The GxxxG-Containing Transmembrane Domain of the CCK4 Oncogene Does Not Encode Preferential Self-Interactions†

Felix J. Kobus and Karen G. Fleming*

T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

Received September 7, 2004; Revised Manuscript Received November 21, 2004

ABSTRACT: The recently cloned colon carcinoma kinase 4 (CCK4) oncogene contains an evolutionarily conserved GxxxG motif in its single transmembrane domain (TMD). It has previously been suggested that this pairwise glycine motif may provide a strong driving force for transmembrane helix−helix interactions. Since CCK4 is thought to represent a new member of the receptor tyrosine kinase family, interactions between the TMDs may be important in receptor self-association and activation of signal transduction pathways. To determine whether this conserved CCK4 TMD can drive protein−protein interactions, we have carried out a thermodynamic study using the TMD expressed as a *Staphylococcal* nuclease (SN) fusion protein. Similar SN−TMD fusion proteins have been used to determine the sequence specificity and thermodynamics of transmembrane helix−helix interactions in a number of membrane proteins, including glycophorin A. Using sedimentation equilibrium in C14 betaine micelles, we discovered that the CCK4 TMD is unable to drive strong protein−protein interactions. At high protein/detergent ratios, the SN−CCK4 fusion protein will dimerize, but a stochastic model for protein association in micelles can explain the observed dimer population. For low-affinity interactions such as the one studied here, an understanding of this discrete stochastic distribution of membrane proteins in micelles is important for distinguishing between preferential and random self-interactions, which can both influence the oligomeric population. The lack of a thermodynamically meaningful self-association propensity for the CCK4 TMDs demonstrates that a GxxxG motif is not sufficient to drive transmembrane helix−helix interactions.

Receptor tyrosine kinases (RTKs)[1] are known to play an important role in the development and/or progression of many forms of cancer. A molecular understanding of the mechanisms by which these receptors promote cellular transformation is critical for the understanding and treatment of human cancer. In normal cells, RTKs are transmembrane signaling proteins that transmit biological signals from the extracellular environment to the interior of the cell. The regulated transmission of these signals is important for many cellular events, such as cell growth and differentiation, axonal growth, epithelial growth and development (*1*). RTKs are misregulated in many human cancers including breast cancer (erbB2/Her2) (*2*, *3*), ovarian cancer (erbB2) (*2*, *3*), melanoma (CCK4) (*4*), and colon cancer (erbB1 and CCK4) (*5*, *6*). The predominant RTK-related genetic alteration in human cancer is not mutation of the receptor protein, but rather it is amplification of the gene. Subsequent overexpression of receptor proteins leads to constitutive stimulation of the RTK activity and uncontrolled cellular signaling.

Recently, colon carcinoma kinase 4 (CCK4), a new member of the receptor tyrosine kinase superfamily was cloned from human colon carcinoma-derived cell lines (*7*). The CCK4 mRNA for this gene was variably expressed in colon carcinoma derived cell lines but not expressed in human adult colon tissues. In contrast, it is expressed in fetal colon of the mouse (*7*). These observations suggest that the normal role for CCK4 may be in development of the colon followed by down regulation in the adult tissue. The unusual expression of CCK4 in colon carcinoma cell lines might indicate that the protein is functioning abnormally in adult colon cells, which suggests a role in colon carcinoma development and/or proliferation.

Orthologues for CCK4 have been cloned from other species (*8*−*10*). Amino acid residues within the transmembrane domain (TMD) are among the conserved features (*9*). The CCK4 TMD is 55% identical to the transmembrane sequences of the chicken *Klg* and Hydra *Lemon* orthologues (*9*), and a pattern of residues with helical periodicity is conserved. Within the conserved residues of the TMD is a GxxxG sequence, a motif thought to promote interactions between transmembrane helices (*11*, *12*). The evolutionary conservation of the CCK4 transmembrane sequence suggests the hypothesis that it may play a functionally important role in signaling that involves transmembrane helix−helix interactions. To test whether this CCK4 TMD may serve as a driving force for protein−protein interactions, we carried out a thermodynamic study of the TMD using sedimentation equilibrium in detergent micelles.

† This work was supported by a grant from the NIH (GM57534) and by a Career Award from the Department of Defense (DAMD17-02-1-0427).

* To whom correspondence should be addressed. Voice: 410-516-7256. Fax: 410-516-4118. E-mail: Karen.Fleming@jhu.edu

[1] Abbreviations: CCK4, colon carcinoma kinase 4; C14 betaine, 3-(*N*,*N*-dimethylmyristyl-ammonio)propanesulfonate; RTK, receptor tyrosine kinases; SN, *Staphylococcal* nuclease; TMD, transmembrane domain; GpA, glycophorin A; SN−CCK4, a fusion protein composed of SN and the transmembrane domain of CCK4; SN−GpA, a fusion protein composed of SN and the transmembrane domain of glycophorin A.

