Supporting Information

Berger et al. 10.1073/pnas.0910873107

SI Text

SI Materials and Methods. ***Design and cloning of ToxRed and DN-ToxRed constructs.*** Plasmids pET-Duet1 and pCDF-Duet were purchased from Novagen. Plasmid pGL3-basic was a gift of Andrew Tsorkas (Department of Bioengineering, University of Pennsylvania), and plasmid pMK8 was a gift of Steven Sandler (Department of Microbiology, University of Massachusetts School of Medicine, Worcester, MA). Plasmids pccGpA, pcc2B, and pccGpAG83I and maltose binding protein (MBP)-deficient MM39 cells have been described (1, 2). MM39 (DE3) cells were generated by using the DE3 lysogenization kit (Novagen) according to the manufacturer's instructions.

T4 DNA ligase, calf-intestinal phosphatase, and restriction enzymes were obtained from New England Biolabs. Platinum TAQ DNA polymerase was from Invitrogen. Primers for mutagenesis were designed by using PrimerX (www.bioinformatics.org/primerx/). All molecular biology techniques were performed according to standard procedures.

To generate mCherry-based reporters, the *ctx* promoter region was amplified from plasmid pccGpA with specific primers that introduced NheI/NcoI restriction sites. The *ctx* promoter fragment and plasmid pGL3-Basic were digested with NheI/NcoI and ligated. mCherry was amplified from pMK8 as a NcoI/BamHI fragment. The mCherry fragment and pctxGluc were digested with NcoI/BamHI and ligated to make plasmid pctxmCherry, replacing Gluc with mCherry under control of the *ctx* promoter.

To generate ToxR expression plasmids, specific primers, introducing SacI/KpnI restriction sites, were used to amplify defined segments of ToxR chimeras, including the ToxR promoter region. Amplified sequences were ligated into SacI/KpnI digested pctxmCherry plasmid, generating pToxRed.

To generate DN ToxR proteins, PCR primers were designed to introduce mutations R96K, R96L, R68K, and R68L mutations into ToxR chimeras. Each of these mutations occurs at a conserved position in the wing-turn-helix domain of ToxR responsible for DNA binding and impairs the ability of ToxR to activate reporter gene expression in *E. coli* (3). ToxR chimera expression and reporter gene activity were assayed as described below. The inactivating ToxR mutation R96K (ToxR*) was chosen for the DN assay (4).

To generate a vector coexpressing ToxR and ToxR* chimeras under control of the T7 promoter, the *ctx::mCherry* fragment from pDNToxRed was amplified as an AgeI fragment, digested along with pCDF-Duet, and ligated to create pCDF-ctxmCherry. Note that the reverse primer for mCherry removes the last 6 amino acids to avoid an internal AgeI site. The ToxR chimeras were amplified as NcoI/PstI fragments, digested, and ligated into MCS1 to create pCDF-ToxRed1. ToxR* chimeras were amplified as NdeI/PacI fragments, digested, and ligated into MCS2 to create pCDF-DNToxRed.

Expression of ToxR chimera. Constructs were introduced into chemically competent MM39 and MM39 (DE3) cells and plated onto selective LB-agar media. Individual colonies were grown overnight in LB medium containing 100 μg/mL ampicillin or 50 μg/mL spectinomycin. The following morning, the T7-based plasmids were diluted to a culture density of $OD_{800} = 0.2$ into selective LB containing 0.25–0.5 mM IPTG (OD_{800} to avoid interference from mCherry fluorescence) and allowed to grow at least 2 h (minimum $OD_{800} = 0.5$). For ToxR-based plasmids, saturated cultures were diluted into selective LB ($OD_{800} = 0.2$) the following morning and allowed to grow to a minimum $OD_{800} = 0.5$.

In vitro mCherry sample preparation and analysis. For mCherry fluorescence measurements, 200-μL aliquots of cell suspensions were transferred to a clear 96-well plate and the wells were adjusted to the equivalent OD_{800}. The cells were pelleted by centrifugation for 5 min at $3,000 \times g$, the supernatant was decanted, being sure to remove all excess liquid from the cell pellet, and the cell pellets were resuspended in 10× initial volume of FastBreak cell lysis reagent (Promega). After transfer to 1.5-mL Eppendorf tubes, the mixture was incubated at room temperature with gentle agitation for 30 min. Samples were centrifuged for 10 min, $12,000 \times g$ to remove cell debris and clarify supernatants for analysis. One hundred and fifty microliters of clarified supernatant was transferred to black, opticlear 96-well plates. mCherry emission spectra were collected by using a Molecular Dynamics plate reader with an excitation wavelength 587 nm and emission wavelengths 610–650 nm using a 610-nm cutoff. Afterwards, aliquots were transferred from black to clear 96-well plates and the absorbance measured from 450 to 750 nm.

CAT ELISAs were performed as described (2). To measure T7-based CAT expression, 200-μL aliquots of cell suspensions were transferred to clear 96-well plates, and OD_{800} of the suspensions were adjusted to equivalent values. For normalization of data, we used a disruption index as described previously (4).

