A dimerization motif for transmembrane α-helices

Mark A. Lemmon[1,3], Herbert R. Treutlein[1,2,4], Paul D. Adams[1], Axel T. Brünger[1,2] and Donald M. Engelman[1]

Specific helix-helix interactions inside lipid bilayers guide the folding and assembly of many integral membrane proteins and their complexes. We report here a pattern of 7 amino acids (LIxxGVxxGVxxT) which when introduced into several hydrophobic transmembrane α-helices promotes their specific dimerization. Dimerization is driven by interactions that are specific, dominated by the helix-helix interface, and involve no potentially ionizable groups. The motif may provide a useful tool for the functional analysis of such interactions in a variety of systems. Further, since this particular motif is rare, whilst specific helix association is not, many other such motifs may exist, which could permit sorting within complex membranes as well as guiding folding and oligomerization.

[1]Department of Molecular Biophysics & Biochemistry and [2]Howard Hughes Medical Institute, Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA
[3]Present address: Department of Pharmacology, New York University Medical Center, 550 First Avenue, New York, NY 10016, USA
[4]Present address: Ludwig Institute for Cancer Research, P.O. Royal Melbourne Hospital, Parkville, Victoria 3050, Australia.

Correspondence should be addressed to D.M.E.

Specific interactions between transmembrane α-helices guide the folding and assembly of many integral membrane proteins and their complexes[1–5]. In the case of bacteriorhodopsin (bR), which has seven transmembrane α-helices[6], the extramembraneous helix-connecting loops are not required for the correct folding of the molecule; independently stable helical fragments of bR can associate with one another to form a functional protein[7,8]. Interactions between transmembrane α-helices are also important in the assembly of a number of cell-surface receptor complexes. The TCRα and CD3δ chains of the T-cell receptor complex associate through interactions between their single transmembrane domains which involve highly polar amino acid side-chains[9,10].

Transmembrane helix association is also important in the assembly of an IgG Fc receptor[11], and specific interactions between the single transmembrane domains of the MHCII α- and β-chains are required for the correct assembly of the MHC class II complex[12]. In addition, a number of studies suggest that transmembrane helix association may play a role in both transmembrane signalling[13–16] and sorting events responsible for the subcellular localization of some integral membrane proteins[17–20]. These results suggest that the amino acid sequence of many transmembrane α-helices contains information which directs specific interactions, with functionally important consequences.

We report here a pattern of 7 amino acids (LIxxGVxxGVxxT) that is sufficient to drive the specific dimerization of several hydrophobic transmembrane α-helices into which it is introduced. The pattern was discerned in previously reported mutagenesis and computational modelling studies of specific dimerization of the single transmembrane α-helix of human glycophorin A (GpA)[21–23]. GpA forms dimers in detergent solution which are disrupted on addition of a peptide corresponding to its transmembrane domain, with concomitant formation of GpA/peptide heterodimers[24]. Heterologous transmembrane peptides have no such effect[21,24]. This behaviour was reproduced in a chimeric protein in which the transmembrane domain of GpA was fused to the carboxy-terminus of the nuclease from *Staphallococcus aureus*[21], which is normally a monomeric soluble protein. Saturation mutagenesis of this chimera at each position of the 23-residue transmembrane α-helix of GpA identified seven (sensitive) positions at which single conservative substitutions disrupted dimerization in SDS[22]. Similar single substitutions at other (insensitive) positions in the transmembrane domain had no effect upon dimerization[22]. It was concluded that the sensitive residues, which exhibit a repeat of approximately 3.9 residues, correspond to the interface of the helix dimer. The same conclusions were reached independently by computational studies of the dimer using simulated annealing techniques[23].

In the best model of the dimer (Fig. 1), the two helices associate in a right-handed supercoil. The essential elements of the interaction between the helices include packing interactions involving the hydrophobic side-chains in the motif, electrostatic interactions between the helix backbones, which are very close in the region of the two glycine residues and interactions involving the side-chain of the threonine residue. The interfacial residues describe a slight right-handed spiral on the sur-

face of each protomer (Fig. 1*a*), and close association of these residues across the dimer interface leads to the formation of a right-handed supercoil of α-helices. The glycine residues of the motif are buried in the centre of the interface (Fig. 1*b*), and aliphatic side-chains of the motif are involved in intimate packing across the dimer interface (Fig. 1*c*). The complement of amino acid residue-types found in this interface is similar to that found in interfaces between transmembrane α-helices deduced from disulphide cross-linking studies of the aspartate chemoreceptor[25] (Tar) and the leader peptidase of *Escherichia coli*[26], as well as that deduced from mutagenesis studies of the dimerization of the M13 coat protein in SDS[27].

Introduction of the motif into poly-leucine

Since individual positions outside the seven key amino acids (LIxxGVxxGVxxT) can be mutated without disrupting dimerization of the GpA transmembrane domain[22], the question arises as to whether interactions mediated by this pattern of seven residues could drive the specific dimerization of any hydrophobic α-helix. All residues in the GpA transmembrane domain (GpATM) of the nuclease/GpA chimera (SN/GpA) other than those that comprise the pattern were mutated, in stages, to leucine to generate SN/$L_{23\text{-}7}$ (Fig. 2*a*). No significant reduction in dimerization was detected at any stage. Furter mutation of Val 80 to Leu, generating SN/$L_{23\text{-}6}$ from SN/$L_{23\text{-}7}$, had no effect upon dimerization of the chimera, as was expected from the mutational analysis of SN/GpA dimerization[22]. The lack of effect of these mutations demonstrates that the sequence pattern LIxxGV/LxxGVxxT can mediate dimerization of a hydrophobic peptide that is otherwise poly-leucine (Fig. 2*b*, lane 4), the extent of dimerization being very similar to that seen with SN/GpA itself (Fig. 2*b*, lane 2).