10.1021/bi048076l CCC: $30.25
Published on Web 01/08/2005

Table 1: Transmembrane Sequences Used in This Study[a]

Protein	SN-TM fusion sequence aligned by the GxxxG motif
CCK4s	QLAVHHFSEPG IGLSVGAAVAYIIAVLGLMFYSKKR
CCK4xs	QLAVHHFSEPGMIQTIGLSVGAAVAYIIAVLGLMFYSKKR
	* * ** * *** * *** **
GpA WT	QLAVHHFSEPEITLIIFGVMAGVIGTILLISYGIRRLI

[a] Sequences are aligned by the GxxxG motif. The predicted transmembrane domains are underlined. The N-terminal sequence shows the junction between the SN sequence and the cloned transmembrane domain of interest. The conserved residues in the transmembrane domain are denoted with a star. In a manner analogous to the GpA construct, positive charges were added to the C-termini of the CCK4 sequences. The SN−CCK4xs contains the extended linker region.

EXPERIMENTAL METHODS

Cloning and Sample Preparation. An image clone for the CCK4 protein was purchased from Invitrogen. The sequence encoding the transmembrane region was amplified using the polymerase chain reaction using primers containing unique SmaI and BamH I sites, and the PCR product was cloned C-terminal to *Staphylococcal* nuclease (SN) in an IPTG-inducible pET11a expression construct from which the open reading frame for the glycophorin A transmembrane had been excised. The DNA sequence was verified by double-stranded DNA sequencing. The Cys-to-Ser mutation (immediately before the KKR) was introduced using the QuickChange mutagenesis kit. The SN−CCK4s and SN−CCK4xs proteins were purified using the published protocol for the SN−GpA proteins (*13*). Immediately before sedimentation equilibrium analysis, samples were exchanged by ion-exchange chromatography into buffer containing 10 mM C14 betaine. SigmaUltra grade C14 betaine was purchased from Sigma.

Sedimentation Equilibrium Analytical Ultracentrifugation. Sedimentation equilibrium experiments were performed at 25 °C using a Beckman XL-A analytical ultracentrifuge using six-sector cells equipped with quartz windows in an AnTi 60 rotor as described previously in detail (*13*−*15*). The samples were centrifuged for lengths of time sufficient to achieve equilibrium determined by using the MATCH algorithm. Equilibrium data obtained from absorbance at 280 nm were analyzed by nonlinear least-squares curve fitting of radial concentration profiles with the Macintosh version of NONLIN (*16*) using the equations describing the reversible association in sedimentation equilibrium. For each global fit, nine data sets were used consisting of three different initial protein concentrations (initial A_{280} absorbance values of 0.9, 0.6, and 0.3 cm^{-1}) analyzed at three significantly different speeds (20000, 24500, 30000 revolutions per minute). The monomeric molecular masses and partial specific volumes were calculated using the program SEDNTERP (*17*), and these parameters were held constant in fitting the absorbance versus radius profiles. The equilibrium constant was converted from absorbance to (bulk molar) units using molar extinction coefficients of 20952 and 21425 cm^{-1} M^{-1} for the SN−CCK4s and SN−CCK4xs proteins, respectively. These values were calculated by adjusting the experimental molar extinction coefficient for SN for the additional mass of the TMD of interest as has been previously done for the glycophorin A transmembrane domain (*18*).

The apparent free energy of association was calculated from the apparent monomer−dimer equilibrium constant in the standard way, $\Delta G^{\circ}_{App} = -RT \ln K_{App}$. The standard state mole fraction free energy of association was calculated using the ideal-dilute assumption as $\Delta G^{\circ}_{x} = -RT \ln K_x$ where $K_x = K_{App}[\text{micellarC14}]_w$ and $[\text{micellarC14}]_w$ is the concentration of micellar C14 betaine expressed on the aqueous molar scale (*14*, *15*).

Sequence Motifs. Sequence motifs were identified using the online TMSTAT calculation (*11*) (http://bioinfo.mbb.yale.edu/tmstat/) with the extended version of the CCK4 TMD.

RESULTS

We expressed the CCK4 TMD as a C-terminal fusion protein with *Staphylococcal* nuclease (SN), an experimental approach that has been successfully used to map the helix−helix interactions of glycophorin A (*13*−*15*, *19*−*22*), BNIP-3 (*23*), and phospholamban (*24*) TMDs. We have previously shown that this method using this type of SN fusion construct has resulted in a quantitative ranking of glycophorin A transmembrane mutants by their ability to dimerize (*13*, *21*, *22*), and the relative stabilities of a series of GpA alanine mutants is conserved between the micellar environment in vitro and bacterial membranes in vivo (*13*). The advantage of using the sedimentation equilibrium method is that the molecular mass and thus stoichiometries are determined. In addition, in the absence of protein−protein association, sedimentation equilibrium will still result in a positive, interpretable result since it will directly report the monomeric molecular weights.