Integrin constructs, reagents, and cell lines. pcDNA 3.1, hygromycin (Hyg), and zeocin (Zeo) were obtained from Invitrogen. Human cDNAs for integrin α_2, α_5, and β_1 were purchased from Open Biosystems. The Jurkat A1 cell line (β_1 integrin-null) and parent E6-1 Jurkat cell line were kindly provided by David Boettinger (Department of Microbiology, University of Pennsylvania) and Yoji Shimizu (Department of Laboratory Medicine and Pathology, University of Minnesota School of Medicine Minneapolis, MN). HEK293 freestyle cells were kindly provided by Lawrence Brass (Hematology/Oncology Division, University of Pennsylvania School of Medicine). Monoclonal antibodies HUTS-4, 6S6, BHA2.1, P1D6, and JBS5 and phycoerythrin (PE)-conjugated mouse anti-human IgG were from Millipore. Fibronectin, laminin, type-I, collagen and type-IV collagen were obtained from Sigma. Purified H1 peptide (specific for integrin $\alpha_4\beta_1$) (7) was synthesized on a Symphony peptide synthesizer (Protein Technologies), and soluble fibronectin was labeled with a Pierce FITC Labeling Kit following the manufactures instructions. Soluble FITC-type-I collagen was obtained from Sigma.

Full-length integrin constructs were amplified as BamHI/XhoI fragments and cloned into pcDNA3.1. Site-directed mutagenesis was used to create leucine and alanine mutations in the transmembrane (TM) domains of each integrin. Plasmid DNA for transfections was prepared by using the HiPrep Maxi Pure kit (Invitrogen).

Jurkat A1 cells were transiently transfected by using the TransIT Jurkat reagent (Mirus Bio) according to the manufacturer's instructions. Cells were then cultured for 72 hours in RPMI 1640, 10% FCS, Glutamax (Invitrogen), and penicillin/streptomycin (complete medium) and assayed for expression level and activity. HEK 293 freestyle cells were transiently transfected by using the FreeStyle 293fectin reagent (Invitrogen) according to the manufacturer's instructions. Cells were cultured for 72 hours in FreeStyle 293 Expression Medium (Invitrogen) and assayed for integrin expression level and activity.

Cell adhesion assay. Solutions of fibronectin (FN; 10 μg/mL), laminin (5 μg/mL), type-I collagen (5 μg/mL), type-IV collagen (5 μg/mL), or H1-peptide (10 μg/mL) were prepared in PBS with 1 mM $CaCl_2$ and 1 mM $MgCl_2$ and incubated in 96-well plates at 4 °C for at least 12 h. Unoccupied protein binding sites on the plates were blocked with 1% BSA in PBS for 1 h at room temperature prior to use. Cell cultures were washed extensively in serum-free medium and resuspended at a density of 10^7 cells/mL in binding buffer [4 mM Hepes (pH 7.4), 135 mM NaCl, 2.7 mM KCl, 3.3 mM PO4, 0.35% BSA, 0.1% glucose, and 1 mM $MgCl_2$]. To each well, 180 mL of cell suspension was added (six samples in duplicate) and eight sets of serial dilutions were made for each sample. Plates were incubated for 30 min at 37 °C, washed with Tris buffered saline [10 mM Tris (pH 7.4), 150 mM NaCl], resuspended in 150 mL of 0.1 M citrate (pH 5.4), 0.1% Triton X-100 containing 4 mM pNPP acid phosphatase substrate, and incubated for 30 min at 37 °C. The reaction was stopped by adding 50 μL of 2 M NaOH and the number of adherent cells determined by measuring the absorbance at 405 nm. As a positive control, cell suspensions containing wild-type integrin constructs were incubated with 1 mM $MnCl_2$ or 0.1 μM phorbol 12-myristate 13-acetate; as a negative control, cell suspensions containing wild-type integrin constructs were incubated without stimulation. To block adhesion, mAb 6S6, JBS5, and BHA2.1 were added at a concentration of 25 μg/mL prior to incubation. Relative adhesion was calculated as the difference between the measured absorbance at 405 nm of a given sample and the negative control (unstimulated cells) divided by the difference between the absorbance at 405 nm of the negative control and the positive control.

FACS assay for integrin activation. Cells were incubated with primary antibody (10 μg/mL) in PBS, 1% BSA for 30 min on ice. Cells were washed with PBS and incubated with PE-conjugated anti-IgG (1:100); for ligand binding, FITC-type-I collagen or FITC-FN were added (50 μg/mL). Cells were incubated for 30 min, washed in PBS, and fixed in 1% paraformaldehyde solution for 15 min. Samples were analyzed by using a FACScalibur (BD), and data analysis (sample gating, histograms, averages) was performed by using BD CellQuest Pro software.

Generating a database of TM helical dimers. The library of TM helical dimers was constructed by using 67 proteins of the 161 proteins in the Orientations of Proteins in Membranes (OPM) crystallographic database (Fig. S7). The following criteria were used in selecting the 67 crystal structures:

1. Structures were selected having a resolution of 3 Å or better.
2. Structures obtained by NMR, theoretical models, and structures with extensive heteroatoms [such as light-harvesting complexes, Protein Data Bank (PDB) 1lgh] were removed.
3. For groups of structures with greater than 90% sequence homology (such as homologues of the same protein), the structure with the highest resolution was chosen as the representative structure for that group.

