Snorkeling of lysine side chains in transmembrane helices: how easy can it get?

Erik Strandberg[*], J. Antoinette Killian[**]

Department of Biochemistry of Membranes, Center for Biomembranes and Lipid Enzymology, Institute of Biomembranes, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Abstract Transmembrane segments of proteins are often flanked by lysine residues. The side chains of these residues may snorkel, i.e. they may bury themselves with their aliphatic part in the hydrophobic region of the lipid bilayer, while positioning the charged amino group in the more polar interface. Here we estimate the free energy cost of snorkeling from thermodynamical calculations based on studies with synthetic transmembrane peptides [Strandberg et al. (2002) Biochemistry 41, 7190–7198]. The value is estimated to be between 0.07 and 0.7 kcal mol^{-1} for a lysine side chain. This very low value indicates that snorkeling may be a common process, which should be taken into consideration both in experimental and in theoretical studies on protein–lipid interactions.
© 2003 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Membrane protein; Lysine side chain; Snorkeling; Lipid phase; Thermodynamical calculation

1. Introduction

Lysines are frequently found near the polar/apolar membrane interface in membrane proteins [1–3]. It has been proposed that the long and flexible side chain of lysine can stretch out of the membrane interior to place the charged amino group in the more polar interface region while keeping the hydrocarbon part of the side chain inside the hydrophobic part of the membrane [4]. This phenomenon is called 'snorkeling' and could be important for the precise location in the lipid bilayer of transmembrane parts of integral membrane proteins and for interactions between proteins and lipids in general.

Fig. 1 shows the general idea about snorkeling in the two main classes of membrane proteins; peripheral proteins having some part bound to a membrane surface, and integral membrane proteins having one or more membrane-spanning segments. These classes of membrane proteins may be modeled by amphipathic peptides and transmembrane helical peptides, respectively, and the snorkeling behavior of lysines in such peptides has been studied [5–7]. For peripheral proteins or amphipathic peptides, snorkeling is proposed to increase

the hydrophobic part of the protein allowing a deeper position in the membrane and thus a stronger binding (Fig. 1A). For transmembrane peptides or protein segments, snorkeling will increase the effective hydrophobic length. For relatively short segments, this can be a way of relieving the strain due to hydrophobic mismatch (Fig. 1B).

The energetic reason for snorkeling would be to place hydrophobic residues in as hydrophobic an environment as possible and still keep the charged groups in a polar environment. It is highly unlikely that the charged groups will partition into the hydrophobic membrane interior, whether or not snorkeling does occur, since the Born energy for this would be too high. However, when snorkeling, the stretching of the long lysine side chains should decrease their conformational entropy by a considerable amount. Therefore, there must be a free energy cost of snorkeling. As far as the authors know, no attempt has so far been made to estimate this free energy cost. Such an energy value could be valuable for modeling of membrane protein structure, giving information about the likelihood of snorkeling and thereby about protein conformation and positioning in the interface region.

Here, a first estimate of the free energy cost of snorkeling of lysine is obtained from thermodynamic calculations based on recent experimental results on transmembrane model peptides [7]. The calculations are based on the observation that the phase behavior of a peptide–lipid system is sensitive to the effective hydrophobic length of transmembrane peptides, in a number of different lipid systems [7–11]. Under matching conditions the peptides can be accommodated in a lamellar (L_α) phase. When the peptide hydrophobic length is less than the lipid hydrophobic thickness, so called negative mismatch, an isotropic phase is formed. For even shorter peptides, an inverted hexagonal (H_{II}) phase is formed. The phases formed depend in a quantitative manner on the extent of mismatch as was demonstrated extensively for peptides with tryptophan as flanking residues, such as WALP peptides (Table 1).

For peptides with lysine as flanking residues the effective length of model peptides depends on whether the lysine side chains are snorkeling. Studies on the effects of such peptides (e.g. KALP peptides, see Table 1) on phase behavior have therefore been useful to provide insight into the snorkeling behavior of lysine side chains [6,7]. From this latter study two key conclusions could be drawn, that each provides valuable information for estimation of the energy cost of snorkeling.

