[16] Consensus-Based Engineering of Protein Stability: From Intrabodies to Thermostable Enzymes

By BORIS STEIPE

Stability Engineering

Protein stability* is not only of interest for biotechnology. Because the function of a protein depends on its structure, and the native structure is presumed to be acquired as the nascent peptide traverses a funnel-like energy landscape to descend into a unique global minimum, we can argue that stability—the goal and driving force of this process—is an experimental metric of the information encoded in a protein sequence. In 1994 we proposed an approach to predict stabilizing mutations in proteins, based on the analysis of sequence profiles.[1] This approach has allowed us to engineer hyperstable immunoglobulin domains that can be expressed as intrabodies. The procedure is entirely general, requiring neither knowledge of protein structure nor elaborate computational procedures; only a number of closely related homologous sequences are needed. Its application by a number of groups to a diverse set of proteins has demonstrated this procedure to be the most reliable strategy for the rational generation of stabilizing mutations. Importantly, the method emphasizes the stochastic nature of residue interactions in natural proteins and thus contributes to our understanding of the principles that govern sequence-folding relationships.

Canonical Sequence Approximation

It is not obvious why it should be possible to improve protein function by consensus mutations, one could equally well argue that every protein might have evolved to acquire a unique set of cooperative, favorable residue interactions and thus consensus mutations must be destabilizing as they would not take this specific context into account. Antibody domains are nature's workbench for the evolutionary engineering of affinity and an excellent model system with which to study the question of consensus versus uniqueness of the encoding of folding information in protein sequence. Clusters of variable residues on a conserved framework of

* Protein stability in this manuscript refers to thermodynamic stability: the free energy difference between the ideally unfolded state and the native state, determined at equilibrium.

[1] B. Steipe, B. Schiller, A. Plückthun, and S. Steinbacher, *J. Mol. Biol.* **240,** 188 (1994).

structure[2] arise from a process of domain hypermutation and selection of successful variants in the B cell.[3,4] This process of affinity maturation ultimately selects unique, high-affinity, high-specificity variants, but the measure of fitness in this process must be a composite of factors that include antigen affinity, domain stability, assembly and interaction of the heavy-and light-chain, variable, and constant domains, protease resistance, and competence for export and secretion. In order to bind its antigen, the antibody must fold first. The relative importance of any residue for these factors depends on its position in the sequence; only the stability requirement has to be fulfilled in every position, as disruptive mutations are eliminated from the pool of functional proteins.

One can model this situation by imagining a large pool of evolved immunoglobulin sequences, each one optimized individually for antigen binding in a stochastic process of mutation and selection to conserve a minimal folding stability and each one folding into approximately the same global conformation. This model suggests an analogy to a thermodynamic canonical ensemble. The individual sequence positions are independent components of a system and can occur in any number of microstates, corresponding to the nature of the individual residue, but the whole system must satisfy global constraints. Accordingly, the distribution of microstates can be analyzed individually in terms of the constraints. Note that we are not explicitly considering covariation in this analysis but treating residues independently. This makes the canonical sequence approximation technically a mean-field approach. Applying Boltzmann's law to such an ensemble predicts that the states that occur most frequently possess the highest fitness of all alternatives. In the case of immunoglobulins, such an averaged ensemble would not include antigen affinity as a component of the fitness function, as we are averaging over all possible antigens. Accordingly, factors that determine the structure, specifically thermodynamic stability, play a dominant role. In the case of enzymes, particularly when using sets of orthologous sequences, functional requirements are represented as well.

The key assumptions made are residues evolve randomly and independently; the collection of sequences analyzed is an unbiased sample of the sequence space of the protein in which each domain is at an equilibrium of random stabilizing and destabilizing mutations; and a requirement for minimum thermodynamic stability is the dominant factor that constrains variability at each position.

[2] T. T. Wu and E. A. Kabat, *J. Exp. Med.* **132,** 211 (1970).
[3] S. Tonegawa, *Nature* **302,** 575 (1983).
[4] C. Berek and M. Ziegner, *Immunol. Today* **8,** 400 (1993).