To determine whether dimerization of SN/$L_{23\text{-}7}$ results from interactions similar to those that mediate GpATM dimerization, we generated a selection of mutations in the motif, for which the effect upon GpATM dimerization is known[22]. The effects of these mutations upon dimerization of SN/$L_{23\text{-}7}$ and SN/GpA are very similar (Table 1). Furthermore, a competition experiment, which provided the first evidence that GpA dimerization occurs via specific interactions[24] between transmembrane domains[28], was performed with SN/$L_{23\text{-}6}$. Like SN/GpA, dimers of SN/$L_{23\text{-}6}$ are disrupted by a peptide corresponding to the transmembrane domain of GpA, with the concomitant formation of peptide-protein heterodimers formed between transmembrane α-helices (Fig. 2*b*, lane 5). Heterologous transmembrane peptides disrupt neither SN/$L_{23\text{-}6}$ (Fig. 2*b*, lanes 6–8) nor SN/GpA[21] to a significant extent, indicating that the interaction is specific, and that SN/$L_{23\text{-}6}$ associates through interactions which are indistinguishable from those driving SN/GpA dimerization. The very small amount of heterodimer formation seen with the *neu*-TM and *erbB*-2TM peptides was also seen with SN/GpA (ref. 21, Fig. 3*a*).

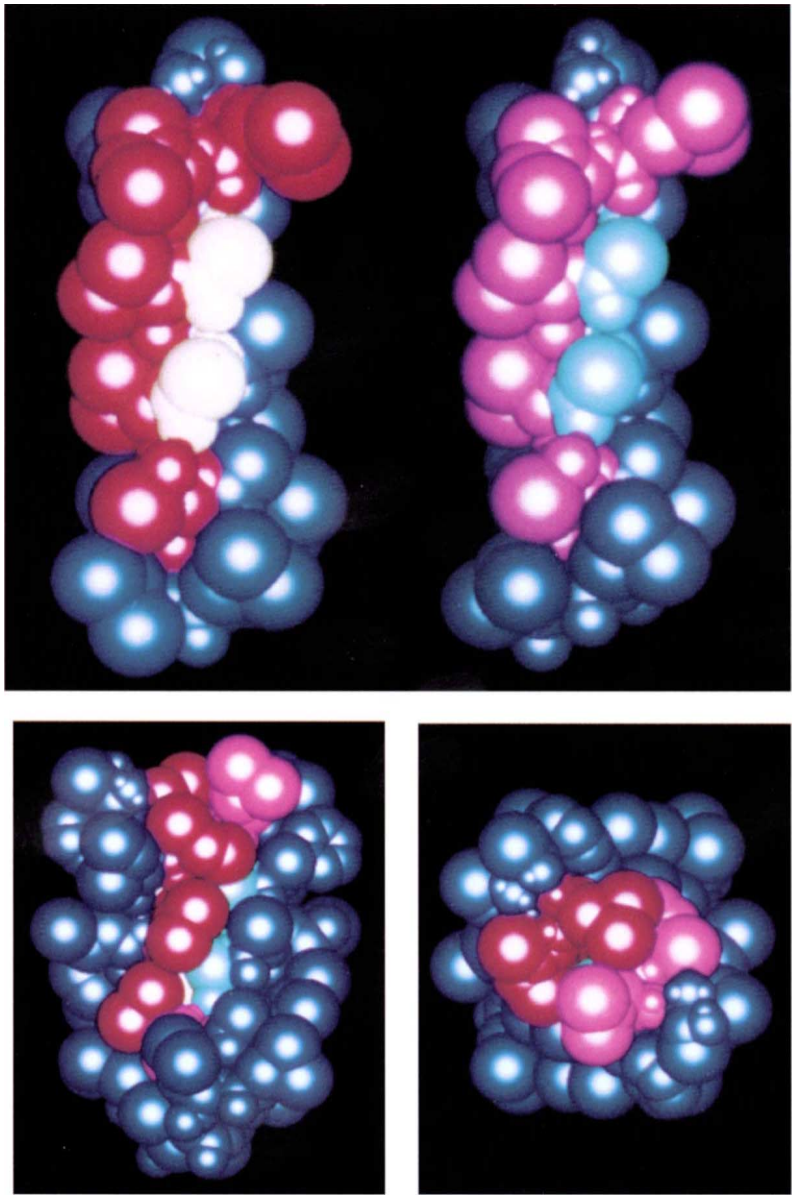
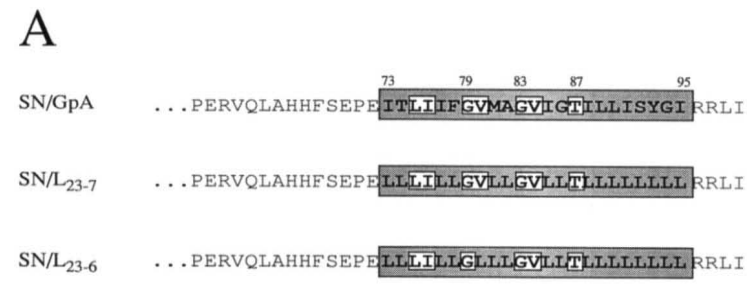


Fig. 1 A model for the dimer of the transmembrane α-helix of human GpA. *a* (top), Side-view of the two protomers, in which the residues of the motif (LIxxGVxxGVxxT) have been coloured differently from the rest of the helix, to illustrate their formation of a right-handed spiral of the surface of each helix. Residues found to be insensitive in the mutagenesis studies of SN/GpA[22] are coloured blue. Residues (except glycines) found to be sensitive are coloured red in one protomer, and purple in the other. Both sensitive glycine residues (G 79 and G 83) are coloured yellow and green respectively in the two protomers. *b* (left), A side-view of the model of the right-handed GpATM dimer[23], showing how the sensitive residues are juxtaposed in the helix-helix interface. The glycine residues are seen to be buried deeply in the centre of the interface, and shielded by the aliphatic side-chains. It is clear that few interactions other than those involving residues in the LIxxGVxxGVxxT pattern occur between the helices. *c* (right), A view of the dimer model from the amino-terminal end of the helices, showing how the leucine and isoleucine residues of the motif are proposed to pack with one another across the interface.



