptible to the proteolytic attack [136]. The principle of protease selection can also be applied to in vitro selection techniques in which the displayed protein is freely accessible [120].

Even though proteolysis seems at first glance to be less correlated to the stability of proteins than temperature, it is suitable for optimizing packing of the hydrophobic core [139] as well as for optimizing electrostatic interactions on the protein surface [59]. One reason for the strength of this approach is based on the fact that proteolytic cleavage is an irreversible reaction, which is not necessarily the case for denaturation by temperature or denaturing agents. Furthermore, protease resistance monitors the flexibility of the polypeptide chain rather than complete denaturation. It is therefore capable not only of selecting against completely unfolded variants but also of detecting local unfolding events. Because sites of local unfolding often initiate the global unfolding process, it can be advantageous to remove such sites. Nevertheless, the effective cleavage of flexible parts of the protein restricts the method to proteins that do not have extended flexible regions in the native state. Furthermore, a selection against the primary recognition sequence of the protease is clearly a possible outcome of such experiments.

None of the described methods can stand completely on its own. Many methods can easily be combined to increase the stringency of a selection. A combination of temperature stress and increasing amounts of denaturing agents may, for example, lead to higher flexibility in certain regions of the protein, thereby increasing the sensitivity of the subsequent proteolytic attack. Additionally, temperature and the concentration of denaturing agents allow a much tighter control of the selection pressure than do increasing concentrations of protease and thus open the possibility of a well-controlled gradual increase of selection stringency.

39.5 Conclusions and Perspectives

Several different approaches are now available for engineering proteins for enhanced biophysical properties. In many cases, few and specific mutations are sufficient to provide proteins of marginal stability with considerably stabilizing features. Thus, the challenge for protein engineering is to identify these positions in a given sequence among the vast number of possible changes and to correctly alter them. Each of the applied techniques has its own merits and bottlenecks. Instead