To the degree that these assumptions hold, the prediction of stabilizing mutations follows a simple rule: a residue that replaces another that is observed significantly less frequently in a set of related sequences is expected to stabilize the protein. Furthermore, individual replacements can be combined, as independent effects are expected to be additive.

Immunoglobulin Domain Engineering

As a proof of principle, we have predicted point mutations in the well-studied V_κ[1] domain of the antibody McPC603[1] using a manually edited alignment of immunoglobulin sequences. We have shown that predictions of stabilizing mutations that replace native residues with the V_κ consensus residue were correct in 6 of 10 cases, three replacements were neutral and only one (Q79E) was found to be somewhat destabilizing. Random mutations are usually estimated to produce a large excess of destabilizing mutations, on the order of 100:1.[5] We have noted that stability effects are equally well predicted in framework regions and in complementarity-determining regions (CDR). Predictions failed predominantly when they were made for residues that would participate in interdomain interactions in the heterodimeric antibody, such as P43S,[†] Q79E, and A100G. Presumably these would be conserved for interaction requirements, rather than global domain stability in the context of the isolated V_κ domain. Had we used our knowledge of antibody structure to exclude such residues from the prediction, the success rate for mutations that stabilize the isolated domain would increase to 7 out of 8. However, the strategy of simply choosing the most prevalent amino acid in every position was found to carry a penalty of less than 2 kJ mol^{-1} at worst. Moreover, an improvement of Fv fragment association might offset the slight loss of folding stability we have observed, while for practical reasons it also might be a goal for optimization in its own right.

Interestingly, loop lengths of the first complementarity determining region of the V_κ domain are also variable and observed with widely differing frequencies in the database. Changing the native loop length to the consensus length further stabilizes the domain.[6] Thus it is not only residue identity that can be analyzed in terms of consensus predictions, but other features as well, as long as they can be quantified in terms of their frequency of occurrence in a set of related sequences.

[5] F. C. Christians, L. Scapozza, A. Crameri, G. Folkers, and W. P. C. Stemmer, *Nature Biotechnol.* **17,** 259 (1999).

[†] To specify mutations, we use the one-letter code for the native residue, followed by the sequence position, followed by the one-letter code for the new residue.

[6] E. C. Ohage and B. Steipe, *J. Mol. Biol.* **291,** 1119 (1999).

Construction of Intrabodies

One of the hallmarks of the immunoglobulin fold is a strictly conserved intramolecular disulfide bridge. Such disulfide bridges normally cannot be formed in the reducing environment of the cytoplasm.[7] Estimating the contribution of the disulfide bond to stability from its effect on the reduction of entropy in the unfolded state[8] and comparing it to typical stabilities of native immunoglobulin domains show that this disulfide bridge alone would contribute more than the net free energy of folding to the stability of each domain. The consequences have been observed experimentally many times over: attempts to express functional antibody domains in a reducing environment fail in general[9] and chemical reduction denatures the domain *in vitro*. While exceptions have been reported[10–12] and, for many practical applications, the problem has been overcome through the clever application of randomization and screening strategies,[13–15] by and large low stability, low expression rates, and unpredictable behavior have limited the application of functional intrabodies.[16–18]

In a study aimed to overcome this problem by rational engineering, we have combined individually stabilizing mutations and showed that this achieves additive stabilization. The hyperstable V_κ we have obtained with this strategy can be expressed in the cytoplasm, solubly and with good yield.[6] As well, predicted point mutations in a V_H domain were similarily stabilizing,[19] allowing us to construct a framework for loop grafting of novel specificities. This strategy led to the successful expression of a