The first conclusion was that upon introduction of negative mismatch by incorporation of too short KALP peptides in the bilayer forming lipid phosphatidylcholine, the mismatch is

_____

*Corresponding author. Present address: Institute of Organic Chemistry, University of Karlsruhe, Fritz-Haber-Weg 6, 761 31 Karlsruhe, Germany. Fax: (49)-721-608 48 23.
**Corresponding author. Fax: (31)-30-253 39 69.
E-mail addresses: erik.strandberg@ioc.uka.de (E. Strandberg), j.a.killian@chem.uu.nl (J.A. Killian).

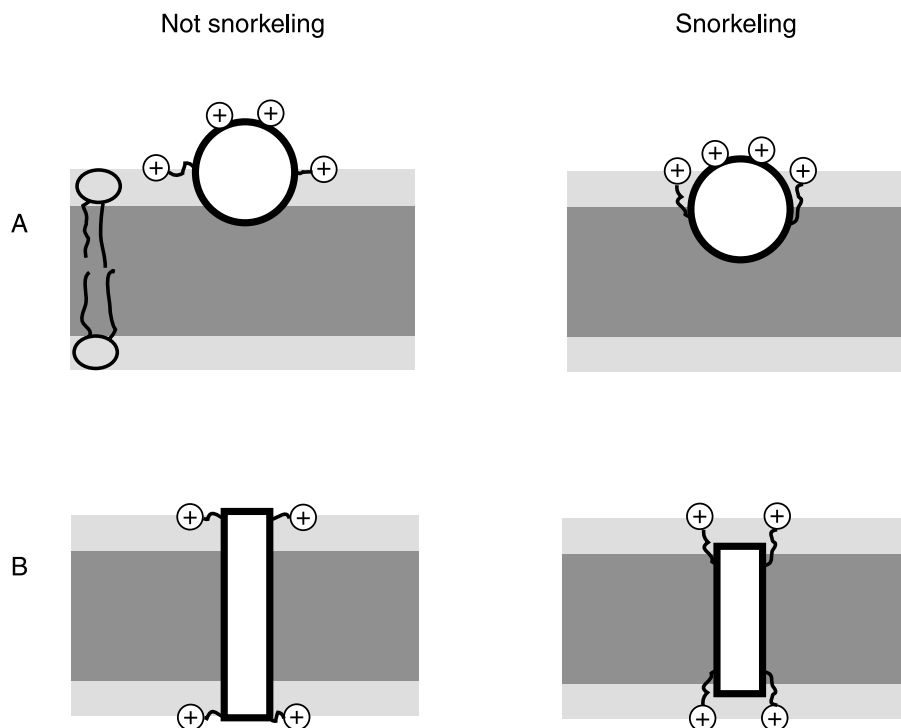Not snorkeling   Snorkeling

Fig. 1. Long charged side chains (lysine and arginine) can snorkel, thereby stretching towards the water phase. In the figure, the dark gray regions are the hydrophobic interior of the membrane, and the light gray regions are the more polar lipid head group regions. Peptides or membrane bound segments of proteins are outlined in black with charges attached on side chains. A: Peripheral membrane segments or amphipathic peptides, lying on the surface of the membrane, can place hydrophobic residues deeper in the lipid bilayer if the side chains are allowed to stretch. B: Transmembrane segments will get a longer effective hydrophobic length upon snorkeling, thereby relieving mismatch for a too short transmembrane segment in a lipid bilayer.

relieved by an increase of the effective peptide length due to snorkeling of the lysine side chains. In contrast, in phosphatidylethanolamine (PE), at temperatures below the $L_\alpha \rightarrow H_{II}$ transition temperature (T_{LH}) of the pure lipid, the mismatch is relieved by the formation of an H_{II} phase, and the lysines do not snorkel. As a lower limit of the energy cost of snorkeling it can thus be assumed that this is higher than the energy cost of H_{II} phase formation in PE, at temperatures below T_{LH}.

