

Thermostability of membrane protein helix–helix interaction elucidated by statistical analysis

Dirk Schneider¹, Yang Liu¹, Mark Gerstein, Donald M. Engelman*

Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, New Haven, CT 06520-8114, USA

Received 14 October 2002; revised 1 November 2002; accepted 4 November 2002

First published online 14 November 2002

Edited by Gunnar von Heijne

Abstract A prerequisite for the survival of (micro)organisms at high temperatures is an adaptation of protein stability to extreme environmental conditions. In contrast to soluble proteins, where many factors have already been identified, the mechanisms by which the thermostability of membrane proteins is enhanced are almost unknown. The hydrophobic membrane environment constrains possible stabilizing factors for transmembrane domains, so that a difference might be expected between soluble and membrane proteins. Here we present sequence analysis of predicted transmembrane helices of the genomes from eight thermophilic and 12 mesophilic organisms. A comparison of the amino acid compositions indicates that more polar residues can be found in the transmembrane helices of thermophilic organisms. Particularly, the amino acids aspartic acid and glutamic acid replace the corresponding amides. Cysteine residues are found to be significantly decreased by about 70% in thermophilic membrane domains suggesting a non-specific function of most cysteine residues in transmembrane domains of mesophilic organisms. By a pair-motif analysis of the two sets of transmembrane helices, we found that the small residues glycine and serine contribute more to transmembrane helix–helix interactions in thermophilic organisms. This may result in a tighter packing of the helices allowing more hydrogen bond formation. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Thermostability; Sequence analysis; Transmembrane helix; Genome; Mesophilic; Thermophilic

1. Introduction

In contrast to soluble proteins, no rigorous analysis has been performed on integral membrane proteins of thermophiles. As with mesophiles, most isolated soluble enzymes from hyperthermophiles show maximal activities at temperatures close to the optimal growth temperatures of the host organisms [1]. Previous biophysical and computational studies have revealed many factors that contribute to the thermostability of soluble proteins, including hydrogen bonding, hydrophobic internal packing, tighter packing and salt bridge optimization (for a recent review see [2]). From these studies, it became obvious that there is no single or most preferred mechanism for increasing thermostability. The hydrophobic environment of the membrane restricts the possibilities for

adaptation of membrane proteins to high temperatures. It has been shown that the loops connecting the individual α -helices of membrane proteins contribute to the stability of the interactions of the individual transmembrane helices (TM) [3], enhancement of the interactions within these loops could increase the stability of helix–helix interactions without changing the helices themselves. It is also known that microorganisms adapt to high temperatures by changing their lipid composition [4]. Since the lipid environment contributes to the stability of membrane protein's helix–helix interactions [5,6], the different lipid composition might modulate the stability of the interaction of transmembrane α -helices. But, what of the contribution of the helices themselves?

Here we present amino acid composition analysis and a statistical pair-motif analysis of the TMs from eight thermophilic and 12 mesophilic organisms. The results indicate that hydrogen bonding and tighter packing of the TMs add extra stability to TM–TM interactions at higher temperatures.

2. Materials and methods

2.1. Dataset

Open reading frames from eight (hyper)thermophilic and 12 mesophilic bacteria from a proteome analysis database [7] (Table 1) were submitted to the TMHMM server [8] for TM prediction. To eliminate most signal sequences, predicted TM sequences were excluded if they had a charged residue within the first seven residues and a stretch of 14 residues with a GES hydrophobicity value under -1 kcal/mol [9,10].

For the prediction of pair motifs, the TM sequences were screened for homology by eliminating each sequence that was highly similar to another sequence in the same TMs, as described in detail by Senes et al. [11]. The pair analysis was performed on all combinations of amino acids separated by one to 10 residues. Pairs of amino acids at the positions i , $i+k$ are separated by $k-1$ residues. In order to limit the pair analysis to the core amino acids of the TMs, the analysis was performed using a hydrophobic window of 18 amino acids instead of the entire TM.

The statistical significance of an identified amino acid pair from the respective average expected value of this pair was calculated by the two-tailed integral of the probably pair distribution as discussed in detail previously [11].