TM dimers were then identified in the 67 selected proteins by using the TM regions defined by the OPM. In instances where the TM region was α-helical at the N and C termini, the TM helix was extended by a helical turn. When more than one chain was present in the biological unit (as defined by OPM), redundant dimers were filtered out by using sequence identity. Finally, dimers were defined by using an interhelical axis distance cutoff of 12 Å.

Searching for helical dimers by using a conserved sequence pattern. The conserved sequence pattern for the set of integrin transmembrane domains is α-chain: small-X_3-small-X_3-large and β-chain: large-X_3-large- X_3-small (small = G, A, T, S; large = all others). We searched the database of parallel TM dimers described above by using this sequence pattern to identify potential heterodimer structures with the key sequence features observed in the integrin heterodimer. We first generated a subfamily of structures that had the observed sequence pattern and then further narrowed this subfamily by using an interhelical distance (12 Å) and crossing angle (parallel right, with a range of 0° to −90°) constraint for the 9-residue window containing the sequence motif. The additional structure filter (interhelical distance and crossing angle) was necessary to remove helical pairs containing the sequence pattern, but not in the heterodimer interface. This filtering resulted in 24 helical pairs, half of which clustered into a tight geometric range of crossing angles (within −20° to −60°) and interhelical distances (within 8 Å).

Searching for helical dimers by using a functional perturbation index. In order to efficiently search the database of dimers, we used a distance matrix representation (where each axis point represents a residue and values are the C-α distances). For each dimer, all pairs of continuous nine residue windows (one on each helix) were correlated with the α- and β-integrin inverse perturbation data and sorted by correlation coefficient for each chain. There were three dimers that had a correlation coefficient of 0.90 on the basis of the perturbation index data. Two of the dimer structures had severe kinks, so the third dimer was chosen as the top-correlating structure. Specifically, residues 18–43 on chain L and residues 12–35 on chain M of the mitochondrial cytochrome c oxidase define the top-correlating dimer (PDB 1V55). In order to obtain a family of structures, the top-correlating dimer was used as an input to a structure-directed search on the basis of interhelical distances (within 8 Å). The resulting 69 structures have tightly clustered geometric parameters (Fig. 5B), and ten of these have the integrin consensus sequence pattern (α-chain: small-X_3-small-X_3-large and β-chain: large-X_3-large-X_3-small).

1. Russ WP, Engelman, DM (1999) TOXCAT: A measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci USA* 96:863–868.
2. Li RH, et al. (2004) Dimerization of the transmembrane domain of integrin α_{IIb} subunit in cell membranes. *J Biol Chem* 279:26666–26673.
3. Ottemann KM, Dirita VJ, Mekalanos JJ (1992) ToxR proteins with substitutions in residues conserved with OmpR fail to activate transcription from the cholera-toxin promoter. *J Bacteriol* 174:6807–6814.
4. Yin H, et al. (2007) Computational design of peptides that target transmembrane helices. *Science* 315:1817–1822.
5. Gurezka R, Langosch D (2001) In vitro selection of membrane-spanning leucine zipper protein-protein interaction motifs using POSSYCCAT. *J Biol Chem* 276:45580–45587.
6. Mccool JD, et al. (2004) Measurement of SOS expression in individual *Escherichia coli* K-12 cells using fluorescence microscopy. *Mol Microbiol* 53:1343–1357.
7. Mould AP, Humphries MJ (1991) Identification of a novel recognition sequence for the integrin $\alpha_4\beta_1$ in the COOH-terminal heparin-binding domain of fibronectin. *EMBO J* 10:4089–4095.
8. Mendrola JM, Berger MB, King MC, Lemmon MA (2002) The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J Biol Chem* 277:4704–12.
9. Duong MT, Jaszewski TM, Fleming KG, MacKenzie KR (2007) Changes in apparent free energy of helix-helix dimerization in a biological membrane due to point mutations. *J Mol Biol* 371:422–434.
10. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM (1992) Sequence specificity in the dimerization of transmembrane α-helices. *Biochemistry* 31:12719–12725.
11. Fleming KG, Engelman DM (2001) Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc Natl Acad Sci USA* 98:14340–14344.
12. Tolia NH, Joshua-Tor, L (2006) Strategies for protein co-expression in *Escherichia coli*. *Nature Methods* 3:55–64.
13. Treutlein HR, Lemmon MA, Engelman DM, Brunger AT (1992) The Glycophorin-A transmembrane domain dimer—Sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* 31:12726–12732.
14. Metcalf DG, Kulp DM, Bennett JS, DeGrado WF (2009) Multiple approaches converge on the model-based structure of the integrin $\alpha_{IIb}\beta_3$ transmembrane heterodimer. *J Mol Biol* 392:1087–1101.