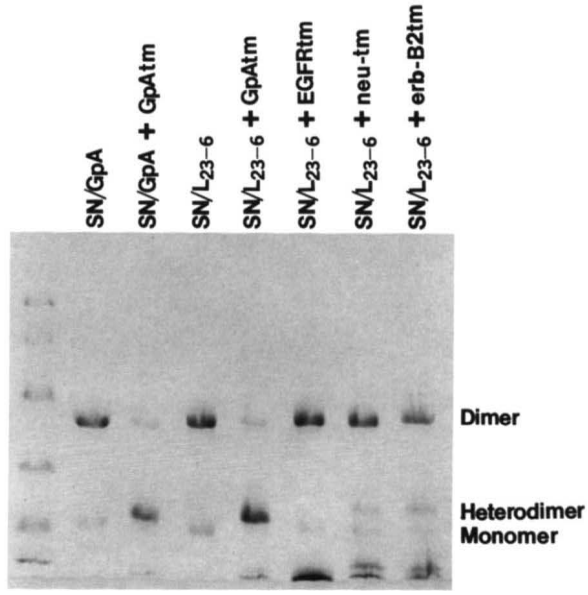


Fig. 2 Dimerization of poly-leucine after introduction of the sequence pattern LIxxGV/LxxGVxxT. *a*, Sequences of the transmembrane domains that were fused to the C-terminus of the nuclease from *S. aureus*. The transmembrane domain is enclosed in a shaded box. Residues at which conservative mutations led to the disruption of SN/GpA dimerization[22] are enclosed in unshaded boxes, and represent L 75, I 76, G 79, V 80, G 83, V 84, and T 87 of human glycophorin A. *b*, SDS-PAGE (12.5%) showing comparable levels of dimerization of SN/GpA (lane 2) and SN/$L_{23\text{-}6}$ (lane 4). Addition of a ten-fold molar excess of a peptide corresponding to the transmembrane domain of GpATM leads to disruption of the homodimer in both cases (lanes 3 & 5), the majority of the protein runs instead at a position that corresponds to a peptide-protein heterodimer[21]. By contrast, addition of the transmembrane domain of EGF-R, Neu, or erb-B2 results in little or no disruption of the dimer of SN/$L_{23\text{-}6}$. Lane 1 shows low molecular mass markers (Bio-Rad), of 97.4, 66.2, 45, 31, 21.5, and 14.4 kDa.

Computational modelling

Using computational approaches developed for the study of the dimerization of GpATM (Fig. 3*c* and ref. 23), conformational space was searched for the most stable parallel, right-handed coiled-coil dimer of poly-leucine containing the deduced dimerization motif. An 18-residue sequence ($L_{18\text{-}7}$) was used, to maintain consistency with previous studies[23]. Searches using the sequence of $L_{18\text{-}7}$ and GpATM, starting from idealized right-handed supercoils of parallel α-helices, both located a well populated global energy minimum in the region corresponding to (α=90 °, β=90 °), according to the convention of Treutlein *et al*[23] (Fig. 3a,b). This corresponds to the best model for the GpATM dimer (Fig. 1*b*), which showed remarkable agreement with the mutational analysis[22,23]. The minimum appears deeper and more clearly defined for the $L_{18\text{-}7}$ dimer than for GpATM, suggesting that it may be more stable. Simulations starting from parallel left-handed coiled-coil structures (not shown) gave minima that were not well populated and were of significantly higher energy. Thus, substitution of the sequence pattern LIxxGVxxGVxxT into poly-leucine results in a transmembrane α-helix that is predicted to dimerize in a specific and defined manner, which appears identical to that observed for GpATM.

Table 1 Comparison of the effects of mutations in the transmembrane domain upon dimerization of SN/$L_{23\text{-}7}$ and SN/GpA.

Nature of Mutation[a]	Effect upon SN/$L_{23\text{-}7}$ Dimerization[b]	Effect upon SN/GpA Dimerization[b]
L75F	+++	+++
G83S	–	–
T87S	++	++
T87V	+++	++
T87F	+	–
T87A	++	+

[a] Mutations are denoted according to the equivalent residue in GpA
[b] Effects upon dimerization were scored as follows:
– complete disruption
\+ detectable dimer remains
++ significant dimer remains
+++ mutation has no effect

Induced dimerization of receptor transmembrane domains

There has been much discussion regarding the potential role of specific interactions between transmembrane α-helices in signalling by growth factor receptors such as those for epidermal growth factor (EGF) and platelet-derived growth factor (PDGF)[2,13]. Receptor dimerization is thought to be the primary ligand-induced event leading to activation of the intracellular tyrosine kinase domain in these cases[16]. In the case of the *neu* oncogene product Neu, which is highly homologous to the EGF-receptor, a valine to glutamic acid mutation (V664E) within the single transmembrane domain results in constitutive activation[29]. This activation has been shown to correlate with enhanced receptor-dimerization[15,30], leading to suggestions that the altered transmembrane α-helix itself forms dimers. A model for the dimerization of the Neu transmembrane domain suggested that a particular sequence pattern, surrounding residue 664 of Neu, may be responsible for this dimerization, and it was argued that a similar pattern, or motif, is found in most other growth factor receptors[31]. In the case of Neu, mutagenesis studies support this suggestion[14]. Furthermore, a similar (Val to Glu) mutation has been shown to activate the *Drosophila* EGF-R homologue[32], although it is not within a proposed motif. However, analogous mutations do not activate the insulin receptor[33–35], although replacement of its entire transmembrane domain with that of the *erbB*-2 oncoprotein has been shown to result in constitutive activation of the receptor[36]. In the case of the EGF-receptor (EGF-R), which is more