The second conclusion was that snorkeling is more favorable than phase separation in lipid systems containing two different peptides. This was based on the observation that when one of the peptides could snorkel, only one lipid phase was formed, but when both peptides were unable to snorkel, two phases were formed. Thus the energy cost of phase separation provides an upper limit of the energy cost of snorkeling. Using this as a starting point a quantitative estimate of the free energy cost of snorkeling of lysine side chains in transmembrane peptides or protein segments will be derived below.

2. Calculations

The calculations use data from the phase behavior of peptide–lipid samples with the well-studied model peptides KALP23 and WALP23 (for peptide sequences see Table 1). Processes will be split into (hypothetical) steps and the total free energy, ΔG, of the process will be the sum of estimated ΔG for different steps.

2.1. Lower limit of ΔG(snorkel) from KALP peptides
As discussed above, the lower limit of the energy cost of

snorkeling can be estimated from the energy cost of peptide-induced H_{II} phase formation in dielaidoylphosphatidylethanolamine (DEPE). This lipid changes from a lamellar (L_α) phase to a hexagonal (H_{II}) liquid crystalline phase at $T_{LH} = 65°C$ [12]. From elementary thermodynamics we have

$$\Delta G = \Delta H - T\Delta S \qquad (1)$$

where H is enthalpy, T is temperature and S is entropy, and since there is a reversible phase transition at the temperature T_{LH} we also know that

$$\Delta G(L_\alpha \rightarrow H_{II}, \ T_{LH}) = 0. \qquad (2)$$

For this phase transition, ΔH was shown from differential scanning calorimetry measurements to be approximately 370 cal mol^{-1} [12]. From Eqs. 1 and 2 follows

$$\Delta S = \Delta H / T_{LH} = 370 \text{ cal mol}^{-1}/338 \text{ K}$$

$$= 1.095 \text{ cal mol}^{-1} \text{ K}^{-1}. \qquad (3)$$

If we assume that ΔH and ΔS are independent of temperature, Eq. 1 will give an expression for $\Delta G(L \rightarrow H_{II})$ at other

Table 1
Amino acid sequences of the peptides used

Peptide	Sequence
WALP23	acetyl-GWW(LA)$_8$LWWA-ethanolamine
WALPn[a]	acetyl-GWW(LA)$_x$LWWA-ethanolamine[b]
KALP23	acetyl-GKK(LA)$_8$LKKA-amide
KALPn[a]	acetyl-GKK(LA)$_x$LKKA-amide

[a] n is the total number of residues in the peptide, $n = 2x+7$.
[b] The end group is either ethanolamine or amide.

temperatures. For example $\Delta G(L_\alpha \to H_{II}, 40°C) = 27.4$ cal mol^{-1}. It was found that when relatively short KALP23 peptides are added to a DEPE system at a peptide to lipid molar ratio of 1:10 at 40°C, the mismatch is relieved by formation of a hexagonal phase rather than by snorkeling [7]. Therefore, we can estimate

ΔG(snorkel, 1 mol KALP23)>

$$\Delta G(L_\alpha \to H_{II}, 40°C, 10 \text{ mol lipids}) \qquad (4)$$

or

$$\Delta G(\text{snorkel}) > 274 \text{ cal (mol peptide)}^{-1} \qquad (5)$$

which is an estimate of the lower limit of the free energy cost of snorkeling. It can be noted that at this peptide concentration there is no defined T_{LH} since the system changes directly from a gel phase to a hexagonal phase. It is therefore likely that the 'true' lowering of the phase transition temperature is larger than 25 K and the value in Eq. 5 is truly a low estimate.

It seems reasonable to assume that the snorkeling process is independent of lipid composition and therefore that the same lower limit of ΔG(snorkel) will apply in other lipid systems.

2.2. Upper limit of ΔG(snorkel) from KALP peptides

We now turn to the more complicated case of the upper limit of snorkeling. This can be estimated from the energy cost of phase separation in lipid systems that contain two peptides with different hydrophobic length. When the tryptophan flanked peptide WALP23 is incorporated into DOPE/DOPG (7:3 molar ratio) at a high peptide concentration an isotropic phase is induced [7]. The lysine flanked peptide KALP23 induces in this lipid system an H_{II} phase, implying that the lysine side chains are not snorkeling (for a more thorough discussion see [7]). However, it was found that when both peptides are present in this lipid system at equimolar amounts, a single isotropic phase is formed, containing both WALP23 and KALP23. From control experiments using peptides with shortened side chains it was concluded that under these conditions the lysine side chains of KALP do snorkel, thus increasing the effective hydrophobic length of the peptide and allowing it to be accommodated in the isotropic phase. Fig. 2A shows a schematic representation of this process of forming one phase containing both peptides upon mixing of two different phases with one type of peptide in each.