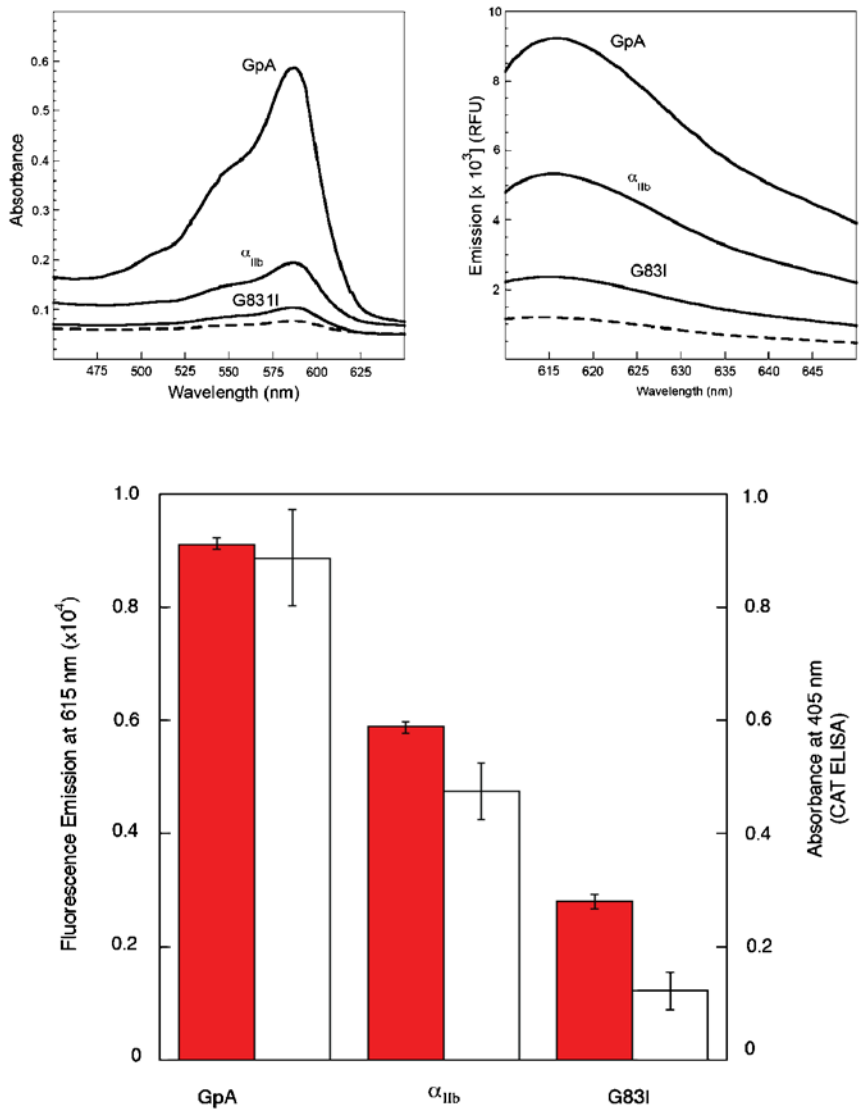

Fig. S1. Emission spectra for the ToxRed assay. Data are shown for GpA, the weakly associating GpA mutant G83I, and the integrin α_{IIb} subunit. For fluorescence measurements, samples from cell lysates were excited at 587 nm and emission spectra were collected with a 610-nm cutoff filter. The dotted line is for whole-cell lysates containing a control plasmid without a *ctx::mCherry* reporter (pcc2B). A comparison of CAT ELISA (white) and mCherry (red) emission data shows good correspondence between results using either reporter gene.