[7] S. Biocca, F. Ruberti, M. Tafani, A. P. Pierandrei, and A. Cattaneo, *Biotechnology* **13,** 1110 (1995).
[8] C. N. Pace, G. R. Grimsley, J. A. Thomson, and B. J. Barnett, *J. Biol. Chem.* **263,** 11820 (1988).
[9] R. Glockshuber, T. Schmidt, and A. Plückthun, *Biochemistry* **31,** 1270 (1992).
[10] J. P. Maciejewski, F. F. Weichold, N. S. Young, A. Cara, D. Zella, M. J. Reitz, and R. C. Gallo, *Nature Medicine* **1,** 667 (1995).
[11] Y. Wu, L. Duan, M. Zhu, B. Hu, S. Kubota, O. Bagasra, and R. J. Pomerantz, *J. Virology* **70,** 3290 (1996).
[12] I. J. Rondon and W. A. Marasco, *Annu. Rev. Microbiol.* **51,** 257 (1997).
[13] A. Auf der Maur, C. Zahnd, F. Fischer, S. Spinelli, A. Honegger, C. Cambillau, D. Escher, A. Plückthun, and A. Barberis, *J. Biol. Chem.* **277,** 45075 (2002).
[14] T. Tanaka and T. H. Rabbitts, *EMBO J.* **22,** 1025 (2003).
[15] F. Gennari, S. Mehta, Y. Wang, A. St. Clair Tallarico, G. Palu, and W. A. Marasco, *J. Mol. Biol.* **335,** 193 (2004).
[16] W. A. Marasco, *Gene Ther.* **4,** 11 (1997).
[17] N. Gargano and A. Cattaneo, *FEBS Lett.* **414,** 537 (1997).
[18] A. Cattaneo and S. Biocca, *Trends Biotech.* **17,** 115 (1999).
[19] P. Wirtz and B. Steipe, *Protein Sci.* **8,** 2245 (1999).

heterodimeric, catalytic intrabody Fv fragment by loop grafting of the CDR regions of the esterolytic antibody 17E8 into our stabilized framework.[20]

Other Examples

A number of similarly successful studies applying this principle have underscored the generality of our approach for the rational engineering of immunoglobulins for a wide range of objectives, from biotechnology to immunotherapy and ranging from individual domains to whole antibodies. These studies include the construction of frameworks for the display of randomized CDR libraries of scFvs,[21] improved production of a Diels–Alder catalytic antibody,[22] stabilization of an internalizing anti-CD22 scFv, with potential application for non-Hodgkin lymphoma therapy,[23] engineering for improved expression yield of an scFv-β-lactamase fusion protein for antibody-directed prodrug activation cancer therapy,[24] improved secretion of whole antibodies in mammalian cells,[25] and thermostabilization of immunoglobulin C_{H3} constant domains.[26]

Due to the generality of the approach, application of the canonical sequence approximation to other proteins is equally straightforward in principle. Successful examples include the stabilization of GroEL minichaperones,[27] stabilization of p53,[28] and SH3 domains.[29]

Perhaps the most convincing example overall has been contributed from the engineering of thermostable phytases for animal feed technology. Lehmann *et al.*[30] compiled a multiple sequence alignment of 13 homologous fungal phytases as a starting point for consensus engineering. Instead of introducing individual point mutations, they generated the entire

[20] E. C. Ohage, P. Wirtz, J. Barnikow, and B. Steipe, *J. Mol. Biol.* **291,** 1129 (1999).
[21] A. Knappik, L. M. Ge, A. Honegger, P. Pack, M. Fischer, G. Wellnhofer, A. Hoess, J. Wolle, A. Plückthun, and B. Virnekäs, *J. Mol. Biol.* **296,** 57 (2000).
[22] A. Piatesi and D. Hilvert, *Can. J. Chem.* **80,** 657 (2002).
[23] M. A. E. Arndt, J. Krauss, R. Schwarzenbacher, B. K. Vu, S. Greene, and S. M. Rybak, *Intl. J. Cancer* **107,** 822 (2003).
[24] C. F. McDonagh, K. S. Beam, G. J. Wu, J. H. Chen, D. F. Chace, P. D. Senter, and J. A. Francisco, *Bioconjug. Chem.* **14,** 860 (2003).
[25] E. A. Whitcomb, T. M. Martin, and M. B. Rittenberg, *J. Immunol.* **170,** 1903 (2003).
[26] S. J. Demarest, J. Rogers, and G. Hansen, *J. Mol. Biol.* **335,** 41 (2004).
[27] Q. H. Wang, A. M. Buckle, N. W. Foster, C. M. Johnson, and A. R. Fersht, *Protein Sci.* **8,** 2186 (1999).
[28] P. V. Nikolova, J. Henckel, D. P. Lane, and A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **95,** 14675 (1998).
[29] K. L. Maxwell and A. R. Davidson, *Biochemistry* **37,** 16172 (1998).
[30] M. Lehmann, L. Pasamontes, S. F. Lassen, and M. Wyss, *Biochim. Biophys. Acta* **1543,** 408 (2000).