The statistical significance of the difference in the amino acid composition of the thermophilic and mesophilic sample means was calculated by a one-tailed t -test. The number of degrees of freedom was 18. The t -test assumes a normally distributed sample. A less than 5% possibility ($|t|$ -values > 1.73) that the difference between the means is due to chance ($P < 0.05$) was accepted to be significant. At a 1% level of significance ($P < 0.01$) $|t|$ -values of < 2.55 were rejected.

3. Results and discussion

Using the TMHMM program, a total of 21 473 TMs for the

*Corresponding author: Fax: (1)-203-432 6381.

E-mail address: don@paradigm.csb.yale.edu (D.M. Engelman).

¹ These authors contributed equally to this work.

mesophilic set of proteins and 13 340 for the thermophilic organism were predicted. The average length of the predicted TMs was 22 for both sets of proteins (thermophilic and mesophilic), which is in good agreement with the average length of TMs predicted in other studies [11,12].

Previous studies of the amino acid composition of all proteins in an organism [2,13] as well as a study of the amino acid compositions of soluble proteins from mesophilic and thermophilic organisms [14] have already indicated trends for the mechanism of thermostabilization of proteins. In these studies an increase in the content of charged residues was found in thermophiles, mostly at the expense of uncharged polar residues [2,14]. In thermophilic α -helical proteins the charged residues are preferentially arranged in an $i, i+4$ helical repeat pattern, suggesting stabilization of the thermophilic proteins via intrahelical salt bridges [13]. However, this trend may not apply to the thermophilic membrane proteins, since acidic and basic amino acids are often uncharged in the membrane environment and therefore are unlikely to form salt bridges. However, the presence of uncharged basic and acid residues could still result in strong interactions of TMs [15,16]. By analyzing the localization of polar residues in the TMs of solved membrane protein structures, it has been shown that these residues form hydrogen bonds between TMs, resulting in strong interactions between pairs of TMs [17,18]. Although the content of ionizable residues is generally very low in the hydrophobic membrane environments, the amino acid analysis of the sets of mesophilic and thermophilic TMs showed that the compositions of the acidic residues, Asp and Glu, are increased in the TMs of thermophiles (Fig. 1) while those of the basic residues, Lys, His, and Arg, remains almost unchanged. As already observed in an overview of thermophilic proteins, the higher content of ionizable residues is coupled to the reduction of the uncharged polar residues Gln and Asn in the TMs. The reduction of the content of Asn and Gln could be explained by the fact that the amides are less stable under high temperature conditions and these residues can be deaminated to Asp and Glu. On the other hand, although Asp and Glu cannot contribute to the membrane protein's stability by

forming salt bridges, they can form stronger interhelical hydrogen bonds than Asn and Gln [16], resulting in a more stable packing between two TMs. The enhanced strength of the hydrogen bonds is most likely caused by the higher electronegativity of the oxygen atom in the acids compared to the nitrogen atom in the corresponding amides. Generally, TMs contain fewer ionizable residues, because the free energy of partitioning ionizable amino acids into the non-polar membrane environment is not favorable [19]. Nevertheless, the membrane incorporation of polar residues is thermodynamically more favorable under higher temperature conditions, which may explain the increased occurrence of the ionizable residues in TM helices of thermophilic organisms compared to the mesophilic TM domains. Since most proteomes of thermophiles are from archaea, we also cannot completely exclude the possibility that the observed differences in the composition of the polar residues is due to generic effects between archaea and bacteria.

A somewhat higher content of Tyr and Trp is found in the TMs of thermophiles (5.44%) versus the mesophiles (4.88%), and Tyr is used more than Trp. It is well known that these amino acids are frequently found near the boundary between the non-polar bilayer interior and the lipid headgroups, where they may stabilize TMs by a combination of non-polar and hydrogen bonding interactions with the bilayer [20–23]. Further, Tyr may contribute to individual helix stability, as it is able to form intrahelical hydrogen bonds [17,18].