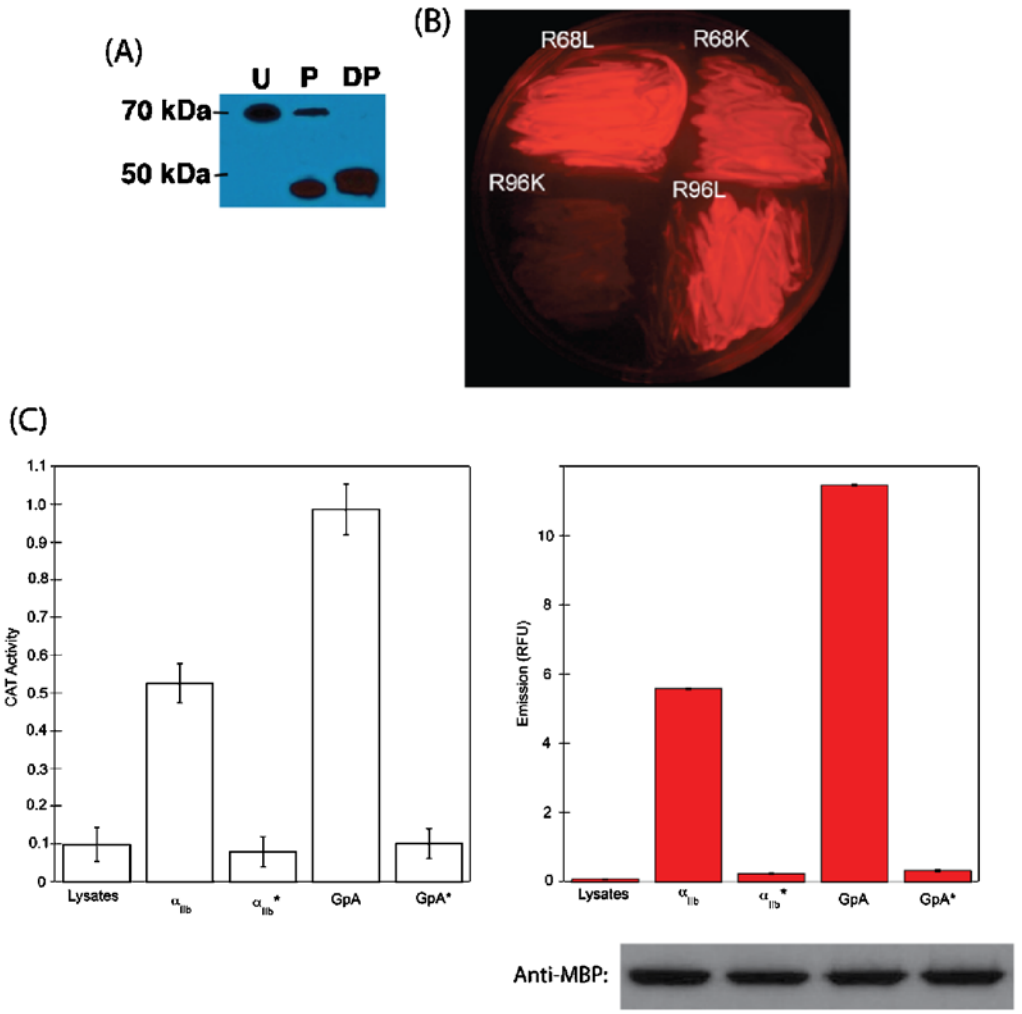

Fig. S2. Characterization of DN-ToxRed mutant expression in *E. coli.* (*A*) A fusion protein containing the ToxR mutation R96K and mCherry is inserted correctly into the inner membrane of *E. coli. E. coli* spheroplasts were prepared by using the PeriPreps Periplasting Kit (Epicentre Biotechnologies) according to manufacturer instructions. The spheroplasts were then incubated on ice for 30 minutes with 20 μg/mL proteinase K (Roche) in the absence or presence of 0.5% NP40. The spheroplasts were then dissolved in SDS, and the immunoblotting was performed by using a monoclonal, HRP-conjugated anti-MBP antibody (New England Biolabs). In the figure, U refers to untreated spheroplasts, P to spheroplasts treated with proteinase K, and DP to spheroplasts treated with NP40 and proteinase K. Incubating intact spheroplasts with proteinase K resulted in MBP cleavage, indicating that the majority of the protein was accessible on the spheroplast surface and that the ToxR-containing fusion protein was properly oriented in the inner membrane. Treatment with NP40 released the chimera from the spheroplast membrane, permitting complete cleavage of the construct by proteinase K. (*B*) The MalE complementation assay (1, 8, 9) indicates that chimeras containing GpA TM and dominant-negative ToxR mutants (R96K, R96L, R68K, R68L) are expressed and integrated into the inner membrane of *E. coli.* The constructs were grown in MM39 cells and streaked onto M9 maltose minimal plates. Since the MM39 cell line is deficient in MBP, only cells with properly integrated ToxR chimera can survive on maltose minimal media. (*C*) The ToxR mutant R96K is unable to activate reporter gene transcription. The data shown demonstrate that introducing an R96K mutation into ToxR suppresses CAT or mCherry expression when expressed as a chimera with the TM domains for α_{IIb} or GpA, unlike the corresponding chimera containing wild-type ToxR. Immunoblotting using anti-MBP confirms that similar levels of wt-ToxR and ToxR R96K are expressed for both the α_{IIb} or GpA TM chimera.

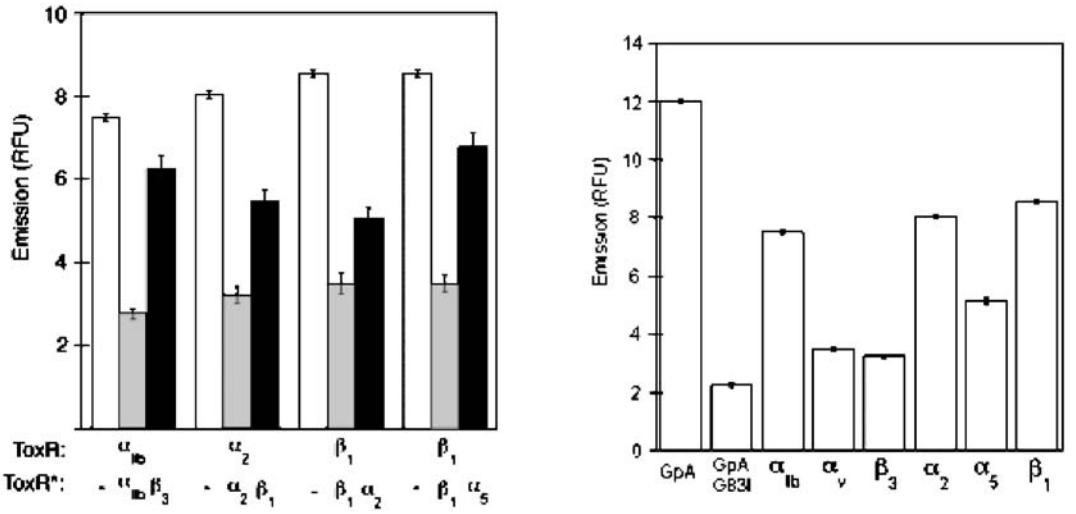

Fig. S3. Homo- and heterooligomerization of integrin α/β TM domains using DN-ToxRed. Emission values from whole-cell lysates are given as an average of at least five replicates with standard deviation.