consensus sequence by gene synthesis. This synthetic protein was not only 15 to 26° more thermostable than any of the progenitor sequences, it was also enzymatically as active. In a follow-up study, the group included further sequence changes that were suggested by the inclusion of additional sequences into the alignment, characterized them individually with respect to the earlier consensus sequence, removed a small number of destabilizing mutations, and finally arrived at a phytase with an unfolding temperature of 90.4°—an increase of almost 30° from the most stable natural phytase of *Aspergillus fumigatus*.[31] The effects of individual mutations are small. Large, global stabilization is achieved by the additive contributions of a number of independent improvements. This makes the strategy complementary to evolutionary engineering methods, which rely on large individual effects to create a screenable signal and which have problems generating saturating libraries with multiple substitutions due to the combinatorially large size of the sequence spaces involved.

The ability to generate synthetic, enzymatically active proteins from consensus sequences emphasizes the statistical nature of natural protein sequences as products of neutral drift and selective pressure. Stabilization is not the only objective for which the approach is useful: others could include the selection of residues for nondisruptive destabilizing mutations to reduce the lifetime of proteins or to generate proteins that can be used to screen for second-site revertants in evolutionary engineering strategies. They could also include guiding the choice of preferred replacements of residues that need to be changed for other purposes, such as engineering of electrostatic properties or alkaline stability.[32]

Work Flow

Consensus engineering is simple and successful predictions do not require elaborate tools. The work flow is represented in Fig. 1 and is discussed in this section.

Sequence Sets

A prerequisite for consensus engineering is the availability of homologous[‡] sequences at moderate to high degrees of similarity, and obviously the very large sequence collections of the immunoglobulins have been an excellent resource in this respect. However, given the rapid growth of sequence databases, it has become more and more likely that a sufficient

[31] M. Lehmann, C. Loch, A. Middendorf, D. Studer, S. F. Lassen, L. Pasamontes, A. P. G. M. van Loon, and M. Wyss, *Protein Eng.* **15,** 403 (2002).
[32] S. Gulich, M. Linhult, S. Stahl, and S. Hober, *Protein Eng.* **15,** 835 (2002).