We found that the content of Cys in TM helices from thermophilic organisms is strikingly decreased, to about 30% of its occurrence in mesophiles. Statistical analysis of the soluble proteins of mesophiles and thermophiles revealed that the occurrence of Cys is, on average, reduced by 10% in thermophiles [14]. Since most thermophiles used in this study are archaea, we compared the Cys content of the thermophilic eubacteria *Aquifex aeolicus* to the average Cys content of the mesophilic eubacteria. Compared to the mesophilic group, the Cys content in the membrane protein domains of *Aquifex* is decreased by over 80%, indicating that the observed difference in the Cys content is not caused by the analysis of the archaea

Table 1
Thermophilic and mesophilic organisms whose proteomes were used for TM prediction and analysis

Organism	Taxonomy	Growth temperature (°C)	Predicted TM helices
Thermophiles			
<i>Aeropyrum pernix K1</i>	Archaea: Crenarchaeota	90–95	1 637
<i>Aquifex aeolicus</i>	Bacteria: Aquificales	95	1 306
<i>Archaeoglobus fulgidus</i>	Archaea: Euryarchaeota	83	2 120
<i>Methanococcus jananashii</i>	Archaea: Euryarchaeota	85	1 143
<i>Methanobacterium thermoautotrophicum</i>	Archaea: Euryarchaeota	65	1 144
<i>Pyrococcus abyssi</i>	Archaea: Euryarchaeota	97	1 826
<i>Pyrococcus horikoshii</i>	Archaea: Euryarchaeota	98	1 932
<i>Thermoplasma acidophilum</i>	Archaea: Euryarchaeota	55	1 950
Mesophiles			
<i>Borellia burgdorferi</i>	Bacteria: Spirochaetales	37	848
<i>Bacillus subtilis</i>	Bacteria: Firmicutes	37	4 929
<i>Chlamydia pneumoniae AR39</i>	Bacteria: Chlamydiales	37	916
<i>Chlamydia trachomatis</i>	Bacteria: Chlamydiales	37	767
<i>Escherichia coli K12</i>	Bacteria: Proteobacteria	37	5 273
<i>Haemophilus influenzae</i>	Bacteria: Proteobacteria	37	1 728
<i>Helicobacter pylori 26695</i>	Bacteria: Proteobacteria	37	1 257
<i>Mycobacterium tuberculosis</i>	Bacteria: Firmicutes	37	2 999
<i>Mycoplasma genitalium</i>	Bacteria: Firmicutes	37	415
<i>Mycoplasma pneumoniae</i>	Bacteria: Firmicutes	37	535
<i>Rickettsia prowazekii</i>	Bacteria: Proteobacteria	37	1 007
<i>Treponema pallidum</i>	Bacteria: Spirochaetales	37	799