For calculation of the energy costs involved, it is useful to describe this process as a series of steps (Fig. 2B): We start with one isotropic phase containing WALP23 [I(W)] and an H_{II} phase containing KALP23 [H_{II}(K)]. There is now a snorkeling step (step 1), a phase transition for the KALP23-containing part of the sample (step 2) and a mixing of the two isotropic phases containing each peptide into a single phase containing both peptides (step 3). Even if the real process does not follow these different steps the starting and final states are determining the difference in free energy, so we can get the total ΔG by summing ΔG for the different (hypothetical) steps:

$$\Delta G(\text{total}) = \Delta G(\text{snorkel}) + \Delta G(H_{II} \to I, K^*)$$

$$+\Delta G(\text{mix}) < 0 \qquad (6)$$

where K* indicates that the phase transition occurs in the presence of snorkeling KALP peptides, and the last term indicates that the total process is spontaneous. This gives

$$\Delta G(\text{snorkel}) < -\Delta G(H_{II} \to I, K^*) - \Delta G(\text{mix}) \qquad (7)$$

We will first estimate ΔG(mix). Since we have the same lipids in the two isotropic phases, only the peptide mixing will give a contribution. We can assume that the enthalpy of mixing is very small, since all peptides are in a lipid surrounding. Thus we can use $\Delta G(\text{mix}) \approx -T\Delta S(\text{mix})$ and assume an ideal mixture. For n mol of molecules, the entropy of mixing between two substances is given by

$$\Delta S(\text{mix}) = -nR(x_A \ln x_A + x_B \ln x_B) \qquad (8)$$

where x_A and x_B are the mole fractions of the substances A and B. Assuming one mole of each substance we get $n = 2$, $x_A = x_B = 0.5$, which gives us

$$\Delta S(\text{mix}) = -2R \ln 0.5 = 11.5 \text{ J K}^{-1} = 2.75 \text{ cal K}^{-1} \qquad (9)$$

This would then be the value of ΔS(mix) per mol of KALP23 peptides. At 30°C this gives

$$\Delta G(\text{mix}) = -T\Delta S(\text{mix}) = -836 \text{ cal (mol KALP23)}^{-1} \quad (10)$$

Using this value in Eq. 7 gives

$$\Delta G(\text{snorkel}) < -\Delta G(H_{II} \to I, K^*)$$

$$+836 \text{ cal (mol KALP)}^{-1} \qquad (11)$$

A
I(W) | H$_{II}$(K) → I(W, K*)

ΔG(total)

B
I(W) | H$_{II}$(K) →(1) I(W) | H$_{II}$(K*) →(2) I(W) | I(K*) →(3) I(W, K*)
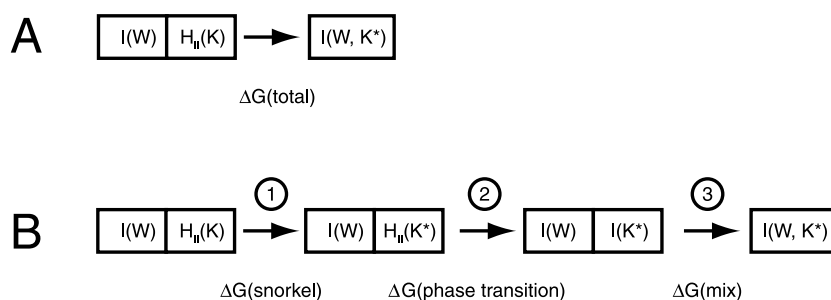
ΔG(snorkel)　　ΔG(phase transition)　　ΔG(mix)