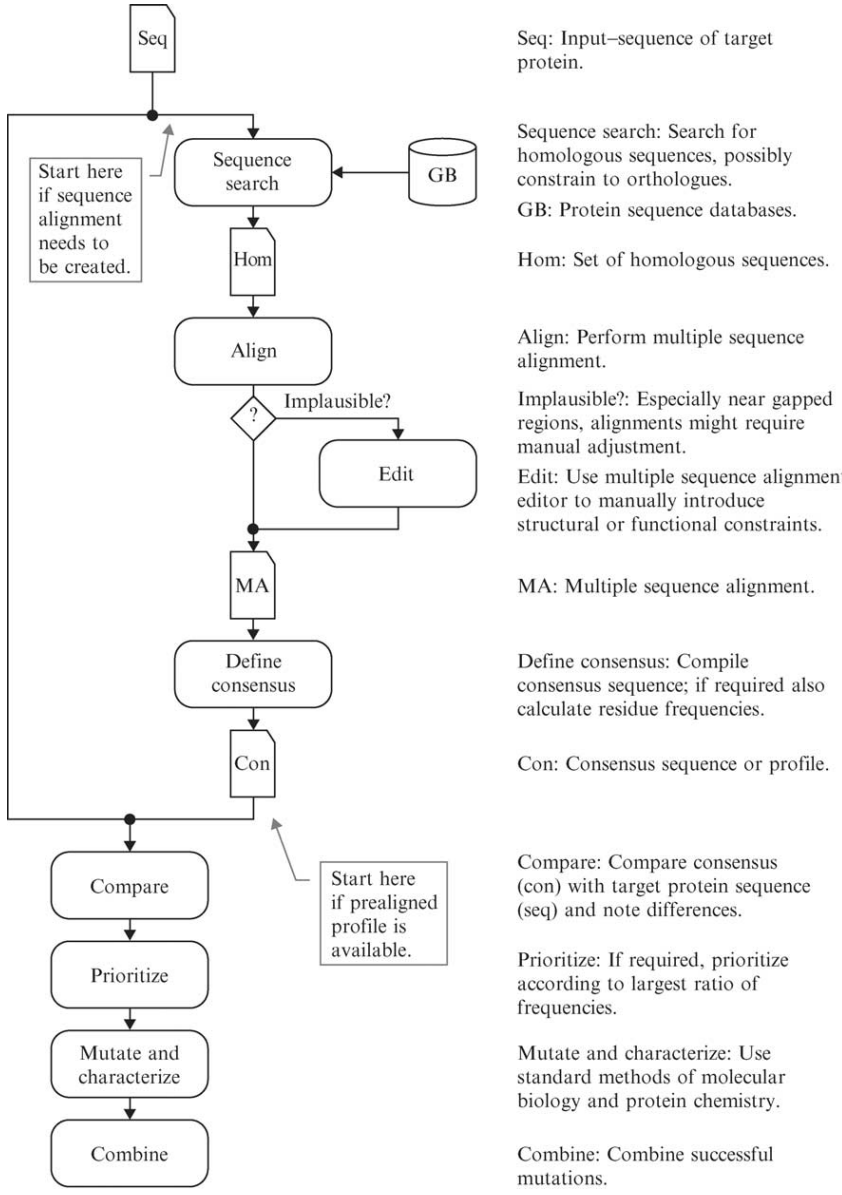

FIG. 1. Work flow diagram of consensus sequence engineering, describing procedures, principal intermediary results or data sets, and their sequential connection. Details of the individual steps are discussed in the text.

number of suitable sequences can be retrieved through homology searches for any protein of interest. Even where this is not possible, the option of randomly collecting homologous sequences from DNA pools by PCR and sequencing can be considered, as identification of the progenitor organism is not required for the procedure. The data sets need not be large, as long as a reasonably unbiased selection can be obtained. It is not the number of sequences, but the ratio of frequencies that is important.

Typically, a protein–protein BLAST search would be the starting point for the collection of homologous sequences. A standard BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) with default parameters should be sufficient to retrieve the sequences of interest. We would preferably not use very divergent sequences retrieved from highly sensitive search strategies, such as PSI-BLAST, as these sequences require introducing additional assumptions about the conserved structural context of individual positions and conserved function.

Sequences can be selected on the output Web form, retrieved and saved to disk, to be used later as input for multiple alignment.

If orthologous sequences are required, e.g., for the engineering of enzymes, three main options are available to select appropriate sequences.

i. Annotations of function can be retrieved from the file headers on the BLAST results page. This is fast, but carries the risk of false positives since function may have been annotated from sequence similarity alone, as well as false negatives, since annotations may be missing, wrong, or using unfamiliar synonyms.

ii. One can attempt to demonstrate that the homologous sequences of interest are reciprocal best matches in their respective genomes. This is tedious, but would give the most convincing support. Obviously, this requires that complete genomes are available that include all sequences that are to be analyzed.

iii. If available, sequences can be obtained from the Clusters of Orthologous Groups (COGs) database (http://www.ncbi.nlm.nih.gov/COG/), which stores several thousand precompiled clusters of reciprocal best matches of genes in bacterial genomes.[33] Note that the source

[‡] Homologous sequences have diverged from a common ancestral sequence. Typically homology is inferred from observing high degrees of sequence similarity, but structural alignments of distantly related proteins may serve as well to define alignments. Orthologous sequences have the same function in different organisms, while paralogous sequences have divergent functions, usually as the result of gene duplications.