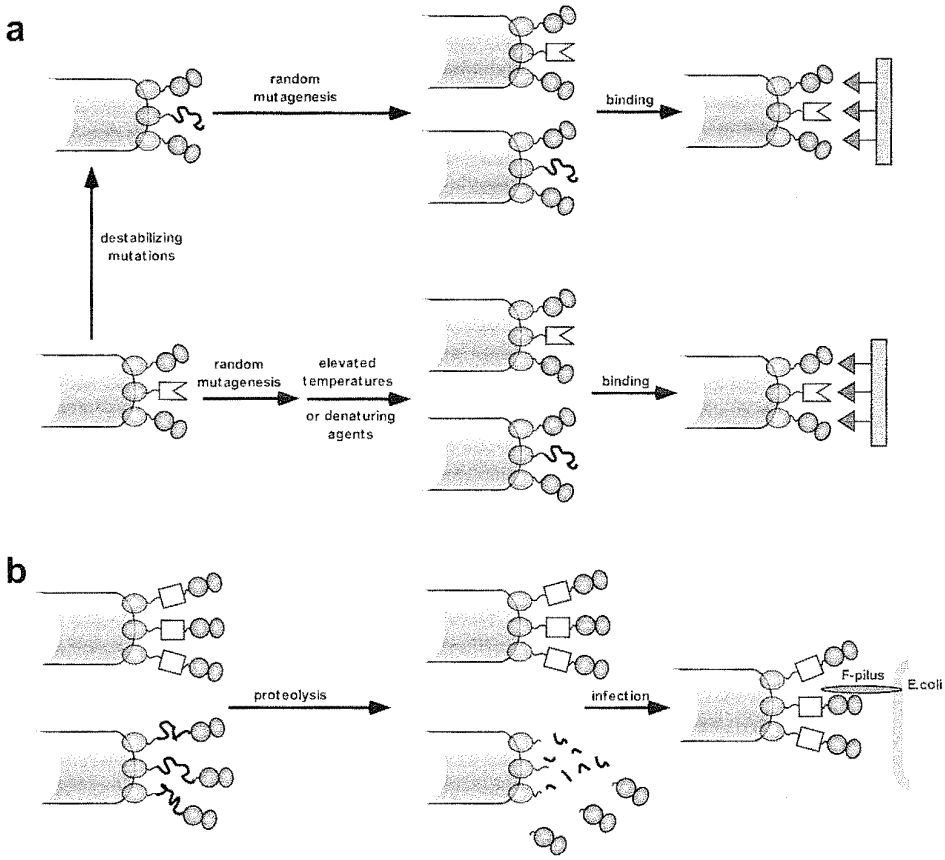

Fig. 39.9. Strategies for selecting for improved protein stability and folding by phage display. (a) Methods utilizing the binding to a given target (affinity selection) as a means for selecting members with improved folding behavior and higher stability from a protein library. Several types of selection pressure can be applied in order to recover mutants with enhanced stability. Destabilizing mutations are deliberately introduced (top) to render the protein "nonfunctional," and alternative stabilizing mutations are identified by selecting variants whose functionality is regained. Alternatively (bottom), elevated temperatures or denaturing agents can be used to render most of the protein variants "nonfunctional," allowing the selection of simply the "fittest" members. (b) The resistance of the target protein to a protease can be combined with the ability of phages to re-infect bacteria. The protein library is inserted between the C-terminal domain and the N-terminal domain in all copies of the g3p protein (cloned into the phage genome), which are needed for re-infection of bacterial cells. By proteolytic cleavage of the inserted protein, these domains are cut off and phage infectivity is lost. In alternative approaches, an N-terminal tag sequence is fused to the protein of interest instead of the N-terminal g3p domains. Upon proteolytic cleavage, the tag is lost. In contrast, phages presenting proteins that are resistant to proteolytic cleavage can subsequently be captured on an affinity matrix binding to the tag sequence. In order to further increase the selection threshold, the shown selection strategies can also be combined.

of playing them off against each other, the future challenge will be to find ways to synergistically use them to improve a given molecular property.

Evolutionary methods have the advantage of being much less biased by theoretical assumptions or working hypotheses. Additionally, proteins stable in new environments may be evolved, e.g., proteins that fulfill a given function in non-aqueous solutions or high concentrations of detergents. Because the biophysical principles in such environments are of a different nature, rational design has to rely on a much smaller empirical dataset. While rational approaches are likely to become more important as more structural and experimental data become available, notably also those from selection experiments, the large number of variations that are potentially able to improve the biophysical properties of a protein often still exceeds the experimentally accessible number. Moreover, because rational engineering has to rely on the available dataset, mutations that lie off the beaten track will rarely be identified. Nevertheless, selection experiments often identify the same "key" mutations in proteins. It is then useful to exploit this information and directly introduce such mutations. A "rational" analysis can also help to recombine important "key" mutations in selected clones and to reduce the effects of a selection-neutral genetic drift. Another combination of rational and combinatorial methods is the creation of "smart" libraries of variants. Library design that is based on such structural considerations and principles will therefore allow more accurate focusing of the selection on specific regions of interest and thereby increase the chances for success.

Acknowledgements

The authors thank Drs. Daniela Röthlisberger, Casim Sarkar and Annemarie Honegger for critical reading of the manuscript. B. S. was the recipient of a Kekulé fellowship from the Fonds der Chemischen Industrie.

References

1 Taverna, D. M. & Goldstein, R. A. (2002). Why are proteins marginally stable? *Proteins* 46, 105–9.

2 Zavodszky, P., Kardos, J., Svingor & Petsko, G. A. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7406–11.

3 Arnold, F. H., Giver, L., Gershenson, A., Zhao, H. & Miyazaki, K. (1999). Directed evolution of mesophilic enzymes into their thermophilic counterparts. *Ann. N.Y. Acad. Sci.* 870, 400–3.

4 Govindarajan, S. & Goldstein, R. A. (1995). Searching for Foldable Protein Structures Using Optimized Energy Functions. *Biopolymers* 36, 43–51.

5 Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry* 29, 7133–55.

6 Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium Thermotoga maritima. *J. Mol. Biol.* 274, 676–83.

7 Sterner, R. & Liebl, W. (2001).

Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 36, 39–106.

8 Zlotnick, A. & Stray, S. J. (2003). How does your virus grow? Understanding and interfering with virus assembly. *Trends Biotechnol.* 21, 536–42.