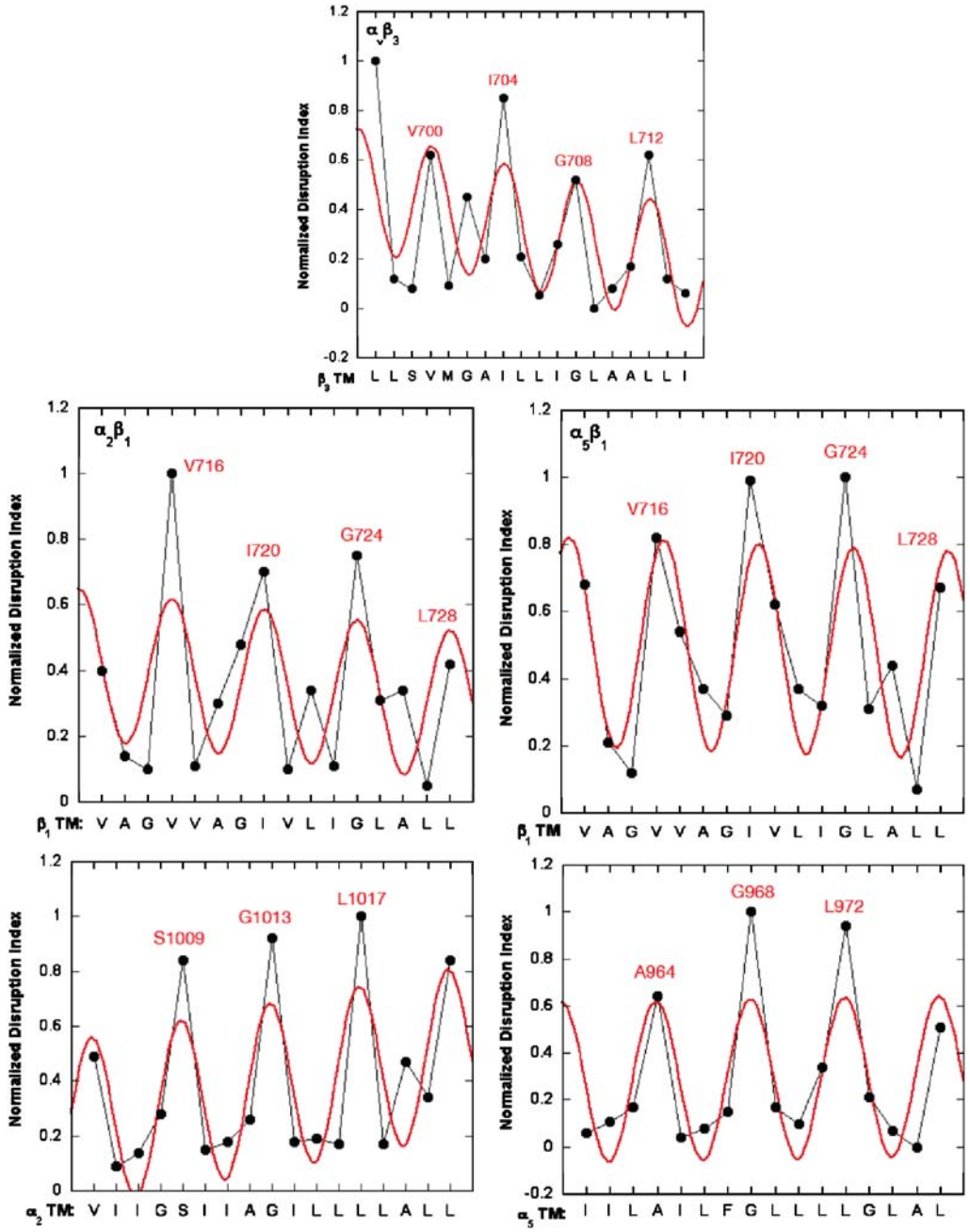

Fig. S4. Summary of alanine- and leucine-scanning mutagenesis results for heterooligomerization of $\alpha_v\beta_3$, $\alpha_2\beta_1$, and $\alpha_5\beta_1$ using DN-ToxRed. The disruption index was averaged at each position for the two mutations (Ala and Leu, except when the WT was one of these two residues), and a sinusoid function fit to the experimental data (2, 13). In each case, a four-residue periodicity was observed, and the key, disruptive residues are highlighted in red for each α/β TM domain. In each case, the sequence of the DN partner, which was subjected to mutagenesis, is indicated.