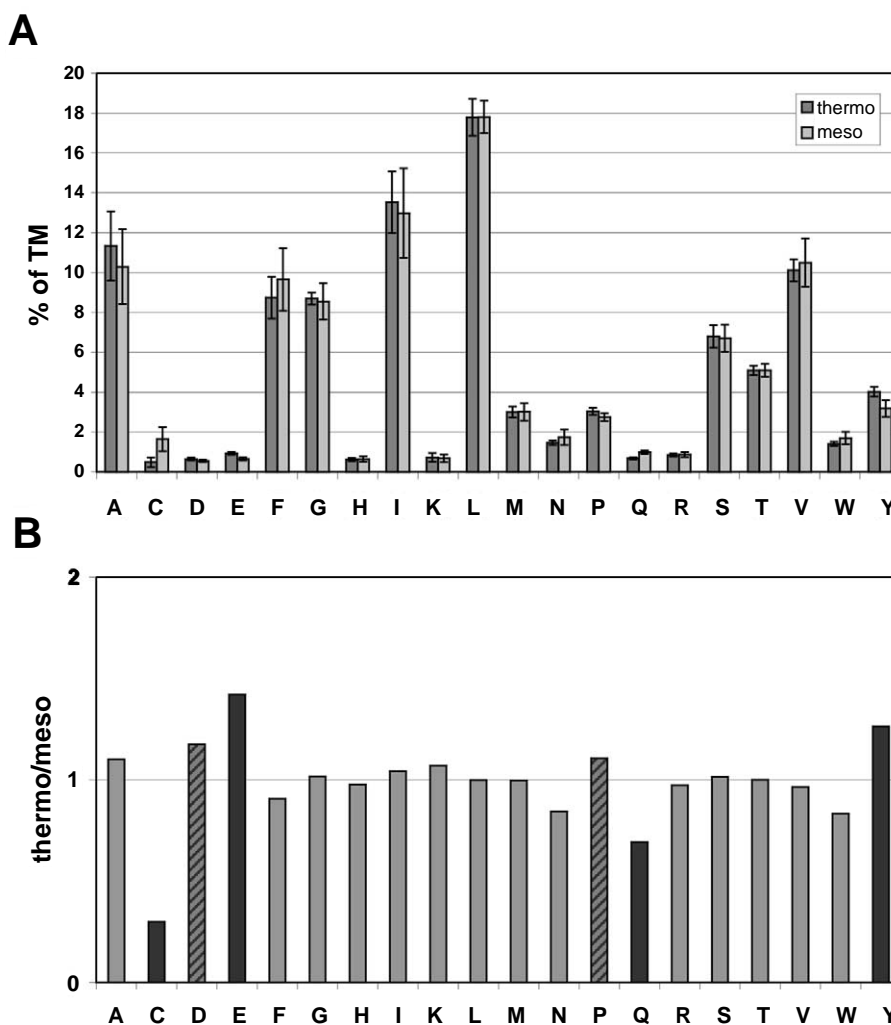


Fig. 1. Relative amino acid composition of TM domains from mesophilic and thermophilic proteomes. A: Amino acid composition (%) of mesophilic and thermophilic TM domains. The variations within the mesophilic and thermophilic proteomes are given as error bars. B: Variation of the amino acid composition in thermophilic relative to mesophilic TM domains. Columns of differences that are statistically significant (t -statistics) at the 5% level ($P < 0.05$) are dashed. Columns of differences statistically significant at a 1% level ($P < 0.01$) are shown in black.

vs. eubacteria. A further differentiation of the thermophilic group into moderately thermophilic (*Methanobacterium thermoautotrophicum* and *Thermoplasma acidophilum*) and hyperthermophilic organisms also did not show significant differences. Although there is a certain variation in the Cys content in mesophiles (0.87–2.98%) as well as in thermophiles (0.14–1.13%), the limits are strongly shifted to much lower percentages in the proteomes from thermophilic organisms. This general decrease may be explained by the higher reactivity and sensitivity of Cys at high temperatures [2]. While in soluble proteins disulfide bonds can contribute to the stability of an extracytoplasmic protein by decreasing the entropy of the protein's unfolded state [24], the occurrence of a disulfide bridge in the membrane environment has not yet been observed and pairs of Cys residues close enough in space to interact with each other can only rarely be found in the transmembrane domains of membrane proteins [18]. In the membrane, Cys can form hydrogen bonds stabilizing a TM (by forming an intrahelical hydrogen bond) or TM–TM interaction [17,18,25]. Previous statistical analysis of the amino acid distribution within 22 protein families showed that a GC bias in the DNA sequence results in an enrichment of some amino

acids while others are under-represented [26,27]. Though Cys is one of the amino acids found to be under-represented in the proteins from extreme thermophiles (Ile, Met, Phe, Ser, Thr, Cys, Trp) [27], only the content of Trp is also decreased in TMs from thermophilic organisms, which suggests that this factor does not generally affect the amino acid composition of TM domains. Since the thermolability of Cys alone cannot explain the different degree of Cys reduction in soluble and TM domains, its decreased occurrence might be explained by some specific function of this amino acid in the membrane environment. Besides contributing to the stability of a TM by forming hydrogen bonds, Cys also fulfills specific functions, like cofactor binding. A recent statistical analysis of Cys clustering in thermophilic versus mesophilic proteomes revealed that, even though the Cys residues involved in specific functions are conserved, the overall content of Cys residues was found to be decreased in the proteome of thermophiles [28]. The high number of replaceable Cys residues in membrane proteins might reflect a minimal usage of this residue in specific functions in membrane proteins compared to soluble proteins.