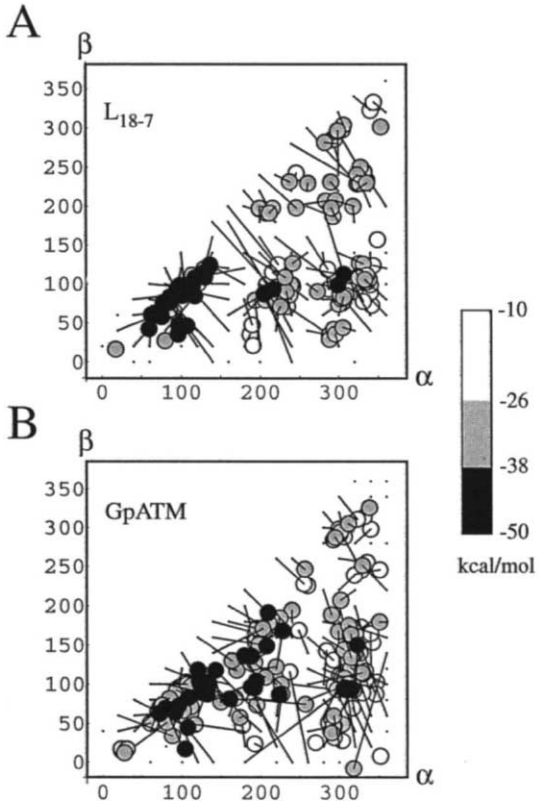


Fig. 3 A comparison of the search for low interaction-energy conformations of the dimer of GpATM and L_{18-7}. Plots of the change in configuration of the dimer of *a* (top), L_{18-7} and *b* (bottom), GpATM during simulated annealing starting from parallel right-handed supercoils with a variety of different relative orientations of the two helices, as depicted in c (right). The large circles represent the final structure, and are shaded according to the interaction energy for that structure. Darker circles represent more stable dimers, as shown in the scale. The area around α=90°, β=90° corresponds to the best model for the dimer of GpATM[23], and both plots show a clustering of low energy dimers in this region. The minimum appears better defined in the case of L_{18-7} (*a*) than of GpATM (*b*). The lowest energy final structures are seen to cluster in the region of α=90°, β=90°, with few structures of such low energy elsewhere. By contrast, as one would expect because of symmetry, no clustering of low-energy structures was seen for L_{18} (not shown). Thus, introduction of the motif LIxxGVxxGVxxT into polyleucine causes it to behave similarly to GpATM in these studies. The sequences used for these studies were: L_{18-7} LLILLGVLLGVLLTLLLL; GpATM TLIIFGVMAGVIGTILLI. *c*, Schematic representation of the methods employed for the search for the most stable dimers of L_{18-7} and GpATM (see *Methods*).

similar to Neu, neither a mutation analogous to the Neu V664E mutation nor a variety of other alterations of the transmembrane domain had any effect upon signalling[37,38], indicating a passive role for this portion of EGF-R. No studies have reported whether the various transmembrane domains are themselves capable of dimerization, and we have been unable to show dimerization of the transmembrane domains of EGF-R or Neu with or without the V664E mutation (M.A.L. & D.M.E., unpublished observations). One way of testing the hypothesis that transmembrane helix association can activate receptors of this type, and that this is the mechanism by which the V664E mutation activates Neu, would be to study receptors containing transmembrane helices that are known to dimerize.

We have introduced the sequence pattern LIxxGxxxGVxxT into the single transmembrane domains of both EGF-R and Neu (Fig. 4*a*,*c*). The pattern was introduced into the nuclease chimerae using the smallest possible number of amino-acid substitutions; where elements of the motif exist in the parent transmembrane domain, these were used to position the rest of the motif. Since no difference in behaviour was observed between SN/L_{23-7} (containing LIxxGVxxGVxxT) and SN/L_{23-6} (containing LIxxGxxxGVxxT), we did not introduce the first valine of the motif into EGF-R. Instead, an alanine residue from the EGF-R transmembrane domain sequence was left at this position, which was previously found not to be significantly disruptive in the context of SN/GpA[22]. In the case of Neu, a valine is found at this position in the native transmembrane domain. An additional mutation was made in each of the transmembrane domains upon introduction of the motif — in both cases, a glycine residue close to the C-terminus of the transmembrane domain was mutated to leucine, since glycine at these positions in SN/GpA was previously found to disrupt dimerization[22].

SN/EGFR-GpA and SN/Neu-GpA2 which both contain the motif dimerize to a significant extent, whereas chimerae containing the parent transmembrane domains (SN/EGFR and SN/Neu)do not (Fig. 4*b*,*d*). Thus the sequence pattern LIxxGxxxGVxxT can also drive the dimerization of these hydrophobic helices. One surprising finding was that SN/Neu into which only part of the motif had been introduced could also form dimers (SN/Neu-GpA1). Alanine and threonine replace leucine and isoleucine of the motif in this chimera. Both of these mutations (L75A and I76T) disrupt dimerization of SN/GpA[22] significantly, and a double mutant (L75A/I76T) also does not dimerize (data not shown). Since SN/Neu-GpA1 dimerizes to a significant extent, whereas SN/Neu does not, additional interactions, involving residues outside the introduced motif, may also contribute to

helix-helix association in this case. Further study is required to determine the origin of this effect.

