Molecular Cell, *Volume 34*

Supplemental Data

The Structure of a Receptor with Two Associating Transmembrane Domains on the Cell Surface: Integrin $\alpha_{IIb}\beta_3$

Jieqing Zhu, Bing-Hao Luo, Patrick Barth, Jack Schonbrun, David Baker, and Timothy A. Springer

Supplemental Experimental Procedures

Plasmid Construction and Transient Transfection

Plasmids encoding full-length human α_{IIb} and β_3 were subcloned into pEF/V5-HisA and pcDNA3.1/Myc-His(+), respectively (Takagi et al., 2002). To make the integrin/glycophorin A chimeras, the transmembrane and cytoplasmic domains of the α_{IIb} and β_3 subunit were replaced with human glycophorin A sequences, E60-Q131 or P71-Q131. Single amino acid substitutions to cysteine were made using site-directed mutagenesis with the QuikChange Kit (Stratagene, La Jolla, CA). The wild type or mutant constructs were transfected into 293T cells using Fugene (Roche Diagnostics, Indianapolis, Indiana).

Ligand Binding Assay

Ligand mimetic IgM PAC-1 (Becton Dickinson, San Jose, CA) and FITC-labeled human fibrinogen (Enzyme Research Laboratories, South Bend, IN) binding to transfected cells was determined as described (Luo et al., 2004). In brief, 293T cell transfectants were incubated with 10 μg/ml PAC-1 or 50 μg/ml FITC-labeled fibrinogen at room temperature for 30 minutes in Hepes-buffered saline (20 mM Hepes, pH 7.4, 150 mM NaCl, 5.5 mM glucose, and 1% bovine serum albumin) containing 5 mM EDTA, or 1 mM Ca^{2+}/Mg^{2+}, or 1mM/Mn^{2+} plus 10 μg/ml αIIbβ3 specific activating mAb PT25-2. Then cells were incubated

with 10 μg/ml of Cy3-labeled mAb AP3 on ice for 30 minutes before subjected to flow cytometry. For PAC-1 binding, cells were incubated with 10 μg/ml of Cy3-labled AP3 and 10 μg/ml of FITC-labeled goat anti-mouse IgM at the same time. Ligand binding ability was expressed as the percentage of mean fluorescence intensity (MFI) of fibrinogen or PAC-1 binding after subtracting the MFI of fibrinogen or PAC-1 binding in EDTA condition relative to the MFI of AP3 binding.

Disulfide Crosslinking and Immunoprecipitation

Twenty-four hours after transfection, 293T cells in 12-well plates with 1.5 ml DMEM medium containing 10% FCS were pre-treated with 15 μg/ml of 2-BP for 1 hour, the medium was replaced with 0.75 ml Met, Cys-free RPMI1640 (Sigma R-7513), supplemented with 10% dialyzed FCS, 10 μl [^{35}S] cysteine/methionine (10mCi/ml, PerkinElmer Life Science), 15 μg/ml 2-BP. After 1.5 h at 37°C, 0.75 ml of RPMI1640 containing 10% FCS, 500 μg/ml cysteine, 100 μg/ml methionine, and 15 μg/ml 2-BP was added, and cells chased for at least 17 hours. Cells were detached by vigorous pipetting, washed, and suspended (10^6 cells in 100 μl) in Tris-buffered saline (TBS, 20 mM Tris-HCl, pH 7.5, 150 mM NaCl) containing 1 mM Ca^{2+}/1 mM Mg^{2+} and proteinase inhibitors (1 μg/ml each aprotinin, leupeptin, and pepstatin). The cells were kept intact or broken by 3 cycles of freezing on dry ice and thawing. Saponin (40 μg/ml) gave results identical to freeze-thawing, but freeze-thawing was adapted as the least membrane-perturbing. After chilling on ice for 5 minutes, 200 μM $CuSO_4$/1000 μM o-phenanthroline was added by 10 fold dilution from stock solution, and cells were incubated on ice for another 10 minutes. N-ethylmaleimide (10 mM) was added and after 10 minutes on ice, cells were lysed with an equal volume of TBS containing 2% Triton X-100 and 0.1% NP-40 for 10 minutes on ice. Cell lysates were cleared by centrifugation at 14000 RPM for 10