While in soluble proteins the amino acid Pro is known not

Table 2
The most significant over-represented pair-motifs of mesophilic and thermophilic TM helices sorted by significance down to 10^{-6}

Pair	Occurrence	Expectation	Standard deviation	Significance	Odds ratio
Mesophiles					
IG1	11956	2569	39.17	6.22×10^{-42}	1.27
II4	24962	3083	42.35	7.59×10^{-41}	1.23
GG4	17888	1647	31.28	4.05×10^{-32}	1.30
IP1	4662	897	23.21	6.56×10^{-22}	1.35
IV4	13885	2416	40.13	4.35×10^{-21}	1.19
IA1	13701	2864	43.00	2.92×10^{-18}	1.15
IL4	18978	3825	50.47	1.11×10^{-17}	1.12
VL4	18272	3643	49.22	3.59×10^{-17}	1.13
AV1	14312	2944	43.73	1.13×10^{-13}	1.13
IG2	11956	2190	38.34	1.85×10^{-13}	1.15
GL3	14949	3040	44.99	2.16×10^{-13}	1.12
VV4	22849	2275	38.65	5.85×10^{-11}	1.13
VI4	13885	2293	40.13	7.97×10^{-11}	1.13
IC5	2653	374	15.48	3.08×10^{-09}	1.34
PL1	5402	1208	28.00	3.68×10^{-09}	1.16
VG2	11714	2009	37.39	4.19×10^{-09}	1.12
GI1	11956	2246	39.17	1.53×10^{-08}	1.11
AA3	23521	2673	41.48	4.09×10^{-08}	1.09
VV3	22849	2384	39.73	4.99×10^{-08}	1.10
IC4	2653	387	15.96	1.34×10^{-07}	1.29
LC1	3128	690	21.02	1.42×10^{-07}	0.01
VY10	5008	405	16.93	1.63×10^{-07}	1.30
LL4	37506	5487	58.77	5.94×10^{-07}	0.11
WL10	4036	441	17.67	7.08×10^{-07}	1.26
IY10	5343	468	18.40	7.40×10^{-07}	1.25
IG3	11956	1966	37.46	1.54×10^{-06}	1.10
II7	24962	2149	37.81	1.69×10^{-06}	1.09
LV5	18272	3220	48.00	2.49×10^{-06}	1.08
LG2	14949	3089	45.94	2.75×10^{-06}	1.07
IS2	9615	1611	33.76	3.06×10^{-06}	1.11
LD10	1099	138	9.16	3.30×10^{-06}	1.47
VC2	2642	399	16.42	3.66×10^{-06}	0.12
LC4	3128	565	19.58	4.30×10^{-06}	1.19
WA7	3435	383	16.42	5.16×10^{-06}	1.25
AG4	11936	1832	36.48	5.16×10^{-06}	1.10
LC5	3128	528	19.03	5.17×10^{-06}	1.20
AC3	2612	381	16.06	5.19×10^{-06}	1.24
GA4	11936	1831	36.48	6.46×10^{-06}	1.10
Thermophiles					
GG4	12225	1264	25.47	1.71×10^{-55}	1.50
II4	20043	2725	38.98	1.08×10^{-47}	1.27
IG2	8832	1845	33.45	1.11×10^{-28}	1.26
IP1	3847	827	21.59	1.24×10^{-26}	1.41
VI4	10380	1833	35.01	6.28×10^{-15}	1.18
AV1	9590	1959	35.43	1.70×10^{-14}	1.16
GG7	12225	825	22.77	3.18×10^{-12}	1.25
SL1	9222	2103	36.69	9.76×10^{-11}	1.13
FF4	13532	1199	27.89	2.77×10^{-10}	1.18
GI1	8832	1771	34.16	3.20×10^{-10}	1.14
AA4	16876	1829	34.06	3.24×10^{-10}	1.13
LL4	27125	4084	50.02	5.84×10^{-09}	1.08
GI2	8832	1660	33.45	5.90×10^{-09}	1.14
SI2	8055	1476	31.60	1.31×10^{-08}	1.14
PI2	3847	674	21.10	1.88×10^{-08}	1.22
LL3	27125	4351	51.31	2.08×10^{-08}	1.07
LA1	13083	3088	44.39	2.14×10^{-08}	1.09
VG2	7680	1296	29.86	2.91×10^{-08}	1.15
IS2	8055	1469	31.60	4.70×10^{-08}	1.14
IA2	10826	2131	37.96	6.60×10^{-08}	1.11
AF1	8151	1507	32.08	1.15×10^{-07}	1.13
GF1	6774	1157	28.22	1.87×10^{-07}	1.15
VL4	12291	2322	40.09	2.75×10^{-07}	1.10
IY7	5104	623	21.03	3.12×10^{-07}	1.22
AI2	10826	2119	37.96	3.67×10^{-07}	1.10
VS1	6839	1140	28.26	4.55×10^{-07}	1.13
IY10	5104	468	18.34	5.05×10^{-07}	1.26
VG1	7680	1340	30.52	7.65×10^{-07}	1.12
SG4	5577	725	22.84	8.28×10^{-07}	1.18
LI1	14983	3563	47.92	1.41×10^{-06}	1.06

Table 2 (Continued).