The CCK4 GxxxG-Containing Transmembrane Domain Self-Associates Only Weakly. Table 1 shows the CCK4 transmembrane sequences considered in this study. Since the GxxxG motif in CCK4 is located closer to the predicted N-terminus of its TMD, we were concerned that self-association might be sterically hindered by the SN portion of the fusion protein. Therefore, we extended the linker region by adding four additional residues from the CCK4 open reading frame. In addition, the C-terminal end of the wild-type CCK4 TMD contains a cysteine residue, which we mutated to serine. For comparison, the sequence of the glycophorin A transmembrane WT TMDs is also shown. We first used SDS−PAGE to screen the purified proteins for the ability of the CCK4 TMD to drive strong self-association.

This method demonstrated that both of the SN−CCK4 fusion proteins migrate as a monomeric species on SDS−PAGE. In contrast, the SN−GpA fusion protein migrates predominantly as a dimer under those same conditions (data not shown).

To further test for a self-association propensity of the CCK4 TMD, we carried out sedimentation equilibrium analytical ultracentrifugation experiments. Sedimentation equilibrium allows detection of protein−protein interactions over a range of association strengths. In particular, we have detected significant dimer populations of the SN-GpA mutant proteins even when these proteins are monomeric as measured by SDS−PAGE (*13*, *18*, *21*, *22*). In addition, unlike SDS−PAGE, sedimentation equilibrium experiments can provide a direct measure of the equilibrium constant for a self-association reaction. We chose to use C14 betaine micelles since this detergent micelle environment has been used for the analysis of several designed transmembrane helices (*25*) along with the fact that the glycophorin A TMD dimer shows in ideal-dilute behavior in these micelles (*15*).

Sedimentation equilibrium experiments carried out in 10 mM C14 betaine detergent micelles suggest that the SN−CCK4s protein can self-associate, albeit weakly. In four independent experiments, the single ideal species fit returned a molecular weight value that is approximately 20% greater than the calculated value for the monomeric protein, which suggests the formation of higher order species. To determine the molecular weights of these self-associated forms, the data were fit to association models including dimers, trimers, or tetramers, In three of four of the experiments, the monomer−dimer fit is the best description of the data. Figure 1 shows a typical species distribution for this monomer−dimer fit, where it can be seen that approximately 20% of the SN−CCK4s protein is dimeric. In the fourth independent experiment, a monomer−trimer fit was equally as good as a monomer−dimer fit; however, the dimerization constant for the monomer−dimer fit of these data was consistent with the other three data sets and we therefore used this value.

The average apparent free energy of association determined in 10 mM C14 betaine micelles at 25 °C, ΔG°_{App}, equals −5.2 (±0.1) kcal mol^{-1} ($N = 4$). Under the ideal-dilute assumption (*14*), this apparent free energy would correspond to a standard state mole fraction association free energy, ΔG°_{x}, equal to −2.40 (±0.1) kcal mol^{-1}. This value is considerably weaker than that of the SN−GpA TMD in either C14 betaine micelles, $\Delta G^{\circ}_{x} = -5.8$ kcal mol^{-1} (*15*) or C_8E_5 micelles, $\Delta G^{\circ}_{x} = -7.0$ kcal mol^{-1} (*14*). Figure 2 shows a comparison of the dimer distributions for SN−CCK4s and SN−GpA calculated from the standard state free energy values in C14 betaine micelles where it can be seen that this free energy difference leads to significant shifts in the concentration range over which the protein population is predominantly dimeric. At a total protein concentration of 10 μM, the SN−GpA fusion protein is >80% dimeric in 10 mM C14 betaine micelles. In contrast, the SN−CCK4s protein is ∼10% dimeric at this same overall protein concentration in 10 mM C14 betaine micelles.

Because of the observation that the self-interactions were so weak for this first SN−CCK4 construct, we also carried out sedimentation equilibrium experiments using the SN−CCK4xs fusion protein containing an extended linker region between SN and the TMD. The results for four independent