9 Jaswal, S. S., Sohl, J. L., Davis, J. H. & Agard, D. A. (2002). Energetic landscape of alpha-lytic protease optimizes longevity through kinetic stability. *Nature* 415, 343–6.

10 Fersht, A. R., Bycroft, M., Horovitz, A., Kellis, J. T., Jr., Matouschek, A. & Serrano, L. (1991). Pathway and stability of protein folding. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 332, 171–6.

11 Jäger, M. & Plückthun, A. (1999). Domain interactions in antibody Fv and scFv fragments: effects on unfolding kinetics and equilibria. *FEBS Lett.* 462, 307–12.

12 Willuda, J., Honegger, A., Waibel, R., Schubiger, P. A., Stahel, R., Zangemeister-Wittke, U. & Plückthun, A. (1999). High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment. *Cancer Res.* 59, 5758–67.

13 Ewert, S., Huber, T., Honegger, A. & Plückthun, A. (2003). Biophysical properties of human antibody variable domains. *J. Mol. Biol.* 325, 531–53.

14 Ellgaard, L., Molinari, M. & Helenius, A. (1999). Setting the standards: quality control in the secretory pathway. *Science* 286, 1882–8.

15 Martineau, P. & Betton, J. M. (1999). In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm. *J. Mol. Biol.* 292, 921–9.

16 Schuler, B. & Seckler, R. (1998). P22 tailspike folding mutants revisited: effects on the thermodynamic stability of the isolated beta-helix domain. *J. Mol. Biol.* 281, 227–34.

17 Proba, K., Wörn, A., Honegger, A. & Plückthun, A. (1998). Antibody scFv fragments without disulfide bonds made by molecular evolution. *J. Mol. Biol.* 275, 245–53.

18 Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Eng.* 8, 81–9.

19 Wörn, A. & Plückthun, A. (1999). Different equilibrium stability behavior of ScFv fragments: identification, classification, and improvement by protein engineering. *Biochemistry* 38, 8739–50.

20 Steipe, B., Schiller, B., Plückthun, A. & Steinbacher, S. (1994). Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* 240, 188–92.

21 Lehmann, M. & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr. Opin. Biotechnol.* 12, 371–5.

22 Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* 296, 57–86.

23 Mosavi, L. K., Minor, D. L., Jr. & Peng, Z. Y. (2002). Consensus-derived structural determinants of the ankyrin repeat motif. *Proc. Natl. Acad. Sci. U.S.A.* 99, 16029–34.

24 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, 2186–93.

25 Lehmann, M., Loch, C., Middendorf, A., Studer, D., Lassen, S. F., Pasamontes, L., van Loon, A. P. & Wyss, M. (2002). The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* 15, 403–11.

26 Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P. & Plückthun, A.

(2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J. Mol. Biol.* 332, 489–503.

27 Main, E. R., Xiong, Y., Cocco, M. J., D'Andrea, L. & Regan, L. (2003). Design of stable alpha-helical arrays from an idealized TPR motif. *Structure (Camb)* 11, 497–508.

28 Stumpp, M. T., Forrer, P., Binz, H. K. & Plückthun, A. (2003). Designing repeat proteins: modular leucine-rich repeat protein libraries based on the mammalian ribonuclease inhibitor family. *J. Mol. Biol.* 332, 471–87.

29 Forrer, P., Binz, H. K., Stumpp, M. T. & Plückthun, A. (2004). Consensus design of repeat proteins. *Chembiochem* 5, 183–9.

30 Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Plückthun, A. & Grütter, M. G. (2003). Designed to be stable: crystal structure of a consensus ankyrin repeat protein. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1700–5.

31 Levitt, M., Gerstein, M., Huang, E., Subbiah, S. & Tsai, J. (1997). Protein folding: the endgame. *Annu. Rev. Biochem.* 66, 549–79.

32 Bohm, G. & Jaenicke, R. (1994). Relevance of sequence statistics for the properties of extremophilic proteins. *Int. J. Pept. Protein Res.* 43, 97–106.

33 Jaenicke, R. & Bohm, G. (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8, 738–48.

34 Vieille, C. & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1–43.

35 van den Burg, B. & Eijsink, V. G. (2002). Selection of mutations for increased protein stability. *Curr. Opin. Biotechnol.* 13, 333–7.