Fig. 2. In the DOPE/DOPG lipid system WALP23 peptides induce an isotropic phase, I(W), while KALP23 peptides induce a hexagonal phase without snorkeling, H$_{II}$(K). A: When the two peptides are mixed in equimolar amounts one isotropic phase is formed, and the KALP23 peptides are snorkeling. This phase is denoted I(W, K*) where K* denotes the presence of snorkeling KALP peptides. B: The process of forming one isotropic phase containing a mixture of WALP23 and KALP23 peptides can be separated into three steps: (1) snorkeling of lysine side chains (H$_{II}$(K)→H$_{II}$(K*)), (2) phase transition from hexagonal to isotropic phase in the presence of snorkeling KALP23 peptides (H$_{II}$(K*)→I(K*)), and (3) mixing of the two isotropic phases to form one phase where WALP23 and KALP23 peptides are mixed (I(W)+I(K*)→I(W, K*)). The total ΔG for this process can be calculated as the sum of ΔG for the three separate steps (Eq. 6).

The next step is then to estimate the free energy for the phase transition from hexagonal to isotropic phase in a system where KALP23 peptides are snorkeling. There are no calorimetric data available for this phase transition and it cannot readily be measured. We can assume that the isotropic phase is spontaneously formed in the presence of snorkeling KALP23 peptides, and hence that the change in free energy is less than zero. It is also probably a rather small value, of the same order as, or even smaller than the free energy of the lamellar to hexagonal phase transition, discussed above, of −274 cal (mol peptide)$^{-1}$. However, to get a 'safe' upper limit of ΔG(snorkel), we will assume that this phase transition from hexagonal to isotropic phase is more favorable than the lamellar to hexagonal phase transition, but not more than twice as favorable. This means $\Delta G(H_{II} \to I, K^*) = -548$ cal (mol peptide)$^{-1}$. Then from Eq. 6 the upper limit would be

$$\Delta G(\text{snorkel}) < 1384 \text{ cal (mol KALP)}^{-1}. \tag{12}$$

2.3. Estimate of ΔG(snorkel) for a lysine side chain

From the phase transition considerations we have now an estimated value of ΔG(snorkel) between 274 and 1384 cal (mol KALP)$^{-1}$. We must now take into account that there are four lysines in a KALP peptide. If all four are snorkeling, we get ΔG(snorkel, lysine) in the interval 68.5–346 cal mol^{-1}. It could equally well be that only two of the lysines are snorkeling, for example only at one side of the peptide, giving ΔG(snorkel, lysine) in the interval 137–692 cal mol^{-1}. However, it seems highly unlikely that only one snorkeling lysine would be enough to relieve a mismatch situation. Therefore the final estimate of ΔG(snorkel, lysine) is between approximately 0.07 and 0.7 kcal mol^{-1}.

2.4. ΔG(snorkel) from entropy calculations

Another way of estimating the snorkeling cost is to look at the molecular details and try to get a value of the entropy loss due to the restricted motion of the lysine side chain. As a crude approximation, we can assume that without snorkeling, the lysine side chain can be anywhere in a half-sphere centered at the C$^\alpha$ of the side chain, and with a radius equal to a fully stretched side chain. From bond lengths and angles, this radius can be estimated to be 7.8 Å giving an available half-sphere volume of 994 Å^3. If we further assume that when snorkeling the side chain is restricted to only one stretched conformation, with the charge pointing up against the water phase, then the available volume is only the volume of the side chain itself. The simplest way of estimating the entropy loss is then by

$$\Delta S = nR \ln(V_{\text{vdW}}/V_{\text{hs}}) \tag{13}$$

where V_{hs} is the volume of the half-sphere and V_{vdW} is the van der Waals volume of a lysine side chain, 122 Å^3 [13]. This gives a ΔG value from entropy loss of about 1.2 kcal mol^{-1}. This value decreases if there is still some conformational flexibility for the side chain while snorkeling, or if the starting volume is less than a half-sphere. It should also be noted that the ΔH term of the free energy is not included here. Although this is a very crude approximation, it does give an idea about the order of magnitude of the entropy loss, and it is comparable with our limits of ΔG(snorkel) above.