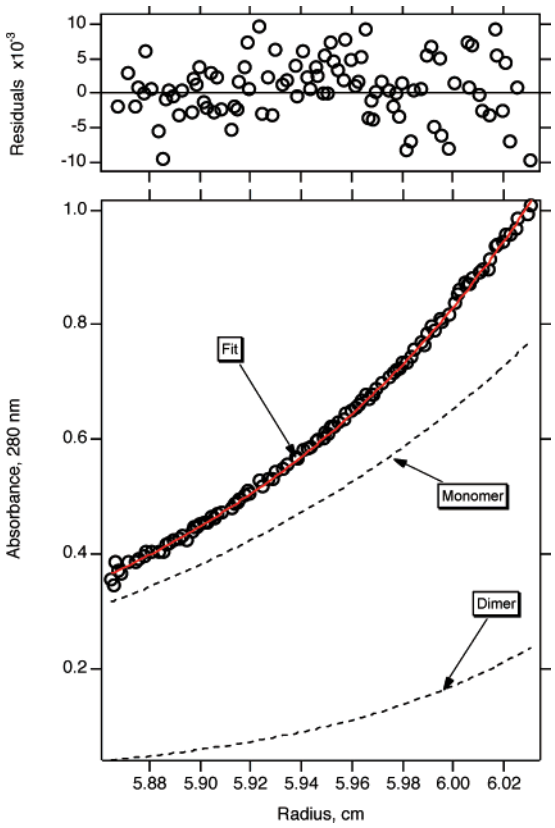

FIGURE 1: SN−CCK4s species distribution in C14 betaine micelles. The lower panel shows a representative data set for 43 μM (bulk) SN−CCK4s collected in 10 mM C14 betaine at 20000 rpm. The open circles are the data, and the line shows the global fit resulting from the global fit of nine data sets for this experiment. The broken lines represent the monomer and dimer contributions as determined from the global analysis, which sum to equal the fit. The upper panel shows the residuals of the fit for this data set, which are small and random.

sedimentation equilibrium experiments on the SN−CCK4xs protein were within error of the original SN−CCK4s fusion protein. The average apparent free energy of association, ΔG°_{App}, was equal to −5.3 (±0.2) kcal mol^{-1} ($N = 4$) in 10 mM C14 betaine micelles. Under the ideal-dilute assumption (*14*), this apparent free energy would correspond to a standard state mole fraction association free energy, ΔG°_{x}, equal to −2.50 (±0.2) kcal mol^{-1}. Thus, extending the linker region of the CCK4 TMD using native sequence resulted in no enhancement of protein−protein interactions.

DISCUSSION

To determine whether the GxxxG-containing TMD of CCK4 encodes the potential for strong helix−helix self-association, we carried out a thermodynamic study using an SN−CCK4 TMD fusion construct. This type of construct has been used in the past to map the sequence dependence of helix−helix oligomerization in several transmembrane proteins, most notably the TMD of glycophorin A (*13*, *19*−*21*). Using both SDS−PAGE and sedimentation equilibrium in detergent micelles, we find that the SN−CCK4 protein shows a very weak propensity for self-association.

The SN−CCK4 Species Distribution Is Well Described by a Random Distribution of Proteins in Micelles. To observe SN−CCK4 interactions, the current experiments were carried out at protein/detergent ratios much higher than were

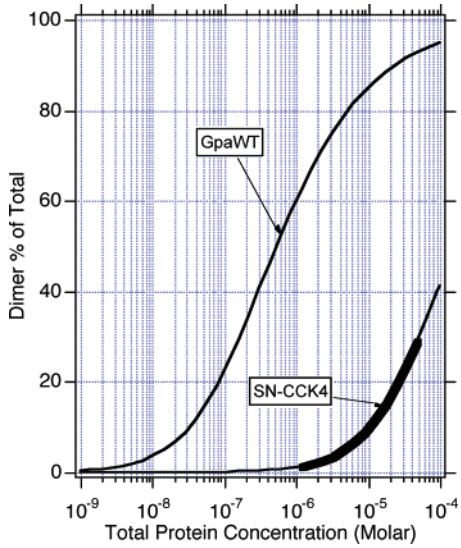

FIGURE 2: CCK4 is destabilized with respect to glycophorin A. A comparison of the dimer species distribution between the SN−GpA and SN−CCK4 transmembrane proteins in 10 mM C14 betaine micelles. The thin curves are calculated based on the free energies of association. The thickened portion of the SN−CCK4 curve represents the region of the data that was observed in the sedimentation equilibrium experiments. The less favorable free energy of association for SN−CCK4 shifts the population of dimer to higher protein concentrations.

previously used for glycophorin A (*15*), and at the highest SN−CCK4 protein concentration the number of protein molecules approached the number of micelles. Since the fundamental measurement in the sedimentation equilibrium under the current experimental conditions is the protein mass, two protein molecules that occupy the same micelle will sediment together and be measured as a dimer whether or not there is a preferential interaction between them. This co-partitioning phenomenon has been referred to in the literature as "artifactual togetherness" (*26*) and should only occur at high protein concentrations coupled with low detergent micelle concentrations. We hypothesize that this co-partitioning is due to an excluded volume effect and that these conditions can be thought of as thermodynamically nonideal. A similar excluded volume effect has been indicated to occur in soluble proteins at high protein concentrations or in the presence of small molecules that induce crowding conditions (*27*). To determine the extent of thermodynamically meaningful preferential interactions between the SN−CCK4 transmembrane proteins, it is first necessary to distinguish them from the protein interactions induced by the limiting solvent conditions.