minutes and immunoprecipitated with anti-β3 mAb AP3 and protein G agarose at 4°C for 1 hour (Luo et al., 2004). The precipitated proteins were subjected to nonreducing 7.5% SDS-PAGE. The SDS-PAGE gel was dried and exposed for 3 h to storage phosphor screens, which were measured with a Storm PhosphorImager (Molecular Dynamics, Sunnyvale, California, United States). Disulfide bond formation was quantitated as the intensity of the disulfide-bonded heterodimer band divided by the sum of the intensity of α_{IIb}, β_3, and heterodimer bands. Specific intensity of each band was determined by subtraction of background intensity.

For constitutively crosslinked extracellular and exofacial residues, crosslinking was also measured in redox buffer and after DTT treatment followed by Cu-phenanthroline. For redox buffer treatment, cells were suspended in pH 8.2 TBS containing 1 mM Ca^{2+}/1 mM Mg^{2+} and 5 mM cysteamine/1 mM cystamine, and incubated at 37 °C for 1 hour. Following addition of 10 mM N-ethylmaleimide, cells were lysed and immunoprecipitated as described above. For DTT treatment, cells were incubated in pH 8.2 TBS containing 1 mM Ca^{2+}/1 mM Mg^{2+} and 10 mM DTT for 10 min at 37 °C, washed three times with TBS containing 1 mM Ca^{2+}/1 mM Mg^{2+}, and broken by freeze/thaw and treated with Cu-phenanthroline as described above.

Constitutive crosslinking for GPA residues 73-78 showed only slight periodicity, whereas redox buffer and DTT/Cu-phenanthroline gave similar, periodic peaks (Fig. S1A). Furthermore, the crosslinking results for residues that were dependent on Cu-phenanthroline, such as 79-80, were similar with or without DTT pretreatment (Fig. S1A). Therefore, results with DTT/Cu-phenanthroline were used for restraint calculation for all crosslinks involving GPA residues 73-78 and are shown in Fig. 2. In contrast, exofacial residues in integrins

showed good periodicity under constitutive conditions (Fig. 2), and we did not use data with reducing agents (Fig. S1B) in restraint calculation because DTT can activate integrins (Peerschke, 1995).

Disulfide-Based Distance Constraints

Crosslinking data and restraints are in Supplementary Materials. To minimize effect of flexibility or cysteine substitution on crosslinking, all values were subjected to near-valley correction before restraint calculation. The % crosslinking of residues i with j was subtracted by the higher of ((the lowest % crosslinking found for residue i with residues j ±2) or (the lowest % crosslinking found for residue j with residues i ±2)). This reduced broadening of crosslinking patterns (Fig. S5). Only crosslinking efficiencies equal to or larger than 20% were used for the calculation of upper boundary C_α-C_α distances. The distance constraints for GPA were symmetrized (Table S4).

Membrane Structure Generation

Structures were generated in three or four stages as shown in Fig. S6. In stage 1, 23-residue TM sequences were divided into three 12-residue, overlapping segments, denoted N-terminal (N) (α_{IIb} I966-L977, β_3 I693-I704, GPA I73-V84), middle (M) (α_{IIb} G972-L983, β_3 S699-A710, GPA G79-L90) and C-terminal (C) (α_{IIb} L977-W988, β_3 I704-W715, GPA V84-I95) (Fig. S6). The 12-residue N α_{IIb} with N β_3, M α_{IIb} with M β_3, etc. ideal helices were docked with an all-atom membrane force field (Barth et al., 2007) with the following command line: *Rosetta.gcc gd 1pdb chain i.d. –spanners 1pdb.span –fake_native –membrane –dock –dock_mcm –randomize1 –randomize2 –ex1 –ex2 –s 1pdb.pdb –nstruct 10000 –paths paths.txt.* This docking protocol samples backbone rigid-body and all side-chain conformational degrees of freedom. The 1,000 docked models with lowest Rosetta all-

atom energy were selected from 10,000 total decoys and pairwise disulfide-based constraint violations were calculated. 100 models with low pairwise constraints violation (in practice all 100 had no violations) were selected as the starting structures for stage 2. For integrins, both N and M helices yielded >100 docked structures with no restraint violations. Very similar 46-residue integrin TM segments (1.2 Å RMSD) were obtained after stage 3 starting with either the N or M 12-residue helices. The C 12-residue helix was not tested because of weaker disulfide restraints in this region. For GPA, 12-residue helix docking with the N, M, and C segments gave for the 100 models with lowest restraint violation, violation scores of 0, 3.96±2.21, and 3.45±0.71, respectively. Therefore, only the N segment was used for GPA structure generation.