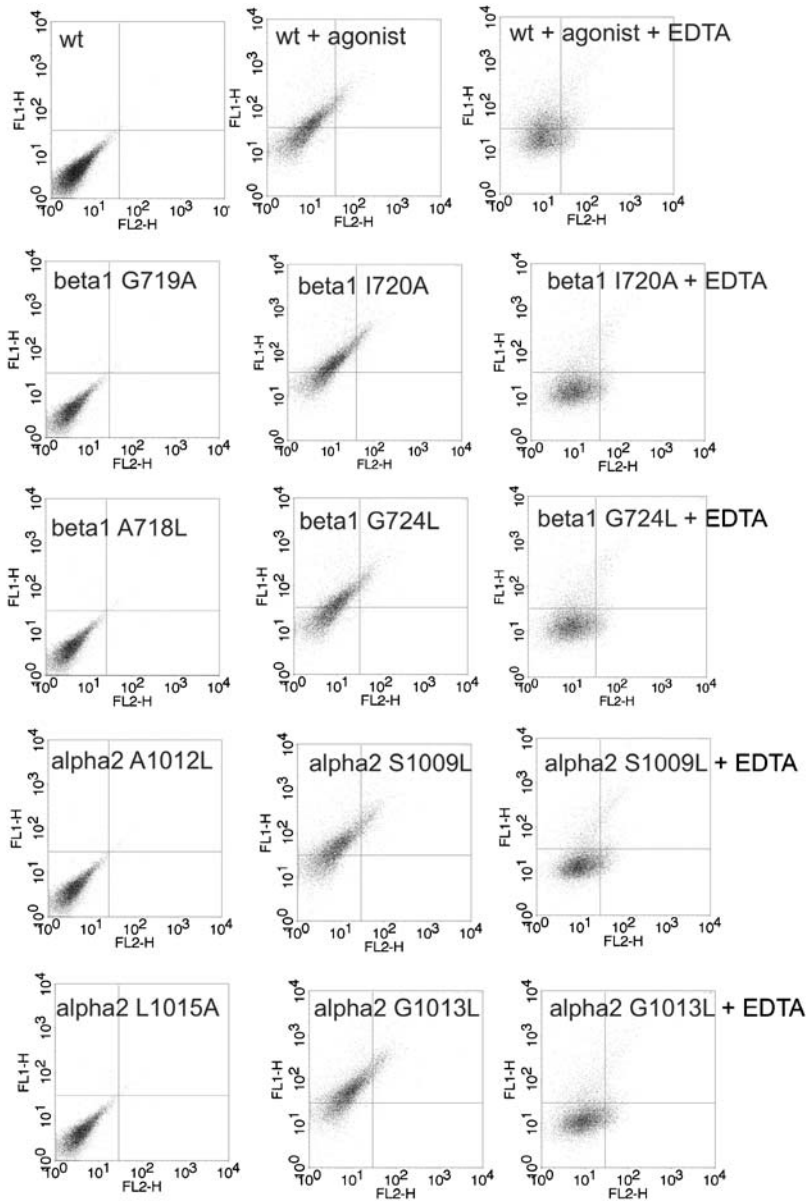

Fig. S5. Two-color FACS assay for $\alpha_2\beta_1$ integrin function. Jurkat A1 and HEK293 cells were transiently transfected with β_1 and α_2 integrin, respectively. Cells were labeled with 6S6 primary mAb, PE-conjugated secondary mAb, and FITC-type-I collagen. $MnCl_2$ was used as an agonist to activate integrins, whereas EDTA was used as an inhibitor for integrin activation. Relative binding and cell surface expression was determined by using FACS. Experimental details are provided in *SI Materials and Methods*.

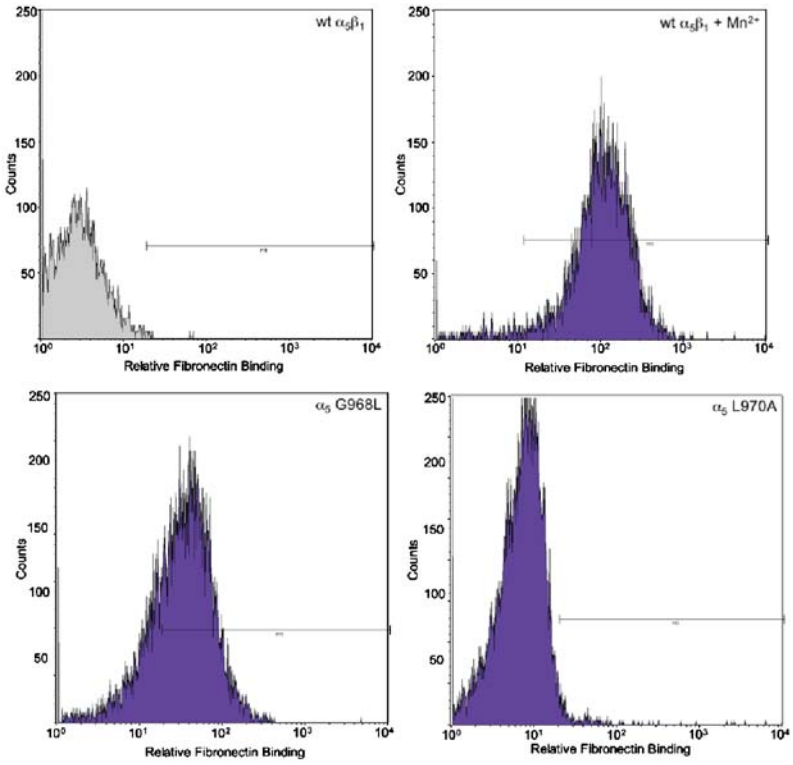

Fig. S6. One-color FACS assay for $\alpha_5\beta_1$ integrin function. HEK293 cells were transiently transfected with wild-type α_5 integrin and transmembrane domain mutants. Cells were treated with FITC-FN(9-11) and the relative binding determined by using FACS. Experimental details are provided in *SI Materials and Methods*.