[33] R. L. Tatusov, D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram, B. S. Rao, B. Kiryutin, M. Y. Galperin, N. D. Fedorova, and E. V. Koonin, *Nucleic Acids Res.* **29,** 22 (2001).

sequences, as well as complete CLUSTAL multiple alignments, can be retrieved from this site.

Alignment

As a next step, multiple sequence alignments are required. A number of options are available; the curated bioinformatics links directory of the Canadian Bioinformatics Workshops provides a good overview (http://www.bioinformatics.ca/). Excellent resources are offered through Web interfaces and do not require local installation of alignment tools. CLUSTAL W has long been the method of first choice (http://www.ebi.ac.uk/clustalw/) but it has become apparent that the T-Coffee algorithm (http://www.ch.embnet.org/software/TCoffee.html) produces alignments that appear biologically more meaningful, especially for sequences that are not closely related.

Multiple alignment algorithms optimize a global score across an alignment. This score is typically based on sequence similarity alone and does not take structural or functional information into account. CLUSTAL can be an exception, as it allows the input of templates that guide alignments to be compatible with knowledge about secondary structure. Accordingly, in many cases, visual inspection of the generated alignment will suggest the need for manual edits to better represent biochemical knowledge about the target protein. CINEMA (http://bioinf.man.ac.uk/dbbrowser/CINEMA2.1/) and JALVIEW (http://www2.ebi.ac.uk/~michele/jalview/) are multiple alignment editors that have been in use for quite some time. An interesting newer alternative in terms of simplicity of access may be GoCore (http:// www.helsinki.fi/project/ritvos/GoCore/), which has been programmed as an Excel plug-in.

Prealigned Profiles

Where available, prealigned sequence collections may be the easiest source to guide consensus engineering projects.

For immunoglobulin variable domains, we have compiled sequence profiles for the major immunoglobulin subtypes based on the machine-readable version of the fifth edition of the Kabat database of immunoglobulin sequences.[34] The sequence alignments have been edited manually to align insertions into the positions where they have been determined crystallographically. The frequency profiles have been posted on the Web (http://biochemistry.utoronto.ca/steipe/research/canonical.html) and allow one to determine the consensus residue itself as well as obtain an

[34] E. A. Kabat, T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller, "Distribution Files of the Fifth Edition of Sequences of Proteins of Immunological Interest," 1992.

overview of how strongly constrained a position is; they can be used to calculate frequency ratios.

Many protein domains have been defined by bioinformatics analysis of sequence databases; these are being made available via the conserved domain database (CDD) at the NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd./cdd.shtml). The display of CDD profiles can be customized through a Web interface. Consensus sequences are calculated and variability can be estimated visually from the alignments or tabulated manually. As a *caveat*, the domain alignments are built primarily for the purpose of database searches and frequently do not extend all the way to the domain termini. In many cases it will thus be desirable to construct a custom multiple alignment. Even for this purpose, the CDD is an efficient way to retrieve source sequences, as the corresponding GenPept records are linked to the sequences displayed on the page. The use of COGs for this purpose was discussed previously.

Generating Consensus Sequences

Consensus sequences are included in our immunoglobulin sequence alignment, with CDDs, or can be created from CINEMA or GoCore output. When the consensus sequence is not obvious (e.g., for very large alignments), the EMBOSS package of molecular biology tools provides a number of options to process multiple alignments from locally installed tools or online (http://bioinfo.pbi.nrc.ca/EMBOSS/). To compile specific frequency distributions, a Perl program can be obtained from the author of this article.