36 Matsumura, M., Signor, G. & Matthews, B. W. (1989). Substantial increase of protein stability by multiple disulphide bonds. *Nature* 342, 291–3.

37 Perry, L. J. & Wetzel, R. (1984). Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation. *Science* 226, 555–7.

38 Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* 263, 11820–5.

39 Zhang, T., Bertelsen, E. & Alber, T. (1994). Entropic effects of disulphide bonds on protein stability. *Nat. Struct. Biol.* 1, 434–8.

40 Doig, A. J. & Williams, D. H. (1991). Is the hydrophobic effect stabilizing or destabilizing in proteins? The contribution of disulphide bonds to protein stability. *J. Mol. Biol.* 217, 389–98.

41 Martensson, L. G., Karlsson, M. & Carlsson, U. (2002). Dramatic stabilization of the native state of human carbonic anhydrase II by an engineered disulfide bond. *Biochemistry* 41, 15867–75.

42 Ivens, A., Mayans, O., Szadkowski, H., Jurgens, C., Wilmanns, M. & Kirschner, K. (2002). Stabilization of a $(\beta\alpha)_8$-barrel protein by an engineered disulfide bridge. *Eur. J. Biochem.* 269, 1145–53.

43 Betz, S. F. & Pielak, G. J. (1992). Introduction of a disulfide bond into cytochrome c stabilizes a compact denatured state. *Biochemistry* 31, 12337–44.

44 Iwai, H. & Plückthun, A. (1999). Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* 459, 166–72.

45 Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C. & Benkovic, S. J. (1999). Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13638–43.

46 Camarero, J. A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D. P. & Muir, T. W. (2001). Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* 308, 1045–62.

47 Nagi, A. D. & Regan, L. (1997). An inverse correlation between loop

length and stability in a four-helix-bundle protein. *Fold. Des.* 2, 67–75.

48 Russell, R. J., Ferguson, J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon pyrococcus furiosus at 1.9 Å resolution. *Biochemistry* 36, 9983–94.

49 Macedo-Ribeiro, S., Darimont, B., Sterner, R. & Huber, R. (1996). Small structural changes account for the high thermostability of 1[4Fe-4S] ferredoxin from the hyperthermophilic bacterium Thermotoga maritima. *Structure* 4, 1291–301.

50 Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl. Acad. Sci. U.S.A.* 84, 6663–7.

51 Sriprapundh, D., Vieille, C. & Zeikus, J. G. (2000). Molecular determinants of xylose isomerase thermal stability and activity: analysis of thermozymes by site-directed mutagenesis. *Protein Eng.* 13, 259–65.

52 Pace, C. N. (1992). Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.* 226, 29–35.

53 Lim, W. A., Hodel, A., Sauer, R. T. & Richards, F. M. (1994). The crystal structure of a mutant protein with altered but improved hydrophobic core packing. *Proc. Natl. Acad. Sci. U.S.A.* 91, 423–7.

54 Ventura, S., Vega, M. C., Lacroix, E., Angrand, I., Spagnolo, L. & Serrano, L. (2002). Conformational strain in the hydrophobic core and its implications for protein folding and design. *Nat. Struct. Biol.* 9, 485–93.

55 Ohmura, T., Ueda, T., Ootsuka, K., Saito, M. & Imoto, T. (2001). Stabilization of hen egg white lysozyme by a cavity-filling mutation. *Protein Sci.* 10, 313–20.

56 Chen, J. & Stites, W. E. (2001). Higher-order packing interactions in triple and quadruple mutants of staphylococcal nuclease. *Biochemistry* 40, 14012–9.

57 Hendsch, Z. S. & Tidor, B. (1994). Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* 3, 211–26.

58 Pace, C. N. (2001). Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry* 40, 310–3.

59 Martin, A., Sieber, V. & Schmid, F. X. (2001). In-vitro selection of highly stabilized protein variants with optimized surface. *J. Mol. Biol.* 309, 717–26.

60 Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M. & Makhatadze, G. I. (1999). Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry* 38, 16419–23.

61 Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. & Baldwin, R. L. (1987). Tests of the helix dipole model for stabilization of alpha-helices. *Nature* 326, 563–7.