All the altered transmembrane domains that were found to dimerize will also interact with GpATM in a peptide competition experiment (Fig. 5); those that do not dimerize (SN/Neu, SN/Neu* and SN/EGFR) did not interact with GpATM (data not shown). Thus, the interactions that drive dimerization of the altered transmembrane domains, as in the case of SN/$L_{23\text{-}6}$, appear to be similar to those that stabilize the GpATM dimer. These results lend further credence to the suggestion that the pattern LIxxGxxxGVxxT is sufficient for the dimerization of any hydrophobic transmembrane α-helix in a detergent environment. Since GpATM has also been shown to form specific dimers in lipid bilayers[39], and substantial changes in the surfaces contacting the lipidic environment can be tolerated, such a motif may be of use in analyzing the functional importance of intramembraneous helix-helix interactions in a number of integral membrane proteins, including the growth factor receptors discussed here. It offers a significant advantage over studies involving the swapping of domains (for example, ref. 40), in that no decision regarding the boundary of the transmembrane domain need be made and rotational relationships can be explored by positioning the motif at a series of positions along the α-helix.

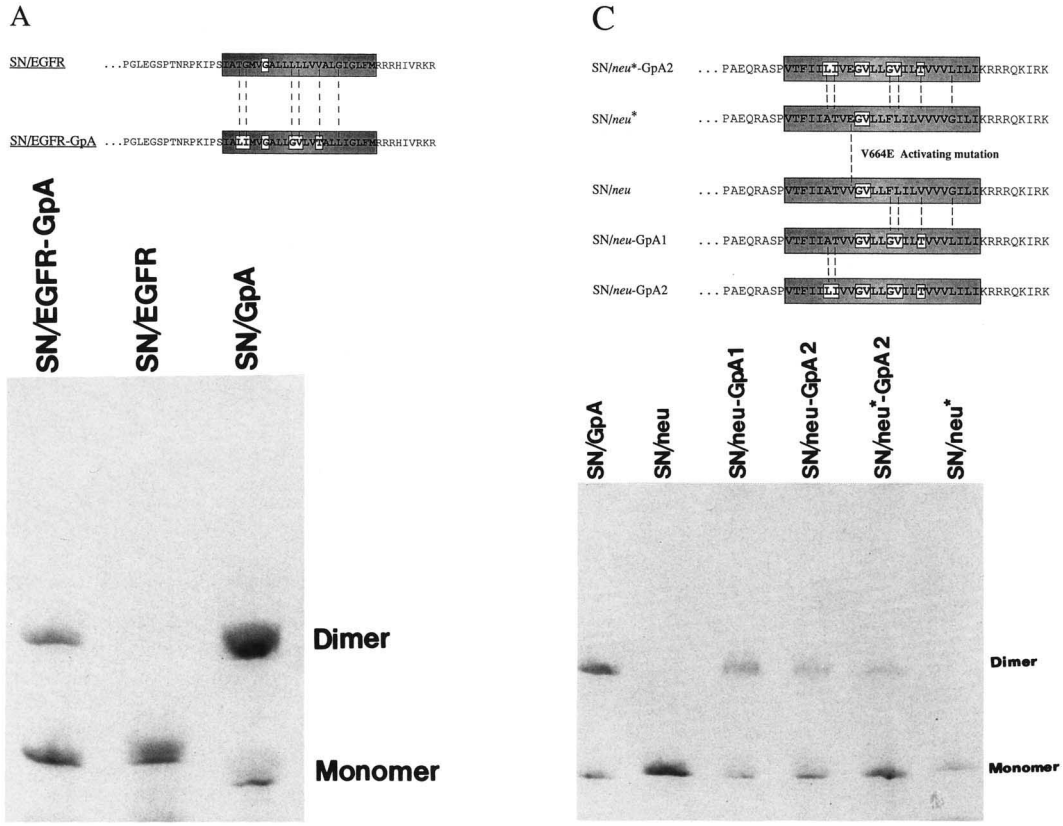


Fig. 4 Dimerization of the single transmembrane domain of the epidermal growth factor receptor (EGFR) and the *neu* oncogene product upon introduction of the sequence pattern LIxxGxxxGVxxT. *a* (top left), Sequence of the transmembrane domain of the EGF receptor that was fused to nuclease to generate SN/EGFR, and the mutations performed in order to introduce the pattern LIxxGxxxGVxxT. In generating SN/EGFR-GpA, five mutations were made to introduce the motif. An additional mutation (G to L) at the position corresponding to L90 of GpA was also made as the mutation L90G disrupted slightly the dimerization of SN/GpA[22]. The first valine residue of the motif was not introduced, since the V80A mutation was only slightly disruptive in SN/GpA[22]. *b* (bottom left), SDS-PAGE (12.5%), showing that while SN/EGFR does not form dimers in SDS (lane 2), introduction of the motif LIxxGxxxGVxxT to yield SN/EGFR-GpA results in significant dimer formation (lane 1). Dimerization of SN/EGFR-GpA is weaker, however, than that seen for SN/GpA (lane 3). *c* (top right), Sequence of the transmembrane domains of $p185^{Neu}$ and $p185^{Neu*}$, that were fused to the C-terminus of nuclease, as well as the mutations performed to introduce the pattern LIxxGVxxGVxxT into these transmembrane domains with the smallest possible number of sequence changes. Again, five substitutions were required to introduce the motif, and an additional mutation (Gly to Leu) was made at the position corresponding to I91 of GpA, since it had previously been found that the mutation I91G disrupted SN/GpA dimerization[22]. *d* (bottom right), A 12.5% polyacrylamide SDS-gel showing dimerization of the SN/Neu chimerae that contain the sequence pattern LIxxGVxxGVxxT, but not those with the native transmembrane sequences. While neither SN/Neu (lane 2) nor SN/Neu* (lane 6) showed any detectable level of dimerization, introduction of the LIxxGVxxGVxxT motif to yield SN/Neu-GpA2 (lane 4) or SN/Neu*-GpA2 (lane 5) resulted in the appearance of significant levels of dimer, although they were reduced compared with the level of SN/GpA dimerization (lane 1). In addition, introduction of part of the motif, to yield SN/Neu-GpA1 (lane 3), resulted in the appearance of a significant amount of dimer.