Sequence Prioritizations

The strongest assumption of our approach is that thermodynamic stability is the dominating component of selective pressure in the positions we would like to mutate. We have noted earlier that this is not necessarily the case for residues that contribute to interdomain interfaces. Generating a homology model (e.g., with the automated SwissModel Web server http://www.expasy.org/swissmod/SWISS-MODEL.html) can help distinguish such residues and prioritize mutations, if necessary. However, there is a second class of residues for which the consensus rule explicitly will not hold: residues involved in the active sites of enzymes. Studies on a number of enzymes suggest a partition of amino acid roles into structural and functional residues, as was demonstrated by Beadle and Shoichet[35] (see also references contained therein). Most residues appear to contribute to structural frameworks through favorable interactions, but a few provide

[35] B. M. Beadle and B. K. Schoichet, *J. Mol. Biol.* **321,** 285 (2002).

specific functions in precisely controlled conformations, frequently at the cost of steric or electrostatic strain, exposed hydrophobic surface, or unsatisfied hydrogen-bonding donors or acceptors. These functional residues are highly conserved, yet significant increases of stability can be gained by replacing them. While this is important to recognize in principle, it is a nonissue for the optimization of enzymes as the example of phytase engineering has shown, since presumably the source sequence has the same functional residues conserved as its homologues, especially if these are orthologues, and the enzymatic activity likely needs to be conserved in the engineering process. Where this is not the case, and stabilization without concern for activity is the objective, other methods than consensus engineering are required. We have had some success in this respect with an extension of the canonical sequence approximation to the convergent evolution of nonhomologous structure motif consensus sequences.[36,37]

Possible mutations may be prioritized. In keeping with the *canonical sequence approximation*, we would not only consider the absolute frequency of the consensus residue but also the ratio of frequencies of the consensus residue and the original residue. If this frequency ratio is used for quantitative predictions as $\Delta\Delta G = -RT\ (\ln f_{\text{consensus}} - \ln f_{\text{original}})$, we typically achieve coefficients of correlation of better than 0.6 (B. Steipe, unpublished), although, given the assumptions that go into the model, such a quantitative treatment has to be treated with caution and, for practical purposes, qualitative predictions are sufficient. The final steps of mutagenesis, characterization and combination of successful residues, are specific for each experimental system at hand and should be straightforward.

Summary

The excellent track record of consensus sequence engineering, together with the ease of obtaining predictions, makes this strategy a method of first choice for biotechnological purposes, as well as for the study of sequence–structure relationships in proteins. The method is general, complements strategies of evolutionary engineering well, and can be combined with alternative methods of stability prediction. It is applicable if nothing but sequence is known for a protein and it can be refined to consider structural and functional knowledge.

[36] E. C. Ohage, W. Graml, M. M. Walter, S. Steinbacher, and B. Steipe, *Protein Sci.* **6,** 233 (1997).
[37] M. Niggemann and B. Steipe, *J. Mol. Biol.* **296,** 181 (2000).

ScienceDirect

Methods in Enzymology

Volume 388, 2004, Pages 176-186

Consensus-Based Engineering of Protein Stability: From Intrabodies to Thermostable Enzymes

Boris Steipe

https://doi.org/10.1016/S0076-6879(04)88016-9 ↗

Get rights and content ↗

Publisher Summary

This chapter elaborates the consensus-based engineering of protein stability. Protein stability is an experimental metric of the information encoded in a protein sequence. This approach has allowed to engineer hyperstable immunoglobulin domains that can be expressed as intrabodies. Its application by a number of groups to a diverse set of proteins has demonstrated this procedure to be the most reliable strategy for the rational generation of stabilizing mutations. The ability to generate synthetic enzymatically active proteins from consensus sequences emphasizes the statistical nature of natural protein sequences as products of neutral drift and selective pressure. Stabilization is not the only objective for which the approach is useful, and others could include the selection of residues for nondisruptive destabilizing mutations to reduce the lifetime of proteins, or to generate proteins that can be used to screen for second-site revertants in evolutionary engineering strategies. It is found that the final steps of mutagenesis, characterization, and combination of successful residues are specific for each experimental system at hand and should be straightforward.