62 Pace, C. N., Shirley, B. A., McNutt, M. & Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *FASEB J.* 10, 75–83.

63 Loladze, V. V., Ermolenko, D. N. & Makhatadze, G. I. (2002). Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior. *J. Mol. Biol.* 320, 343–57.

64 McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238, 777–93.

65 Hennig, M., Darimont, B., Sterner, R., Kirschner, K. & Jansonius, J. N. (1995). 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile Sulfolobus solfataricus: possible determinants of protein stability. *Structure* 3, 1295–306.

66 Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7, 95–9.

67 Kimura, S., Kanaya, S. & Nakamura, H. (1992). Thermostabilization of

Escherichia coli ribonuclease HI by replacing left-handed helical Lys95 with Gly or Asn. *J. Biol. Chem.* 267, 22014–7.

68 Takano, K., Yamagata, Y. & Yutani, K. (2001). Role of amino acid residues in left-handed helical conformation for the conformational stability of a protein. *Proteins* 45, 274–80.

69 Stites, W. E., Meeker, A. K. & Shortle, D. (1994). Evidence for strained interactions between side-chains and the polypeptide backbone. *J. Mol. Biol.* 235, 27–32.

70 Vega, M. C., Martinez, J. C. & Serrano, L. (2000). Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II′ region of the Ramachandran plot. *Protein Sci.* 9, 2322–8.

71 Gunasekaran, K., Ramakrishnan, C. & Balaram, P. (1996). Disallowed Ramachandran conformations of amino acid residues in protein structures. *J. Mol. Biol.* 264, 191–8.

72 Pal, D. & Chakrabarti, P. (2002). On residues in the disallowed region of the Ramachandran map. *Biopolymers* 63, 195–206.

73 Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin V_H domains with a generalizable approach. *Biochemistry* 42, 1517–28.

74 Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng.* 10, 435–44.

75 Brandts, J. F., Hu, C. Q., Lin, L. N. & Mos, M. T. (1989). A simple model for proteins with interacting domains. Applications to scanning calorimetry data. *Biochemistry* 28, 8588–96.

76 Wörn, A. & Plückthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. *J. Mol. Biol.* 305, 989–1010.

77 Gokhale, R. S., Agarwalla, S., Francis, V. S., Santi, D. V. & Balaram, P. (1994). Thermal stabilization of thymidylate synthase by engineering two disulfide bridges across the dimer interface. *J. Mol. Biol.* 235, 89–94.

78 Hudson, P. J. (1999). Recombinant antibody constructs in cancer therapy. *Curr. Opin. Immunol.* 11, 548–57.

79 Plückthun, A., Krebber, A., Krebber, C., Horn, U., Knüpfer, U., Wenderoth, R., Nieba, L., Proba, K. & Riesenberg, D. (1996). Producing antibodies in *Escherichia coli*: From PCR to fermentation. In *Antibody Engineering* (McCafferty, J., Hoogenboom, H. R. & Chiswell, D. J., eds.), pp. 203–252. IRL Press, Oxford.

80 Honegger, A. & Plückthun, A. (2001). The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains. *J. Mol. Biol.* 309, 687–99.

81 Jung, S., Spinelli, S., Schimmele, B., Honegger, A., Pugliese, L., Cambillau, C. & Plückthun, A. (2001). The importance of framework residues H6, H7 and H10 in antibody heavy chains: experimental evidence for a new structural subclassification of antibody V_H domains. *J. Mol. Biol.* 309, 701–16.

82 Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. (1990). A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* 29, 1362–7.

83 Jung, S. H., Pastan, I. & Lee, B. (1994). Design of interchain disulfide bonds in the framework region of the Fv fragment of the monoclonal antibody B3. *Proteins* 19, 35–47.

84 Wörn, A. & Plückthun, A. (1998). Mutual stabilization of V_L and V_H in single-chain antibody fragments, investigated with mutants engineered for stability. *Biochemistry* 37, 13120–7.

85 Young, N. M., MacKenzie, C. R., Narang, S. A., Oomen, R. P. & Baenziger, J. E. (1995). Thermal stabilization of a single-chain Fv antibody fragment by introduction of a disulphide bond. *FEBS Lett.* 377, 135–9.