In subsequent "chain growing" stages, the backbone coordinates of two or more residues were kept from the previous stage and remaining residues were rebuilt from N-terminal to C-terminal or from C-terminal to N-terminal at low-resolution by Monte-Carlo-based peptide fragment insertion and then refined at all-atom using disulfide distance constraints in both low resolution and all atom potentials (Barth et al., 2007). After use of the integrin M-helices in stage 1, residues L974-W988, M701-W715 (stage 2), L978-L959 and L705-P688 (stage 3), and L970-P998 and L697-A737 (stage 4) were grown in α_{IIb} and β_3, respectively. GPA 12-residue N-helices were I73-V84 in stage 1, and in stages 2 and 3, residues A82-I73 and G79-I95 were grown, respectively.

Fragments used in fragment assembly were generated as described (Rohl et al., 2004), except that only SAM-T02 (Karplus et al., 2003) secondary structure prediction method was used during fragment selection. The "chain growing" stage was performed with the following command line: *Rosetta.gcc lg 1pdb chain i.d. –fake_native –membrane –*

no_filters –loops –fold –l pdblist –nstruct 10000 –paths paths.txt –hb_srbb_reweight 0.0 –hb_lrbb_reweight 0.0 –pc_reweight 1.0 –FA_pc_weight 1.0 –short_rang_hb_weight 0.5 –use_fold_constrints_no_minimize –minimize_exclude_helix –fa_refine –ex1 –ex2 –exlaro –extrachi_cutoff 0 –thickness 15 –steepness 10 –mem_solv –memb_hb –mem_env –Wmbenv 0.462289. In each stage, 10,000 models were generated (100 decoys for each of the 100 starting structures) and 100 low energy, low constraint violation models were selected as starting structures for the following stage.

The final 10% lowest energy models were clustered based on Cα-RMSD (Bonneau et al., 2002). Clustering was robust to the RMSD cutoff used. In clustering with 1.5, 1.8, 2.0, 2.5, and 3.0 Å cutoffs, similar top (most populous) clusters were found, and the structures at the center of the 1.8, 2.0, 2.5 and 3.0 Å top clusters were identical (Fig. S7). Clustering used TM+JM+CT α_{IIb} I966-P998 and β_3 I693-E731 residues or GPA TM residues 73-95, and results are summarized in Fig. S6B,D.

For comparison, the coordinates for the GPA solid state NMR structure (Smith et al., 2001) were kindly provided by Dr. S. O. Smith, SUNY Stony Brook, Stony Brook NY. Coordinates for previous integrin models were provided by Dr. K. Gottschalk, Ludwig-Maximilians U., München, Germany (Gottschalk, 2005), and Dr. J. Bowie, UCLA, Los Angeles, CA (Partridge et al., 2005).

Cross-Validation by Restraint Omission

Restraints were randomly partioned into approximately equal subsets (before symmetrization for GPA) of 1/8, 1/4, or 1/2 of restraints. These smaller subsets of 12.5%, 25%, or 50% of restraints, or no restraints, were used in all stages of structure generation, exactly as when all restraints were used. At least two groups of each size of restraint set were

used in structure generation. In each run of structure generation, the final 10% lowest energy structures were clustered at a 3.0 Å cutoff. The central structure in the largest cluster was examined by superposition to the GPA NMR model or GPA or integrin models made with complete restraints to calculate RMSD. Models were also scored for violation of the omitted restraints. RMSD and restraint violations were calculated over the TM segment (46 residues) or TM+cytoplasmic segments (78 residues). The root mean square distance violation (above the upper bound) per restraint was calculated for all omitted restraints (whether violated or not).