Section snippets

Stability Engineering

Protein stability* is not only of interest for biotechnology. Because the function of a protein depends on its structure, and the native structure is presumed to be acquired as the nascent peptide traverses a funnel-like energy landscape to descend into a unique global minimum, we can argue that stability—the goal …

Canonical Sequence Approximation

It is not obvious why it should be possible to improve protein function by consensus mutations, one could equally well argue that every protein might have evolved to acquire a unique set of cooperative, favorable residue interactions and thus consensus mutations must be destabilizing as they would not take this specific context into account. Antibody domains are nature's workbench for the evolutionary engineering of affinity and an excellent model system with which to study the question of …

Immunoglobulin Domain Engineering

As a proof of principle, we have predicted point mutations in the well-studied V_κ[1] domain of the antibody McPC603[1] using a manually edited alignment of immunoglobulin sequences. We have shown that predictions of stabilizing mutations that replace native residues with the V_κ consensus residue were correct in 6 of 10 cases, three replacements were neutral and only one (Q79E) was found to be somewhat destabilizing. Random mutations are usually estimated to produce a large excess of destabilizing …

Construction of Intrabodies

One of the hallmarks of the immunoglobulin fold is a strictly conserved intramolecular disulfide bridge. Such disulfide bridges normally cannot be formed in the reducing environment of the cytoplasm.[7] Estimating the contribution of the disulfide bond to stability from its effect on the reduction of entropy in the unfolded state[8] and comparing it to typical stabilities of native immunoglobulin domains show that this disulfide bridge alone would contribute more than the net free energy of folding …

Other Examples

A number of similarly successful studies applying this principle have underscored the generality of our approach for the rational engineering of immunoglobulins for a wide range of objectives, from biotechnology to immunotherapy and ranging from individual domains to whole antibodies. These studies include the construction of frameworks for the display of randomized CDR libraries of scFvs,[21] improved production of a Diels–Alder catalytic antibody,[22] stabilization of an internalizing anti-CD22 …

Work Flow

Consensus engineering is simple and successful predictions do not require elaborate tools. The work flow is represented in Fig. 1 and is discussed in this section. …

Summary

The excellent track record of consensus sequence engineering, together with the ease of obtaining predictions, makes this strategy a method of first choice for biotechnological purposes, as well as for the study of sequence–structure relationships in proteins. The method is general, complements strategies of evolutionary engineering well, and can be combined with alternative methods of stability prediction. It is applicable if nothing but sequence is known for a protein and it can be refined to …

Volume chapters Recommended articles

References (37)

B Steipe *et al.*
J. Mol. Biol. (1994)

C Berek *et al.*
Immunol. Today (1993)

E.C Ohage *et al.*
J. Mol. Biol. (1999)

C.N Pace *et al.*
J. Biol. Chem. (1988)

F Gennari *et al.*
J. Mol. Biol. (2004)

N Gargano *et al.*
FEBS Lett. (1997)

A Cattaneo *et al.*
Trends Biotech. (1999)

E.C Ohage *et al.*
J. Mol. Biol. (1999)

A Knappik *et al.*
J. Mol. Biol. (2000)

S.J Demarest *et al.*
J. Mol. Biol. (2004)

View more references

Cited by (64)

A broad range of Fab stabilities within a host of therapeutic IgGs

2007, Biochemical and Biophysical Research Communications

Citation Excerpt :

...We first investigated the level of unusual amino acid usage within the BIIB1–17 V-gene frameworks and CDR1/CDR2 regions resulting from the germline or from hypersomatic mutation. Protein engineering using consensus approaches supports the idea that mutating to common amino acid types found in multiple members of a protein family and away from unusual amino acid types can improve

> stability [36,41–44]. High levels of infrequently used amino acids within a particular fold may limit protein stability....

Show abstract

The Crucial Role of Methodology Development in Directed Evolution of Selective Enzymes ↗

2020, Angewandte Chemie - International Edition

Design and production of bispecific antibodies ↗

2019, Antibodies

Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability ↗

2019, Chemical Reviews

Experimental evidence for the thermophilicity of ancestral life ↗

2013, Proceedings of the National Academy of Sciences of the United States of America

Antibody therapeutics, antibody engineering, and the merits of protein stability ↗

2008, Current Opinion in Drug Discovery and Development

View all citing articles on Scopus ↗

View full text