To develop a model describing the distribution of non-interacting proteins in micelles, we employed a statistical approach that considers distribution of proteins in micelles in terms of the random occupancy of placing balls in bins, where the protein molecules represent balls and the individual micelles represent bins. Using this formalism, the random occupancy of any particular micelle by any particular protein molecule can be described as a Bernoulli event (e.g., either the micelle contains that protein or not), which can be described by a binomial distribution:

$$\Pr[X_i = k] = \binom{m}{k} p^k (1-p)^{m-k} \tag{1}$$

where X_i is the random variable that counts the number of proteins in a particular micelle, i, k is the number of proteins occupying a micelle (e.g., the "oligomeric state"), m is the number of protein molecules, p is the probability of each Bernouilli trial, which equals $1/n$, and n is the number of micelles. When the number of proteins and the number of micelles both exceed the allowed occupancy of a particular micelle, k, this binomial distribution is approximated by the Poisson distribution:

$$\lim\,(m,n \gg k):\Pr[X_i = k] \approx \frac{1}{k!}\left(\frac{m}{n}\right)^k e^{-m/n} \tag{2}$$

To consider the statistical protein distribution, we write the partition function, P, as the sum all those probabilities for which $k > 0$. (e.g., for this purpose, we are not interested in the distribution of "empty" micelles). To convert each probability to a relative concentration, each term is multiplied by the number of proteins that the probability represents.

$$P = \sum_{k>0} k\Pr[X_i = k] \tag{3}$$

The fraction of any particular species can then be calculated as a function of the mole fraction protein (m/n) in the micellar phase from this partition function by dividing its contribution by the sum of all species. For example, the fraction dimer can be calculated as

$$f_{\text{Dimer}} = \frac{2\Pr[X_i = 2]}{\sum_{k>0} k\Pr[X_i = k]} \tag{4}$$

The exact form of the partition function and the distribution of species will depend on the allowed values of k in the model. In this study, we limited the species to monomer and dimer. In addition, we postulate that an unlimited number of proteins in a micelle would not be physically possible. Experimentally, this would behave as an aggregated protein, and we approximated this as a micelle containing 10 monomers. While we did not observe any aggregation in our sedimentation equilibrium experiments, we wished to calculate the statistical distribution over wide range of mole fraction values. Including this parameter in the Poisson model allows us to approximate when this might occur, and by including this term in the partition function, we confirmed that decamer is not a significant species within the concentration range shown in Figure 3 (data not shown). Thus, for clarity, we subsequently removed this term, and the final partition function contains only monomeric and dimeric species.

A final parameter in this statistical model is required to convert micelles to total micellar detergent concentration, where the latter is the experimental variable. For protein-free micelles, we assumed an aggregation number for C14 betaine micelles of 100, which lies intermediate between the published values for C12 betaine and C16 betaine micelles (*28*, *29*). For protein-containing micelles, we recognize that the number of detergents bound may or may not be similar to the protein-free aggregation number, and it may also be that the monomeric and dimeric forms of the protein would be solvated by a micelle of a different size. Under the

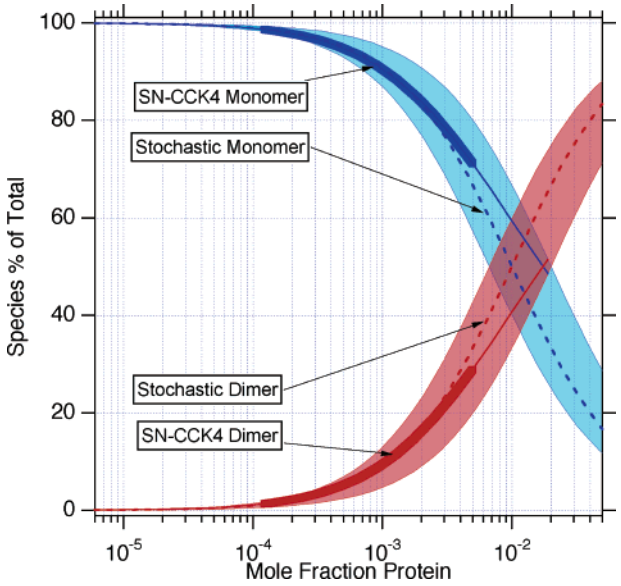

FIGURE 3: The SN−CCK4s species distribution is similar to the stochastic distribution in micelles. The solid blue and red lines are the experimentally determined monomer and dimer distributions, respectively. The thickened portions of these curves represent those concentrations that were experimentally observed, and the thin portions of these curves represent extrapolations of those distributions to the equilibrium midpoint. The broken blue and red lines are the monomer and dimer distributions predicted from the Poisson distribution of proteins in micelles assuming no preferential interaction with an aggregation number of 100. The shaded regions represent the area covered by the distributions allowing the aggregation number to vary from 50 to 150. Mole fraction on the abscissa is defined as the ratio of moles of protein to moles of micellar detergent.

density-matching conditions used for sedimentation equilibrium, there is no experimental information about the size of the micelle required to solvate the transmembrane regions of the protein molecules. For large membrane proteins le Maire and colleagues have postulated that the amount of detergent bound depends on the hydrophobic protein surface area (*30*−*32*). However, for smaller membrane proteins, such as the single transmembrane helix of CCK4, it is thought that the transmembrane region may serve as an organizing hydrophobic nucleus surrounded by detergent molecules that bind in a micellar fashion, and the amount of detergent bound to the glycophorin A transmembrane helix is similar to the aggregation number for protein-free micelles (*33*). We therefore used the protein-free number as a starting approximation of the number of detergent molecules bound, and then we varied this number from 50 to 150 to account for our uncertainty. As long as the number of detergents bound is within this range, our results shown in Figure 3 indicate that the exact value is not an extremely sensitive factor in this analysis, and the fundamental conclusion is the same.