86 Arkin, M. R. & Wells, J. A. (1998). Probing the importance of second sphere residues in an esterolytic antibody by phage display. *J. Mol. Biol.* 284, 1083–94.

87 Low, N. M., Holliger, P. H. & Winter, G. (1996). Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *J. Mol. Biol.* 260, 359–68.

88 Stemmer, W. P. (1994). Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370, 389–91.

89 Zhao, H., Giver, L., Shao, Z., Affholter, J. A. & Arnold, F. H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* 16, 258–61.

90 Amstutz, P., Forrer, P., Zahnd, C. & Plückthun, A. (2001). In vitro display technologies: novel developments and applications. *Curr. Opin. Biotechnol.* 12, 400–5.

91 Forrer, P., Jung, S. & Plückthun, A. (1999). Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins. *Curr. Opin. Struct. Biol.* 9, 514–20.

92 Zhao, H. & Arnold, F. H. (1997). Combinatorial protein design: strategies for screening protein libraries. *Curr. Opin. Struct. Biol.* 7, 480–5.

93 Lehmann, M., Pasamontes, L., Lassen, S. F. & Wyss, M. (2000). The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta* 1543, 408–415.

94 Mattheakis, L. C., Bhatt, R. R. & Dower, W. J. (1994). An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl. Acad. Sci. U.S.A.* 91, 9022–6.

95 Hanes, J. & Plückthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. U.S.A.* 94, 4937–4942.

96 He, M. & Taussig, M. J. (1997). Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for in vitro display and evolution of antibody combining sites. *Nucleic Acids Res* 25, 5132–4.

97 Roberts, R. W. & Szostak, J. W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U.S.A.* 94, 12297–302.

98 Kurz, M., Gu, K., Al-Gawari, A. & Lohse, P. A. (2001). cDNA – protein fusions: covalent protein–gene conjugates for the in vitro selection of peptides and proteins. *ChemBioChem* 2, 666–72.

99 Tawfik, D. S. & Griffiths, A. D. (1998). Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* 16, 652–6.

100 Doi, N. & Yanagawa, H. (1999). STABLE: protein-DNA fusion system for screening of combinatorial protein libraries in vitro. *FEBS Lett.* 457, 227–30.

101 Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228, 1315–7.

102 Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. & Curtiss, R., 3rd. (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotechnol.* 15, 29–34.

103 Boder, E. T. & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* 15, 553–7.

104 Jung, S. & Plückthun, A. (1997). Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng.* 10, 959–66.

105 Hoess, R. H. (2001). Protein design and phage display. *Chem. Rev.* 101, 3205–18.

106 Finucane, M. D., Tuna, M., Lees, J. H. & Woolfson, D. N. (1999). Core-directed protein design. I. An experimental method for selecting

stable proteins from combinatorial libraries. *Biochemistry* 38, 11604–12.

107 Sieber, V., Plückthun, A. & Schmid, F. X. (1998). Selecting proteins with improved stability by a phage-based method. *Nat. Biotechnol.* 16, 955–60.

108 Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* 340, 245–6.

109 Bai, C. & Elledge, S. J. (1996). Gene identification using the yeast two-hybrid system. *Methods Enzymol.* 273, 331–47.

110 Johnsson, N. & Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 91, 10340–4.

111 Michnick, S. W. (2001). Exploring protein interactions by interaction-induced folding of proteins from complementary peptide fragments. *Curr. Opin. Struct. Biol.* 11, 472–7.

112 Waldo, G. S. (2003). Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.* 7, 33–8.

113 Wigley, W. C., Stidham, R. D., Smith, N. M., Hunt, J. F. & Thomas, P. J. (2001). Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. *Nat. Biotechnol.* 19, 131–6.

114 Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* 17, 691–5.

115 Pedelacq, J. D., Piltch, E., Liong, E. C., Berendzen, J., Kim, C. Y., Rho, B. S., Park, M. S., Terwilliger, T. C. & Waldo, G. S. (2002). Engineering soluble proteins for structural genomics. *Nat. Biotechnol.* 20, 927–32.

116 Wurth, C., Guimard, N. K. & Hecht, M. H. (2002). Mutations that reduce aggregation of the Alzheimer's Aβ42 peptide: an unbiased search for the sequence determinants of Aβ amyloidogenesis. *J. Mol. Biol.* 319, 1279–90.