Modeling of the Full-Length Integrin Receptor in the Context of the Membrane Bilayer

A modeling protocol was developed to assemble different domains in a membrane environment, to find low-energy conformations for the linkers consistent with disulfide crosslinking data, and to optimize the rigid-body orientations of each domain with respect to the membrane. The full-length integrin receptor structure was modeled in four stages.

In the first stage, the structures of A958-W968 of α_{IIb} and E686-V695 of β_3 were rebuilt by chain growing protocol from C to N with the structures of the TM+cytoplasmic domains in the largest cluster (52 structures, Figure 4A and Table S6B) as initial stubs. Redox buffer disulfide restraints were used for the residues L959-P695 of α_{IIb} and P688-D692 of β_3 (Table S3B). These restraints were validated based on crosslinking between nearby residues defined in $\alpha_{IIb}\beta_3$ and $\alpha_V\beta_3$ crystal structures; because of flexibility in crystal structures of the β_3 tail domain (Zhu et al., 2008), restraints were loosened relative to the TM region (Fig. S12). Additional restraints were added from the $\alpha_{IIb}\beta_3$ crystal structure. Cα-CαCβ-Cβ□ □C̃C and O-O atom distances between residues α_{IIb} A958 and β_3 E686 and between residues α_{IIb} L959 and β_3 C687 were used. The upper and lower bound distance

restraints were obtained by adding 2 Å and subtracting 1 Å, respectively, from distances in the $\alpha_{IIb}\beta_3$ crystal structure (Zhu et al., 2008). All the restraints for TM+cytoplasmic residues were also kept. 100 models were generated for each starting structure. The 10% lowest energy models were selected and clustered (RMSD cutoff of 2 Å) based on the structures of the rebuilt residues. The center models of top 10 clusters were used for the next stage in full length model generation.

In a second stage, the crystal structure was joined to the linker segments. The lowest energy cluster center models of the TM and cytoplasmic (TM+CT) regions were assembled to the crystal structure of the ectodomains (chains A and B) with one extended linker connecting the domains in chain A (command line used: *rosetta.intel -assemble -regions -extend -s A2B3_xtal_tmhAB.pdb -nstruct 1*). This step involved selecting residues on chain A (residues 958 to 959) connecting the TM+CT and extracellular regions and modeling these residues as an extended polypeptide chain connecting the two regions. The other chain was left unclosed.

In a third stage, the linkers were remodeled simultaneously to optimize the rigid-body orientation of the ectodomain with respect to the membrane bilayer and the TM domains, and the remaining cut was closed. The redox buffer disulfide crosslinking restraints for the linker segments (Table S3B) were used as one of the energy terms (command line used: *rosetta_assemble_new.intel CE A2B3 -pose1 -pose_memb -jumping -pairing_file pairings.dat -membrane_cst_reweight 0 -membrane_exposure_reweight 0 -s A2B3.pdb -regions -regionfile region -spanfile A2B3.span -increase_cycles 0.5 -cut_for_assemble -memb_assemble -n A2B3.pdb -num_TMH_pairings 2 -skip_closure -skip_cut 953 -chain_break_reweight 50.0 -nstruct 20 -dump_pdb -paths paths.txt.)*. To keep the relative

backbone orientations of each chain in both TM and ectodomains fixed, a “fold tree” was constructed for the polypeptide chain where two Cα positions in each domain were connected and fixed in space during folding . The fold tree was designed so that the TM regions stayed fixed and embedded in the membrane while the ectodomain region was allowed to adopt different conformations with respect to the membrane by sampling the space as a rigid-body. The conformational space sampled by the ectodomain was dictated by the conformations of the loops connecting the TM+CT to the ectodomain. The connecting loops were modeled by Monte-Carlo-based insertion of 3-mer peptide fragments into α_{IIb} and β_3 residues 958-965 and 688-692, respectively, while simultaneously scoring the entire receptor structure. This procedure allowed discarding loop conformations that would embed the ectodomain deeply in the membrane. A total of 6000 coarse-grained models were generated by modeling the loops connecting the TM+CT to the ectodomain. During this procedure, the original orientations of the linker fragments were altered. The models were clustered and the lowest energy cluster centers were refined at all-atom by sampling backbone / side-chain and side-chain conformational degrees of freedom for the remodeled loops and the remaining regions of the entire integrin structure, respectively.