Index	Structural Description	PDB	Resolution	Num.Helix
1	Archaerhodopsin-2 trimeric	2ei4	2.10Å	21
2	Halorhodopsin	1e12	1.80Å	21
3	Bacteriorhodopsin trimer	1m0l	1.47Å	21
4	Sensory Rhodopsin	1xio	2.00Å	7
5	Sensory rhodopsin II tetramer	1h2s	1.93Å	18
6	Rhodopsin	1gzm	2.70Å	7
7	Photosynthetic reaction center	1eys	2.20Å	11
8	Photosynthetic reaction center	1rzh	1.80Å	11
9	Photosynthetic reaction center	1dxr	2.00Å	11
10	Photosystem I	1jb0	2.50Å	32
11	Photosystem II	2axt	3.00Å	72
12	Light-harvesting complex LH2	1nkz	2.00Å	18
13	Light-harvesting complex	1lgh	2.40Å	16
14	Light-harvesting complex LH3	1ijd	3.00Å	18
15	Light-Harvesting Complex II	2bhw	2.50Å	9
16	Light-Harvesting Complex II	1rwt	2.70Å	9
17	Cytochrome b6f	2e74	3.00Å	26
18	Cytochrome bc1 mitochondrial	1pp9	2.10Å	26
19	Cytochrome bc1 mitochondrial	1kb9	2.30Å	24
20	Formate dehydrogenase	1kqf	1.60Å	15
21	Respiratory Nitrate Reductase	1q16	1.90Å	10
22	Fumarate reductase	2bs2	1.78Å	10
23	Succinate dehydrogenase	1nek	2.60Å	18
24	Fumarate reductase	1kf6	2.70Å	6
25	Succinate dehydrogenase	1yq3	2.20Å	6
26	Succinate dehydrogenase	1zoy	2.40Å	6
27	Bacterial cytochrome c oxidase	1ehk	2.40Å	15
28	Bacterial cytochrome c oxidase	1m56	2.30Å	22
29	Bacterial cytochrome c oxidase	1qle	3.00Å	22
30	Mitochondrial cytochrome c oxidase	1v55	1.90Å	56
31	F-type Sodium ATPase	1yce	2.40Å	22
32	V-type Sodium ATPase	2bl2	2.10Å	40
33	Particulate Methane Monooxygenase	1yew	2.80Å	42
34	Calcium ATPase E2-Pi state	1wpg	2.30Å	10
35	ABC transporter BtuCD	2nq2	2.40Å	20
36	Multidrug ABC transporter SAV1866	2hyd	3.00Å	12
37	Mitochondrial ADP/ATP carrier	1okc	2.20Å	6
38	Lactose permease	2cfq	2.95Å	12
39	Multidrug efflux transporter AcrB	2gif	2.90Å	36
40	Proton Glutamate Symport Protein	2nwl	3.00Å	39
41	Neurotrasmitter symporter	2a65	1.65Å	28
42	Ammonium transporter Amt-1	2b2f	1.72Å	33
43	Ammonia Channel	1u7g	1.40Å	33
44	Potassium channel KvAP sensor doma	1ors	1.90Å	5
45	Potassium channel KcsA	1r3j	1.90Å	12
46	Potassium channel	2ahy	2.40Å	12
47	Potassium channel KvAP membr-like er	2r9r	2.40Å	28
48	Potassium channel Kirbac3.1	1xl6	2.85Å	12
49	Aquaporin Aqpm	2f2b	1.70Å	32
50	Aquaporin-1	1j4n	2.20Å	32
51	Aquaporin-0	2b6o	1.90Å	32
52	Aquaporin Sopip2	1z98	2.10Å	32
53	Glycerol uptake facilitator	1ldf	2.10Å	32
54	Aquaporin Z	1rc2	2.50Å	32
55	Clc chloride channel	1kpl	3.00Å	28
56	Clc chloride channel	1ots	2.50Å	28
57	Steryl-sulfatase	1p49	2.60Å	2
58	Protease glpG	2nr9	2.20Å	6
59	Protease glpG	2ic8	2.10Å	6
60	Outer membrane lipoprotein Wza	2j58	2.25Å	8
61	Cytochrome c nitrite reductase comple:	2j7a	2.30Å	2
62	Lekotriene C4 synthase	2uuh	2.15Å	12
63	Acid-sensing ion channel	2qts	1.90Å	6
64	Maltose transporter	2r6g	2.80Å	14
65	Beta2-adrenergic receptor	2rh1	2.40Å	7
66	Rh-like protein	3b9w	1.30Å	24
67	M2 protein TM tetramer ligand-free	3bkd	2.05Å	4

Fig. S7. Structures used to generate the pair library. PDB identifiers, resolution, and the number of transmembrane helices used in each structure are given.