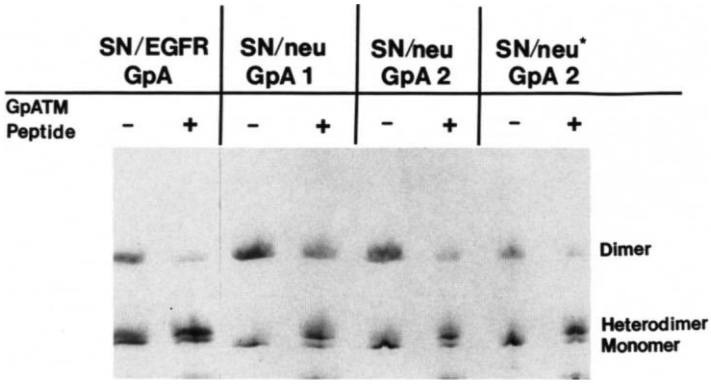


Fig. 5 12.5 % polyacrylamide SDS-gel showing that dimers of the EGFR-GpA and Neu-GpA chimerae are disrupted by addition of GpATM. To each chimera, a synthetic peptide corresponding to GpATM, with the sequence PEITLIIFGVMAGVIGTILLISYGIRRLI was added at a ten-fold molar excess. In each case, it is clear that significant disruption of the dimer results, with the concomitant formation of a peptide-protein heterodimer. Synthetic peptides corresponding to the transmembrane domains of Neu and EGFR have no such effect (see Fig. 2*b*). These observations suggest that dimerization of the chimerae occurs through interactions similar to those that stabilize the dimer of glycophorin A.

Functional Implications

A search for the sequence LIxxGVxxGVxxT in the GenBank database[41], as well as those variants predicted to dimerize from mutational analysis (Table 1 and ref 22) failed to locate any matches, although in a few cases the majority of the motif was found in presumed transmembrane domains of other integral membrane proteins. The probability of finding this motif at random in a 23-residue transmembrane α-helix may be estimated at approximately 10^{-7} if its position along the helix does not matter (10^{-6} if all tolerated mutations in the motif[22] are considered). These probabilities, and the absence from the database of any other transmembrane domain that we predict will dimerize in a manner similar to GpATM, indicate that the specificity of the interaction is sufficient to be biologically relevant. There is some biological precedent for the importance of such interactions, the signals for Golgi retention of several glycosyltransferases have recently been reported to reside in the transmembrane domains of these proteins (reviewed in refs 1, 5 and 17) and it has recently been shown that SDS-resistant oligomerization of a membrane protein, mediated by its transmembrane domain, correlates with its retention in the Golgi[18]. Furthermore, there is evidence that interactions involving transmembrane domains are important in the sorting of two integral membrane proteins of the nuclear envelope[19–20].

It is possible that a large number of oligomerization motifs exist in transmembrane a-helices (up to about 10^7 of the type described here) and that these posses degrees of specificity similar to that seen in the case of GpA. One other example is suggested by studies of interactions between the single transmembrane domains of the MHCII α- and β- chains[12]. If interactions involving all such motifs are as specific as those described here, helix-helix interactions within membranes may well be implicated in biological sorting events.

Methods

Generation of constructs. SN/$L_{23\text{-}7}$ and SN/$L_{23\text{-}6}$, were generated from pT7SN/GpA (I85L)[22] using PCR-mediated *in vitro* mutagenesis as described[22], or four-primer PCR[42]. The mutations were made in stages such that if dimerization were abolished, the mutations responsible could readily be determined. The stages were: stage 1. I88L/S92L/Y93L/G94L; stage 2. F78L/M81L/A82L; stage 3. I73L/T74L/I77L; stage 4. G86L; stage 5. I91L/I95L; stage 6. V80L; with the product of each stage serving as the template for the next. The sequence of the chimera was verified at each stage using the Sequenase kit (U.S. Biochemicals) as described[22]. At each stage, dimerization of the resulting chimera was assayed by SDS-PAGE, and found to be unaffected. The final chimerae have transmembrane domains that represent poly-leucine, into which has been introduced the sequence pattern LIxxGVxxGVxxT (SN/$L_{23\text{-}7}$), or LIxxGxxxGVxxT (SN/$L_{23\text{-}6}$) (Fig 2*a*).

Chimerae in which the transmembrane domain of Neu or EGF-R were fused to staphylococcal nuclease were generated as follows. PCR was used to amplify fragments corresponding to the transmembrane domain of Neu or EGF-R with an *Apa I* site at the 5' end and a *Bam HI* site at the 3' end. These were then digested and subcloned into the vector pSN/GpA from which the *Apa I / Bam HI* fragment corresponding to the GpA transmembrane domain had been removed[21]. The motif was introduced by two rounds of four-primer PCR mutagenesis[42] (Fig. 4*a,c*). The resulting chimerae were subcloned into the T7 expression vector pET11a (ref. 22) for high levels of expression in *E. coli*, and the sequences of the resulting constructs were verified.

Expression and purification of chimerae. The chimeric proteins were expressed in *E. coli* (ref. 21). SN/GpA, SN/EGFR, SN/EGFR-GpA, SN/Neu, and SN/Neu*, were detergent-extracted and purified as described[21,22]. SN/$L_{23\text{-}7}$, SN/$L_{23\text{-}6}$, and the altered forms of SN/Neu were refractory to this approach, and were extracted in 100 mM NaCl, 50 mM Tris-HCl, pH 8.0, containing 8M urea, from the pellet obtained after cell lysis. After clarification of the extract by centrifugation, protein was purified by reversed-phase HPLC[21,22]. Fractions containing the chimerae were lyophilized and the protein dissolved in 1% SDS solution in Tris buffer. This material could also be solubilized in β-octylglucoside (β-OG) by dissolving it in 50% aqueous trifluoroethanol containing β-OG, re-lyophilizing, and dissolving the residue in buffer. Chimerae treated in this way could be reconstituted into lipid bilayers by detergent dialysis.