Figure 3 shows the mole fraction distribution of species for SN−CCK4 overlaid upon the mole fraction distribution of species from the random distribution model containing monomer and dimer forms. It can be seen that the decrease in experimentally observed monomeric SN−CCK4s tracks well with the decrease in the statistical distribution of monomer using an aggregation number of 100 as depicted by the blue broken line. Concomitantly, the increase in experimentally observed dimeric SN−CCK4 is well approximated by the expected statistical distribution of dimer. Even if the number of detergents bound varies by ±50, the experimentally observed SN−CCK4s species deviate only slightly from the random distributions, suggesting that the number of detergent molecules solvating each species is a relatively insensitive parameter. The results are very similar when the partition function for the random model includes an additional trimeric form (data not shown).

We can consider the free energy of preferential interactions for SN−CCK4 as a reflection of the deviation of the experimental monomer−dimer midpoint from that of the stochastic distribution. Within the range of distributions shown by the shaded regions of Figure 3, the comparison demonstrates that the free energy of any preferential self-interactions between SN−CCK4 proteins in micelles is minimal and certainly less than one kcal mol^{-1}. By comparison with the calculated random distributions, it can be seen that the midpoints overlay almost exactly if the average aggregation number equals 50. In this case, we can conclude that the dimerization propensity of SN−CCK4s arises solely from the random distribution in micelles, e.g., association of SN−CCK4 is completely due to overloading of the micelles with protein and not from any preferential interactions between the helices. On the other hand, if the average aggregation number is 150, the midpoint of the experimental curve (1.5×10^{-2}) falls to the right of the stochastic curve (5×10^{-3}). This result might suggest the presence of slight repulsive interactions between the SN−CCK4 proteins. Physically, this is possible using this stochastic model, because there is a significant population of protein-free micelles at the midpoint of the random distribution. In this latter case, the mole fraction concentration difference between the midpoints of the random and experimental populations is approximately 3-fold. From this, we can estimate the free energy of repulsion as being on the order of the available thermal energy at 25 °C. Thus, even if with the uncertainty of knowing the size of the protein−micelle complex, we can conclude that the TMDs of CCK4, unlike those of glycophorin A, do not encode a thermodynamically meaningful preferential self-interaction in micelles.

Sequence Motifs Do Not Necessarily Encode Strong Self-Association. The lack of a preferential self-interaction between the helices was initially surprising to us in light of the evolutionary conservation of the transmembrane sequence, which includes a conserved GxxxG motif. Since it has been suggested that this pairwise glycine motif might be a sequence signature indicative of transmembrane α-helix dimerization (*12*), we hypothesized that the CCK4 TMDs would self-associate. The presence of a GxxxG motif is important for the most stable dimer formation of the glycophorin A TMD (*21*), and the GpA dimer has been invoked as a structural model for the way in which the glycines in a GxxxG motif might interact in a transmembrane dimer (*11*, *12*, *34*). Moreover, the leucine following the first glycine in the CCK4 GG4 motif was found to be the only stabilizing single point mutation in the context of the glycophorin A TMD (*21*), and we would therefore not predict that this small-to-large mutation would be inhibitory to protein interactions. However, a rationalization of the cumulative differences between the GpA and CCK4 transmembrane sequences is not straightforward, since it has been shown by double-mutant cycle analysis that the thermo-

dynamic interactions between the sites flanking and including the GpA GxxxG motif are cooperative and not additive (*22*). Glycophorin A TMD mutant sequences lacking a GxxxG motif can still dimerize, whereas other SN-GpA mutant sequences containing a GxxxG motif do not dimerize (*22*). As in glycophorin A, we hypothesize that the role of the GxxxG motif in the CCK4 TMD may be to allow a close approach of the helices in order for flanking sequences to interact. This might be important for disulfide bond formation in the full-length CCK4 protein, which contains a conserved cysteine in the flanking region C-terminal to the TMD. Alternatively, extramembranous regions of CCK4 may encode the protein−protein interactions for this receptor, and the TMD sequence need only permit a close approach of two or more proteins. The small amino acid glycine would facilitate this, and the remainder of the sequence need only be optimized to allow it.