Pair	Occurrence	Expectation	Standard deviation	Significance	Odds ratio
IV4	10380	1725	35.01	1.48×10^{-06}	1.11
GV1	7680	1338	30.52	1.94×10^{-06}	1.12
LG1	10242	2279	38.77	2.51×10^{-06}	1.09
IG1	8832	1704	34.15	6.83×10^{-06}	1.10
PA4	3439	473	18.17	7.81×10^{-06}	1.21
GA4	8055	1233	29.73	9.11×10^{-06}	1.12

to be tolerated in β -sheets as well as in α -helices, this amino acid can frequently be found in TM α -helices. The relatively high occurrence of Pro in TMs can be explained by distinct roles of Pro in membranes as discussed in [29]. In addition, recent data suggest that Pro residues in TMs are destabilizers of misfolded states of TM domains helping to prevent misfolding of membrane proteins [30]. In this context, the significantly higher content of Pro in TMs of thermophilic organisms (Fig. 1) can be explained by the higher possibility of hydrophobic proteins to aggregate and misfold under high temperature conditions. The higher occurrence of Pro in TMs could facilitate to prevent TM protein misfolding in thermophiles.

To find possible pair-motifs that mediate TM–TM interactions in membrane domains from thermophiles vs. mesophiles, we analyzed frequently occurring pairs of residues in the two datasets of TM domains (Table 2). In the group of over-represented pair-motifs for mesophilic TMs, seven motifs can be found that contain a Cys residue, while in thermophilic TMs none can be found containing this residue. The occurrence of Cys residues in over-represented pair motifs suggests an architectural function of these residues in the TMs of mesophilic organisms, and the absence of Cys in the TM domains of thermophilic organisms suggests that the architectural function of the Cys residue containing motifs can be fulfilled by other motifs at high temperatures. This assumption again shows the strong disfavor of Cys in thermophilic proteins.

We also found that Ala, Gly and Ser residues have higher occurrences in pair motifs in thermophilic TM domains. At a 10^{-6} significance level 24 of 36 pairs (67%) contain one of these three in thermophiles, whereas only 16 of 38 pair motifs (42%) do so in mesophiles. The role of Gly residues for the interaction of membrane proteins was recently described [11], and it was also shown that pair-motifs containing Gly residues stabilize proteins under higher temperature conditions [31]. It is likely that the presence of small residues (Ala, Gly, Ser) favors a tighter packing of TM helices which can allow the formation of more hydrogen bonds, especially with the backbone of a TM [32]. Ser residues can also drive the interaction of TMs by forming hydrogen bonds [17,33], possibly explaining its higher occurrence in thermophilic helix motifs. Although seven out of the eight thermophilic proteomes are from archaea, while all mesophiles are bacteria, the generally higher preference for small residues due to generic effects of genetic drift between archaea and bacteria can be excluded since *t*-statistics on the amino acid composition did not prove any significant differences of these residues in thermophiles vs. mesophiles.

4. Conclusions

By analyzing eight thermophilic and 12 mesophilic pro-

teomes, we found a significant increase in the relative amount of the charged residues Asp and Glu and a decrease of the polar, non-ionizable amino acids Asn and Gln. It is likely that the ionizable residues contribute to a network of hydrogen bonds for stabilizing the TM–TM interactions under high temperature conditions. The preference for the small residues Ala, Gly and Ser in pair-motifs suggests a more tightly packed TM interface in membrane proteins of thermophilic bacteria.

As observed already for soluble proteins, the content of Cys residues is reduced in helical membrane proteins of thermophilic bacteria. This most likely reflects a non-specific function of most Cys residues in helical membrane proteins, allowing a replacement of this residue by less reactive residues.

Acknowledgements: We thank the members of the research groups for helpful comments and critically reading of the manuscript and especially A. Senes for his help and advice. This work was supported by grants from the NIH, the NSF and the Bundesministerium für Bildung und Forschung (Leopoldina fellowship to D.S.).