SDS-PAGE. Solutions of chimerae were loaded on to 12.5 % homogeneous Phastgels (Pharmacia) at 0.3 mg/ml (15 μM) for SDS-PAGE. The loading buffer contained 2% SDS in all cases, and samples were boiled for 5 minutes prior to electrophoresis. 2–4 μl of sample were loaded. When peptides were added (Figs 2*b* & 5), they were present at a 10-fold molar excess (150 μM). Gels were stained with Coomassie blue.

Modeling. A detailed description of the methods employed in this study is presented elsewhere[23]. All simulations were carried out *in vacuo*. An initial right-handed supercoil of parallel α-helices was built, with the relative orientations of the two helices defined by the angles α and β (Fig. 3*c*). A variety of symmetric (α≈β) and asymmetric (α≠β) starting structures were tested. After energy minimization, further refinement was achieved by simulated annealing consisting of a 20 psec molecular dynamics simulation at 300 K. All atoms were free to move, although helical hydrogen-

bond restraints and restraints on the distance between helix axes were maintained. The relative orientation of the two helices changed in many cases, from (α,β) to the final (α',β') configuration. The total interaction energy between the helices was calculated for this final configuration. Lines are drawn in Fig. 3*a,b* from the starting to the final configurations, (as shown in Fig. 3*c*), with the energy of the final structure represented by the darkness of shading. Since the plots in Fig. 3*a,b* show only the region from 0° to 360°, there were cases, for starting configurations close to the extremes of the axes, in which the final configuration would appear distant from the starting structure since α' and/or β' exceeded 360° or fell below 0°. These cases would result in lines across the entire plot. Thus, for clarity, initial and final configurations are joined by lines only when α and/or β have changed by less than 200°. Lone dots in the plots therefore represent starting configurations that have led to final structures either outside the graph, or represented by lone shaded circles across the graph. Areas in which final structures appear to cluster represent energy minima. The clusters in which the energies of the final structures are lowest are likely to represent global energy minima, such as the area around α=90°, β=90° (Fig 3*a,b*).

Received 7 October 1993; accepted 27 January 1994.

Acknowledgements
We thank K.M. Ferguson and D.F. Stern for valuable discussions, and gratefully acknowledge the assistance of J. Zhang and L. Fisher. This work was supported by grants from the NIH and NSF (D.M.E.), the NSF (P.D.A.), the Howard Hughes Medical Institute (H.R.T. & A.T.B.), as well as by funds from Boehringer Ingelheim Inc., and the National Foundation for Cancer Research (D.M.E.). M.A.L. was a Predoctoral Fellow of the Howard Hughes Medical Institute, and H.R.T. was supported by the NATO Scientific Affairs Division managed by the DAAD.