In addition to a conserved GxxxG motif, TMSTAT analysis (*11*) of the CCK4 TMD reveals that it contains 25 pair wise motifs that are significantly overrepresented ($p < 0.05$) and six that are significantly underrepresented. Interestingly, only 31% of the overrepresented pairs involve evolutionarily conserved residues in the TMD, whereas this value is 50% for the underrepresented pairs. From the current data, we can conclude that the presence of a GxxxG and other pairwise motifs as arranged in the CCK4 TMD do not appear to be sufficient to encode strong helix−helix interactions. Additional experimentation is required to elucidate the functional importance of these motifs in transmembrane helices.

Functional Significance of the Absence of Preferential Self-Interactions between the CCK4 Transmembrane Domains. The lack a preferential self-association propensity for the CCK4 TMD may have several functional implications. First, we can postulate that the CCK4 TMD may play a relatively passive role in signaling, as long as the TMD sequence allows other regions of the full-length receptor to appropriately interact. Alternatively, the CCK4 TMD may encode an ability to interact with the TMDs of other receptor tyrosine kinases. The kinase domain of CCK4 has been shown to lack activity, suggesting that that the full-length CCK4 receptor may function like the erbB3 receptor tyrosine kinase, which also lacks kinase activity (*35*). It has been proposed that the evolutionary pressure for this lack of kinase activity might be that their cytoplasmic domains have extremely potent mitogenic power for recruiting downstream signaling molecules (*36*). Overactive signaling would be deleterious to cells if they also were able to transphosphorylate themselves in homodimeric complexes. Therefore, like erbB3, it is likely that signaling active complexes of CCK4 must be co-receptors containing another kinase-active partner. The TMD of CCK4 might be involved in mediating such hetero-interactions.

Alternatively, a second explanation for the lack of self-interactions interactions in the current experiments could be that that the functional properties of the TMD are not compatible with the micellar environment required for the sedimentation equilibrium analysis. In a cell, the biological lipid bilayer may impose additional structural and/or dynamic constraints upon the CCK4 TMD, which might stabilize conformations that would encode specific interactions in bilayers.

CONCLUSIONS

Regulated signaling of tyrosine kinase receptors is crucial for controlled cell growth. The observation that the CCK4 TMD self-associates much more weakly than the glycophorin A TMD is consistent with the current expectation that a tyrosine kinase receptor would not normally be constitutively found in a dimeric state. However, the complete lack of preferential self-interactions in micelles may suggest that the CCK4 TMD plays a supportive role in signaling. Future studies in lipid bilayers might possibly reveal specific interactions between the helices, and such a finding would emphasize the importance of the hydrophobic environment in modulating the structural and/or thermodynamic properties of this membrane protein. Alternatively, the conserved CCK4 TMD may mediate hetero-interactions. An understanding of the discrete stochastic distribution of membrane proteins in micelles and lipid bilayers is important for distinguishing between preferential and random interactions, both of which can influence the oligomeric population and perhaps the functional properties as well.

ACKNOWLEDGMENT

We thank Ann Marie Stanley for helpful discussions and critical reading of the manuscript.

REFERENCES

1. Prenzel, N., Fischer, O. M., Streit, S., Hart, S., and Ullrich, A. (2001) The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification, *Endocr. Relat. Cancer 8*, 11−31.
2. Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., and Ullrich, A. (1989) Studies of the HER-2/neu proto-oncogene in human breast cancer and ovarian cancer, *Science 244*.
3. Slamon, D., Clark, G., Wong, S., Levin, W., Ullrich, A., and McGuire, W. (1987) Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2/*neu* Oncogene, *Science 235*, 177−182.
4. Easty, D. J., Mitchell, P. J., Patel, K., Florenes, V. A., Spritz, R. A., and Bennett, D. C. (1997) Loss of expression of receptor tyrosine kinase family genes PTK7 and SEK in metastatic melanoma, *Int. J. Cancer 71*, 1061−1065.
5. Radinsky, R., Risin, Fan, Dong, Bielenberg, Bucana, and Fidler. (1995) Level and function of epidermal growth factor receptor predict the metastatic potential of human colon carcinoma cells, *Clin. Cancer Res. 1*, 19−31.
6. Shaheen, R. M., Ahmad, S. A., Liu, W., Reinmuth, N., Jung, Y. D., Tseng, W. W., Drazan, K. E., Bucana, C. D., Hicklin, D. J., and Ellis, L. M. (2001) Inhibited growth of colon cancer carcinomatosis by antibodies to vascular endothelial and epidermal growth factor receptors, *Br. J. Cancer 85*, 584−589.
7. Mossie, K., Jallal, B., Alves, F., Sures, I., Plowman, G. D., and Ullrich, A. (1995) Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosine kinase family, *Oncogene 11*, 2179−2184.
8. Chou, Y. H., and Hayman, M. J. (1991) Characterization of a member of the immunoglobulin gene superfamily that possibly represents an additional class of growth factor receptor, *Proc. Natl. Acad. Sci. U.S.A. 88*, 4897−4901.
9. Miller, M. A., and Steele, R. E. (2000) Lemon encodes an unusual receptor protein-tyrosine kinase expressed during gametogenesis in Hydra, *Dev. Biol. 224*, 286−298.
10. Pulido, D., Campuzano, S., Koda, T., Modolell, J., and Barbacid, M. (1992) Dtrk, a Drosophila gene related to the trk family of neurotrophin receptors, encodes a novel class of neural cell adhesion molecule, *EMBO J. 11*, 391−404.
11. Senes, A., Gerstein, M., and Engelman, D. M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the

GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions, *J. Mol. Biol. 296*, 921−936.
12. Russ, W. P., and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association, *J. Mol. Biol. 296*, 911−919.
13. Fleming, K. G., and Engelman, D. M. (2001) Specificity in transmembrane helix-helix interactions defines a hierarchy of stability for sequence variants, *Proc. Natl. Acad. Sci. U.S.A. 98*, 14340−14344.
14. Fleming, K. G. (2002) Standardizing the free energy change of transmembrane helix-helix interactions, *J. Mol. Biol. 323*, 563−571.
15. Fleming, K. G., Ren, C. C., Doura, A. K., Kobus, F. J., Eisley, M. E., and Stanley, A. M. (2004) Thermodynamics of glycophorin A transmembrane helix-helix association in C14 betaine micelles, *Biophys. Chem. 108*, 43−49.
16. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of Data from the Analytical Ultracentrifuge by Nonlinear Least-Squares Techniques, *Biophys. J. 36*, 575−588.
17. Laue, T. M., Shah, B., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge.
18. Fleming, K. G., Ackerman, A. L., and Engelman, D. M. (1997) The Effect of Point Mutations on the Free Energy of Transmembrane α-Helix Dimerization, *J. Mol. Biol. 272*, 266−275.
19. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E., and Engelman, D. M. (1992) Glycophorin A Dimerization is Driven by Specific Interactions between Transmembrane α-Helices, *J. Biol. Chem. 267*, 7683−7689.
20. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J., and Engelman, D. M. (1992) Sequence Specificity in the Dimerization of Transmembrane α-Helices, *Biochemistry 31*, 12719−12725.
21. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E., and Fleming, K. G. (2004) Sequence context modulates the stability of a GxxxG mediated transmembrane helix-helix dimer, *J. Mol. Biol. 341*, 991−998.
22. Doura, A. K., and Fleming, K. G. (2004) Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer, *J. Mol. Biol. 343*, 1498−1497.
23. Sulistijo, E. S., Jaszewski, T. M., and MacKenzie, K. R. (2003) Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent, *J. Biol. Chem. 278*, 51950−51956.
24. Arkin, I. T., Adams, P. D., MacKenzie, K. R., Lemmon, M. A., Brunger, A. T., and Engelman, D. M. (1994) Structural organization of the pentameric transmembrane α-helices of phospholamban, a cardiac ion channel, *EMBO J. 13*, 4757−4764.
25. Lear, J. D., Gratkowski, H., and DeGrado, W. F. (2001) Membrane active peptides, *Biochem. Soc. Trans. 29*, 559−564.
26. Tanford, C., and Reynolds, J. A. (1976) Characterization of Membrane Proteins in Detergent Solutions, *Biochim. Biophys. Acta 457*, 133−170.
27. Hall, D., and Minton, A. P. (2003) Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges, *Biochim. Biophys. Acta 1649*, 127−139.
28. Lichtenberg, D., Robson, R. J., and Dennis, E. A. (1983) Solubilization of phospholipids by detergents. Structural and kinetic aspects, *Biochim. Biophys. Acta 737*, 285−304.
29. Neugebauer, J. (1990) *A Guide to the Properties and Uses of Detergents in Biology and Biochemistry*, Calbiochem Corporation, San Diego.
30. Møller, J. V., and le Maire, M. (1993) Detergent binding as a measure of hydrophobic surface area of integral membrane proteins, *J. Biol. Chem. 268*, 18659−18672.
31. le Maire, M., Champeil, P., and Møller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents, *Biochim. Biophys. Acta 1508*, 86−111.
32. le Maire, M., Kwee, K., Anderson, J. P., and Møller, J. V. (1983) *Eur. J. Biochem. 129*, 525−532.
33. Grefrath, S. P., and Reynolds, J. A. (1974) The molecular weight of the major glycoprotein from the human erythrocyte membrane, *Proc. Natl. Acad. Sci. U.S.A. 71*, 3913−3916.
34. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The Calpha - - -H···O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9056−9061.
35. Mossie, K., Jallal, B., Alves, F., Sures, I., Plowman, G. D., and Ullrich, A. (1995) Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosiine kinase family, *Oncogene 16*, 2179−2185.
36. Waterman, H., Alroy, I., Strano, S., Seger, R., and Yarden, Y. (1999) The C-terminus of the kinase-defective neuregulin receptor ErbB-3 confers mitogenic superiority and dictates endocytic routing, *EMBO J. 18*, 3348−3358.

BI048076L