In the fourth and final stage, two membrane proximal loops that are missing in density in calf-2 domain were remodeled in the context of the entire receptor embedded in the membrane. Residues 838-855 and 858-874 in a loop cleaved by furin, and 762-777 in another loop were added, and modeled as flexible loops that could conform to the orientation with respect to the membrane.

Supplemental Tables

Table S1. Comparison of Omitted Distance Restraints from Active Integrin Cysteine Mutants and the Average Cα-Cα Distance from the Cluster 1 Structure Ensemble

αIIb	β3	Crosslinking (%)	Corrected crosslinking (%)	Idealized distance (Å)	upper bound distance (Å)	Cα-Cα distance (52 models) (Å)
G976	**M701**	62.0	49.0	6.03	13.43	4.29±0.51
G976	**L698**	34.0	31.0	7.82	15.85	8.67±0.59
G976	**G702**	53.7	30.7	6.56	15.89	7.26±0.53
G976	**V700**	41.3	28.3	7.36	16.21	7.10±0.58
G976	**L697**	36.0	22.0	7.70	17.05	8.81±0.64
G976	**I707**	33.5	22.5	7.86	17.25	10.26±0.50
L978	**G708**	34.7	20.2	7.78	17.29	11.40±0.50
L979	**G708**	37.0	28.0	7.63	16.25	9.75±0.61
L980	**G708**	58.0	39.5	6.29	14.71	6.67±0.56
L983	**G708**	41.5	37.5	7.34	14.98	9.18±0.86
V984	**G708**	41.5	37.5	7.34	14.98	9.25±0.76
K994	**I719**	27.5	27.5	8.24	16.32	9.88±1.28
K994	**D723**	21.5	20.0	8.62	17.32	11.25±1.10
R995	**D723**	33.3	28.3	7.87	16.21	8.09±1.14

Table S2. Integrin Juxtamembrane Structure Characteristics among the 10% Lowest-Energy (500 Models) and Cluster 1 Ensemble

	β3 K716 ε-amino hydrogen bonding to αIIb F992 or K994 carbonyl oxygen[a]	αIIb R995 to β3 D723 side chain hydrogen bonding[a, b]	Turn at αIIb G991[c]	αIIb F992 or F993 membrane embedding[d]	β3 K716 and αIIb F992 or F993 side chain contact at 6 (5, 4) Å[e]
All 500 models	31%	16%	91%	81%	72% (60%, 38%)
Cluster 1 [f] (52 models)	65%	30%	100%	92%	90% (73%, 53%)

[a] Distance for hydrogen bonding cutoff is 3.5 Å.
[b] At least one hydrogen bond between the three guanido N atoms of R995 and two side chain oxygens of D723.
[c] Turn at G991 was defined by dihedral angle -60° <ψ<180° and 0°<φ<180°.
[d] Membrane embedding was defined as a distance of the C_ζ atom of F992 or F993 from the membrane center of less than 17 Å.
[e] At least one contact between K716 side chain carbons and F992 or F993 side chain carbons less than the indicated distance.
[f] The largest cluster with 2.0 Å cutoff.