1. Lemmon, M.A. & Engelman, D.M. Helix-helix interactions inside lipid bilayers. *Curr. Op. struct. Biol.* **2,** 511–518 (1992).
2. Bormann, B.-J. & Engelman, D.M. Intramembrane helix association in oligomerization and transmembrane signalling. *A. Rev. Biophys. biomolec. Struct.* **21,** 223–242 (1992).
3. Popot, J.-L. & Engelman, D.M. Membrane protein folding and oligomerization: The two-stage model. *Biochemistry* **29,** 4031–4037 (1990).
4. Popot, J.-L. Integral membrane protein structure: Transmembrane α-helices as autonomous folding domains. *Curr. Op. struct. Biol.* **3,** 532–540 (1993).
5. Lemmon, M.A. & Engelman, D.M. Specificity and promiscuity in membrane helix interactions. *Quart. Rev. Biophys.* (in the press, 1994).
6. Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. molec. Biol.* **213,** 899–929 (1990).
7. Popot, J.-L., Gerchman, S.E. & Engelman, D.M. Refolding of bacteriorhodopsin in lipid bilayers: A thermodynamically controlled two-stage process. *J. molec. Biol.* **198,** 655–676 (1987).
8. Kahn, T.W. & Engelman, D.M. Bacteriorhodopsin can be refolded from two independently stable transmembrane helices and the complementary five-helix fragment. *Biochemistry* **31,** 6144–6151 (1992).
9. Cosson, P., Lankford, S.P., Bonifacino, J.S. & Klausner, R.D. Membrane protein association by potential intramembrane charge pairs. *Nature* **351,** 414–416 (1991).
10. Manolios, N., Bonifacino, J.S. & Klausner, R.D. Transmembrane helical interactions and the assembly of the T-cell receptor complex. *Science* **249,** 274–277 (1990).
11. Kurosaki, T., Gander, I. & Ravetch, J.V. (1991). A subunit common to an IgG Fc receptor and the T-cell receptor mediates assembly through different interactions. *Proc. natn. Acad. Sci. U.S.A.* **88,** 3837–3841 (1991).
12. Cosson, P. & Bonifacino, J.S. Role of transmembrane domain interactions in the assembly of class II MHC molecules. *Science* **258,** 659–662 (1992).
13. Hendrickson, W.A. Modes of transduction. *Curr. Biol.* **2,** 57–59 (1992).
14. Cao, H., Bangalore, L., Bormann, B.-J. & Stern, D.F. A subdomain in the transmembrane domain is necessary for $p185^{neu*}$ activation. *EMBO J.* **11,** 923–932 (1992).
15. Weiner, D.B., Liu, J., Cohen, J.A., Williams, W.V. & Green, M.I. A point mutation in the *neu* oncogene mimics ligand induction of receptor aggregation. *Nature* **339,** 230–231 (1989).
16. Schlessinger, J. & Ullrich, A. Growth factor signalling by receptor tyrosine kinases. *Neuron* **9,** 383–391 (1992).
17. Nilsson, T., Slusarewicz, P., Hoe, M.H. & Warren, G. Kin recognition: A model for the retention of Golgi enzymes. *FEBS Lett.* **330,** 1–4 (1993).
18. Weisz, O.A., Swift, A.M. & Machamer, C.E. Oligomerization of a membrane protein correlates with its retention in the Golgi complex. *J. Cell Biol.* **122,** 1185–1196 (1993).
19. Smith, S. & Blobel, G. The first membrane spanning region of the lamin B receptor is sufficient for sorting to the inner nuclear membrane. *J. Cell Biol.* **120,** 631–637 (1993).
20. Wozniak, R.W. & Blobel, G. The single transmembrane segment of gp210 is sufficient for sorting to the pore membrane domain of the nuclear envelope. *J. Cell Biol.* **119.** 1441–1449 (1992).
21. Lemmon, M.A. *et al.* Glycophorin A dimerization is driven by specific interactions between transmembrane α-helices. *J. biol. Chem.* **267,** 7683–7689 (1992).
22. Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J. & Engelman, D.M. Sequence-specific dimerization of transmembrane α-helices. *Biochemistry* **31,** 12719–12725 (1992).
23. Treutlein, H.R., Lemmon, M.A., Engelman, D.M. & Brünger, A.T. The glycophorin A transmembrane domain dimer: Sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* **31,** 12726–12732 (1992).
24. Bormann, B.-J., Knowles, W. J. & Marchesi, V.T. Synthetic peptides mimic the assembly of transmembrane glycoproteins. *J. biol. Chem.* **264,** 4033–4037 (1989).
25. Pakula, A.A. & Simon, M.I. Determination of transmembrane protein structure by disulfide cross-linking: The *Escherichia coli* Tar receptor. *Proc. natn. Acad. Sci. U.S.A.* **89,** 4144–4148 (1992).
26. Whitley, P., Nilsson, L. & von Heijne, G. Three-dimensional model for the membrane domain of *Escherichia coli* leader peptidase based on disulfide mapping. *Biochemistry* **32,** 8534–8539 (1993).
27. Deber, C.M., Khan, A.R., Li, Z., Joensson, C., Glibowicka, M. & Wang, J. Val to Ala mutations selectively alter helix-helix packing in the transmembrane segment of phage M13 coat protein. *Proc. natn. Acad. Sci. U.S.A.* **90,** 11648–11652 (1993).
28. Furthmayr, H. & Marchesi, V.T. Subunit structure of human erythrocyte glycophorin A. *Biochemistry* **15,** 1137–1144 (1976).
29. Bargmann, C.I., Hung, M.-C. & Weinberg, R.A. Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain of p185. *Cell* **45,** 649–647 (1986).
30. Cao, H., Bangalore, L., Dompé, C., Bormann, B.-J. & Stern, D.F. An extra cysteine proximal to the transmembrane domain induces differential cross-linking of $p185^{neu}$ and $p185^{neu*}$. *J. biol. Chem.* **267,** 20489–20492 (1992).
31. Sternberg, M.J.E. & Gullick, W.J. A sequence motif in the transmembrane region of growth factor receptors with tyrosine kinase activity mediates dimerization. *Protein Eng.* **3,** 245–248 (1990).
32. Wides, R.J. Zak, N.B. & Shilo, B.-Z. Enhancement of tyrosine kinase activity of the *Drosophila* epidermal growth factor homolog by alterations of the transmembrane domain. *Eur. J. Biochem.* **189,** 637–645 (1990).
33. Frattali, A.L., Treadway, J.L. & Pessin, J.E. Evidence supporting a passive role for the insulin receptor transmembrane domain in insulin-dependent signal transduction. *J. biol. Chem.* **266,** 9829–9834 (1991).
34. Longo, N., Shuster, R.C., Griffin, L.R., Langley, S.D. & Elsas, L.J. Activation of insulin receptor signaling by a single amino-acid substitution in the transmembrane domain. *J. biol. Chem.* **267,** 12416–12419 (1992).
35. Yamada, K., Goncalves, E., Kahn, C.R. & Shoelson, S.E. Substitution of the insulin receptor transmembrane domain with the *c-neu/erbB2* transmembrane domain constitutively activates the insulin receptor *in vitro*. *J. biol. Chem.* **267,** 12452–12461 (1992).
36. Cheatham, B., Shoelson, S.E., Yamada, K., Goncalves, E. & Kahn, C.R. Substitution of the *erbB-2* oncoprotein transmembrane domain activates the insulin receptor and modulates the action of insulin and insulin-receptor substrate 1. *Proc. natn. Acad. Sci. U.S.A.* **90,** 7336–7340 (1993).
37. Carpenter, C.D. *et al.* Structural analysis of the transmembrane domain of the epidermal growth factor receptor. *J. biol. Chem.* **266,** 5750–5755 (1991).
38. Kashles, O. *et al.* Ligand-induced stimulation of epidermal growth factor receptor mutants with altered transmembrane regions. *Proc. natn. Acad. Sci. U.S.A.* **85,** 9567–9571 (1988).
39. Adair, B.D. An investigation into the effects of packing on the folding of integral membrane proteins. Ph.D. Thesis, Yale University, New Haven, CT, U.S.A. (1993).
40. Lee, A.W. & Nienhuis, A.W. Functional dissection of structural domains in the receptor for colony-stimulating factor-1. *J. biol. Chem.* **267,** 16472–16483 (1992).
41. Devereux, J., Haeberli, P. & Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395 (1984).
42. Higuchi, R. Recombinant PCR. in *PCR protocols, a guide to methods and applications* (eds Innis, M.A. *et al.*) 177–183 (Academic Press, San Diego, 1990)