117 Maxwell, K. L., Mittermaier, A. K., Forman-Kay, J. D. & Davidson, A. R. (1999). A simple in vivo assay for increased protein solubility. *Protein Sci.* 8, 1908–11.

118 Lesley, S. A., Graziano, J., Cho, C. Y., Knuth, M. W. & Klock, H. E. (2002). Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng.* 15, 153–60.

119 Minard, P., Scalley-Kim, M., Watters, A. & Baker, D. (2001). A "loop entropy reduction" phage-display selection for folded amino acid sequences. *Protein Sci.* 10, 129–34.

120 Matsuura, T. & Plückthun, A. (2003). Selection based on the folding properties of proteins with ribosome display. *FEBS Lett.* 539, 24–8.

121 Gu, H., Yi, Q., Bray, S. T., Riddle, D. S., Shiau, A. K. & Baker, D. (1995). A phage display system for studying the sequence determinants of protein folding. *Protein Sci.* 4, 1108–17.

122 Kim, D. E., Gu, H. & Baker, D. (1998). The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4982–6.

123 Kowalski, J. M., Parekh, R. N., Mao, J. & Wittrup, K. D. (1998). Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control. *J. Biol. Chem.* 273, 19453–8.

124 Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M. & Wittrup, K. D. (1999). Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J. Mol. Biol.* 292, 949–56.

125 Shusta, E. V., Holler, P. D., Kieke, M. C., Kranz, D. M. & Wittrup, K. D. (2000). Directed evolution of a stable scaffold for T-cell receptor engineering. *Nat. Biotechnol.* 18, 754–9.

126 Boder, E. T. & Wittrup, K. D. (2000). Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol.* 328, 430–44.

127 Ruan, B., Hoskins, J., Wang, L. & Bryan, P. N. (1998). Stabilizing the subtilisin BPN′ pro-domain by phage display selection: how restrictive is the amino acid code for maximum protein stability? *Protein Sci.* 7, 2345–53.

128 Zhou, H. X., Hoess, R. H. & DeGrado, W. F. (1996). In vitro evolution of thermodynamically stable turns. *Nat. Struct. Biol.* 3, 446–51.

129 Goto, Y. & Hamaguchi, K. (1979). The role of the intrachain disulfide bond in the conformation and stability of the constant fragment of the immunoglobulin light chain. *J. Biochem. (Tokyo).* 86, 1433–41.

130 Akanuma, S., Yamagishi, A., Tanaka, N. & Oshima, T. (1999). Further improvement of the thermal stability of a partially stabilized Bacillus subtilis 3-isopropylmalate dehydrogenase variant by random and site-directed mutagenesis. *Eur. J. Biochem.* 260, 499–504.

131 Farinas, E. T., Bulter, T. & Arnold, F. H. (2001). Directed enzyme evolution. *Curr. Opin. Biotechnol.* 12, 545–51.

132 Zhao, H. & Arnold, F. H. (1999). Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng.* 12, 47–53.

133 Giver, L., Gershenson, A., Freskgard, P. O. & Arnold, F. H. (1998). Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12809–13.

134 Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Plückthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. *Proc. Natl. Acad. Sci. U.S.A.* 98, 75–80.

135 Jung, S., Honegger, A. & Plückthun, A. (1999). Selection for improved protein stability by phage display. *J. Mol. Biol.* 294, 163–80.

136 Kristensen, P. & Winter, G. (1998). Proteolytic selection for protein folding using filamentous bacteriophages. *Fold. Des.* 3, 321–8.

137 Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E. & Zambonin, M. (1997). Probing the partly folded states of proteins by limited proteolysis. *Fold. Des.* 2, R17–26.

138 Krebber, C., Spada, S., Desplancq, D. & Plückthun, A. (1995). Co-selection of cognate antibody-antigen pairs by selectively-infective phages. *FEBS Lett.* 377, 227–31.

139 Finucane, M. D. & Woolfson, D. N. (1999). Core-directed protein design. II. Rescue of a multiply mutated and destabilized variant of ubiquitin. *Biochemistry* 38, 11613–23.