3. Discussion

The value for ΔG(snorkel) given by this study is derived from studies of transmembrane peptides (Fig. 1B). However, the snorkeling mechanism is the same for amphipathic peptides, and the same entropy loss is predicted there, so it is reasonable that this value is valid also for snorkeling in such peptides (Fig. 1A) and membrane proteins in general.

The estimate for ΔG(snorkel) gives a small value, 0.7 kcal mol^{-1} as an upper limit. This means that snorkeling is probably a common process in membrane proteins as a way of placing charges of arginine and lysine side chains in a preferable environment. This has several implications, in particular for calculations and modeling of membrane proteins, but also for experimental studies on processes involving for example mismatch adaptation or dynamics of transmembrane protein segments at the membrane/water interface.

The free energy of snorkeling is not zero and it cannot be assumed that snorkeling always takes place. The optimal position of a peptide segment in a lipid bilayer depends on how many residues can be placed in a favorable environment and what the energy terms are. For example, the energy cost of transferring a whole lysine residue of a peptide from inside the membrane to the interface has been determined to be about −1.8 kcal mol^{-1} [14], mainly due to the more favorable positioning of the charged group. If only the hydrophobicity of residues is taken into account, then it is unlikely that a charged lysine will be predicted to be in a transmembrane segment of a protein. But if the lysine is snorkeling, it easily can be accommodated in such a segment, and it may even be followed by a hydrophobic residue that also will reside in the membrane, as proposed for M13 [15]. Snorkeling is a way of favorably positioning the charged group without moving the whole residue. Thus, the cost of snorkeling should be taken into account both in experimental and theoretical studies on protein–lipid interactions.

References

[1] Landolt-Marticorena, C., Williams, K.A., Deber, C.M. and Reithmeier, R.A.F. (1993) J. Mol. Biol. 229, 602–608.

[2] Von Heijne, G. (1994) Annu. Rev. Biophys. Biomol. Struct. 23, 167–192.

[3] Arkin, I.T. and Brungen, A.T. (1998) Biochim. Biophys. Acta 1429, 113–128.

[4] Segrest, J.P., De Loof, H., Dohlman, J.G., Brouillette, C.G. and Anantharamaiah, G.M. (1990) Proteins 8, 103–117.

[5] Segrest, J.P., Jones, M.K., De Loof, H., Brouillette, C.G., Venkatachalapathi, Y.V. and Anantharamaiah, G.M. (1992) J. Lipid Res. 33, 141–166.

[6] De Planque, M.R.R., Boots, J.-W.P., Rijkers, D.T.S., Liskamp, R.M.J., Greathouse, D.V. and Killian, J.A. (2002) Biochemistry 41, 8396–8404.

[7] Strandberg, E., Morein, S., Rijkers, D.T.S., Liskamp, R.M.J., van der Wel, P.C.A. and Killian, J.A. (2002) Biochemistry 41, 7190–7198.

[8] Killian, J.A., Salemink, I., de Planque, M.R.R., Lindblom, G., Koeppe II, R.E. and Greathouse, D.V. (1996) Biochemistry 35, 1037–1045.

[9] De Planque, M.R.R., Kruijtzer, J.A.W., Liskamp, R.M.J.,

Marsh, D., Greathouse, D.V., Koeppe II, R.E., de Kruijff, B. and Killian, J.A. (1999) J. Biol. Chem. 274, 20839–20846.

[10] Morein, S., Koeppe II, R.E., Lindblom, G., de Kruijff, B. and Killian, J.A. (2000) Biophys. J. 78, 2475–2485.

[11] Van der Wel, P.C.A., Pott, T., Morein, S., Greathouse, D.V., Koeppe II, R.E. and Killian, J.A. (2000) Biochemistry 39, 3124–3133.

[12] Epand, R.M. (1990) Chem. Phys. Lipids 52, 227–230.

[13] Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) J. Comp. Chem. 4, 187–217.

[14] White, S.H. and Wimley, W.C. (1998) Biochim. Biophys. Acta 1376, 339–352.

[15] Stopar, D., Spruijt, R.B., Wolfs, C.J.A.M. and Hemminga, M.A. (1996) Biochemistry 35, 15467–15473.