References

- [1] Vieille, C., Burdette, D.S. and Zeikus, J.G. (1996) *Biotechnol. Annu. Rev.* 2, 1–83.
- [2] Vieille, C. and Zeikus, G.J. (2001) *Microbiol. Mol. Biol. Rev.* 65, 1–43.
- [3] Kahn, T.W., Sturtevant, J.M. and Engelman, D.M. (1992) *Biochemistry* 31, 8829–8839.
- [4] Tolner, B., Poolman, B. and Konings, W.N. (1997) *Comp. Biochem. Physiol. A: Physiol.* 118, 423–428.
- [5] Lague, P., Zuckermann, M.J. and Roux, B. (2001) *Biophys. J.* 81, 276–284.
- [6] Mall, S., Broadbridge, R., Sharma, R.P., East, J.M. and Lee, A.G. (2001) *Biochemistry* 40, 12379–12386.
- [7] Apweiler, R. et al. (2001) *Nucleic Acids Res.* 29, 44–48.
- [8] Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) *J. Mol. Biol.* 305, 567–580.
- [9] Gerstein, M. (1998) *Proteins* 33, 518–534.
- [10] Engelman, D.M., Steitz, T.A. and Goldman, A. (1986) *Annu. Rev. Biophys. Chem.* 15, 321–353.
- [11] Senes, A., Gerstein, M. and Engelman, D.M. (2000) *J. Mol. Biol.* 296, 921–936.
- [12] Wallin, E. and Von Heijne, G. (1998) *Protein Sci.* 7, 1029–1038.
- [13] Rajdeep, D. and Gerstein, M. (2000) *Funct. Integr. Genomics* 1, 76–88.
- [14] Chakravarty, S. and Varadarajan, R. (2000) *FEBS Lett.* 470, 65–69.
- [15] Zhou, F.X., Cocco, M.J., Russ, W.P., Brunger, A.T. and Engelman, D.M. (2000) *Nat. Struct. Biol.* 7, 154–160.
- [16] Zhou, F.X., Merianos, H.J., Brunger, A.T. and Engelman, D.M. (2001) *Proc. Natl. Acad. Sci. USA* 98, 2250–2255.
- [17] Adamian, L. and Liang, J. (2002) *Proteins* 47, 209–218.
- [18] Adamian, L. and Liang, J. (2001) *J. Mol. Biol.* 311, 891–907.
- [19] Wimley, W.C. and White, S.H. (1996) *Nat. Struct. Biol.* 3, 842–848.
- [20] Arkin, I.T. and Brunger, A.T. (1998) *Biochim. Biophys. Acta* 1429, 113–128.
- [21] Reithmeier, R.A. (1995) *Curr. Opin. Struct. Biol.* 5, 491–500.
- [22] Rinia, H.A. et al. (2002) *Biochemistry* 41, 2814–2824.

- [23] Yau, W.M., Wimley, W.C., Gawrisch, K. and White, S.H. (1998) *Biochemistry* 37, 14713–14718.
- [24] Matsumura, M., Signor, G. and Matthews, B.W. (1989) *Nature* 342, 291–293.
- [25] Popot, J.L. and Engelman, D.M. (2000) *Annu. Rev. Biochem.* 69, 881–922.
- [26] Di Giulio, M. (2000) *Gene* 261, 189–195.
- [27] Wilquet, V. and Van de Casteele, M. (1999) *Res. Microbiol.* 150, 21–32.
- [28] Rosato, V., Pucello, N. and Giuliano, G. (2002) *Trends Genet.* 18, 278–281.
- [29] Deber, C.M. and Therien, A.G. (2002) *Nat. Struct. Biol.* 9, 318–319.
- [30] Wigley, W.C. et al. (2002) *Nat. Struct. Biol.* 9, 381–388.
- [31] Kleiger, G., Grothe, R., Mallick, P. and Eisenberg, D. (2002) *Biochemistry* 41, 5990–5997.
- [32] Senes, A., Ubarretxena-Belandia, I. and Engelman, D.M. (2001) *Proc. Natl. Acad. Sci. USA* 98, 9056–9061.
- [33] Dawson, J.P., Weinger, J.S. and Engelman, D.M. (2002) *J. Mol. Biol.* 316, 799–805.