Table S3A. Crosslinking efficiency and calculated Cα-Cα distance

Integrin TM+JM+CT					
αIIb	β3	Crosslinking(%)	Corrected[a] crosslinking(%)	idealized[b] Cα-Cα distance(Å)	Upper boundary[c] Cα-Cα distance(Å)
W968	I693	100.0	100.0	3.60	6.60
G972	L697	77.0	77.0	5.07	9.68
G972	V700	76.0	63.0	5.14	11.56
W968	L697	62.3	62.3	6.01	11.65
V971	L697	61.5	61.5	6.06	11.76
V969	V700	62.5	59.5	6.00	12.03
G972	V696	56.7	56.7	6.37	12.40
G972	M701	67.7	54.7	5.67	12.67
G972	L698	66.3	53.3	5.76	12.86
L983	L709	54.5	53.0	6.51	12.90
V984	L712	55.0	52.0	6.48	13.03
P998	E731	80.0	49.0	4.88	13.43
V969	L697	48.5	44.0	6.90	14.10
V973	V700	69.3	44.3	5.56	14.06
W968	V696	48.3	48.3	6.91	13.53
G975	M701	51.3	44.0	6.72	14.10
N996	D723	65.0	43.5	5.84	14.17
R997	F730	60.0	41.0	6.16	14.51
P998	F730	75.0	41.0	5.20	14.51
R997	I719	49.0	40.0	6.86	14.64
L979	L705	44.3	38.3	7.16	14.87
R997	D723	70.0	37.0	5.52	15.04
L979	M701	51.0	34.0	6.74	15.44
R997	E726	63.0	34.0	5.97	15.44
V973	L697	47.3	33.3	6.97	15.54
L980	L705	38.0	32.0	7.57	15.71
R997	F727	61.0	32.0	6.10	15.71
P998	F727	82.0	32.0	4.75	15.71
N996	E726	52.0	31.0	6.67	15.85
N996	F727	50.0	31.0	6.80	15.85
G975	L697	43.0	29.0	7.25	16.11
L980	A711	51.7	28.7	6.69	16.15
V984	A711	29.5	28.0	8.11	16.25
G975	V700	51.3	26.3	6.72	16.48
V969	V696	32.0	27.5	7.95	16.32
V969	L698	30.5	25.0	8.05	16.65
V971	V700	25.0	25.0	8.40	16.65
L980	L712	34.5	25.0	7.79	16.65
V984	L713	34.5	25.0	7.79	16.65
R997	E731	43.0	24.0	7.25	16.78
L983	L713	25.5	23.5	8.37	16.85
L980	I704	28.5	22.5	8.18	16.99
L983	L712	25.5	22.5	8.37	16.99
W968	L698	30.3	22.0	8.06	17.05
N996	A728	41.0	22.0	7.38	17.05
N996	F730	34.0	22.0	7.82	17.05
G972	L694	20.3	20.3	8.70	17.28
N996	R724	41.0	20.0	7.32	17.32

GPA					
GPA TM in					
αIIb	β3	Crosslinking(%)	Corrected[a] crosslinking(%)	idealized[b] Cα-Cα distance(Å)	Upper boundary[c] Cα-Cα distance(Å)
G79	V80	93.4	88.9	4.02	8.09
G79	G79	91.8	87.3	4.13	8.31
G83	G83	86.3	83.0	4.47	8.88
G83	V84	74.5	67.2	5.23	11.00
V80	V80	73.3	66.7	5.31	11.07
V80	G83	65.3	61.7	5.82	11.74
T87	T87	62.3	59.7	6.01	12.00
I88	I91	60.5	58.5	6.13	12.16
I91	I91	60.0	53.0	6.16	12.90
I88	L90	65.0	52.0	5.84	13.03
G79	G83	52.3	47.8	6.65	13.59
L90	I91	63.0	45.5	5.97	13.90
L75	I76	54.3	44.8	6.53	14.00
L75	L75	63.0	44.5	5.97	14.04
V84	V84	41.7	38.0	7.33	14.91
I88	I88	48.0	38.0	6.93	14.91
L90	L90	61.0	38.0	6.10	14.91
I76	G79	35.0	35.0	7.76	15.31
T87	I88	37.3	31.7	7.61	15.76
I76	I76	33.0	29.3	7.89	16.08
G86	G86	28.3	28.0	8.19	16.25
A82	G83	29.0	26.5	8.14	16.45
G86	T87	23.8	23.4	8.48	16.86
I91	S92	30.0	23.0	8.08	16.92
S92	G94	31.5	22.0	7.98	17.05
G79	A82	26.3	21.8	8.31	17.07
I73	L75	35.0	21.0	7.76	17.19

[a] Corrected crosslinking efficiency was calculated as described in the supplement and Fig. S5; only ≥20% crosslinking was used for restraint generation.
[b] Idealized Cα-Cα distance was calculated with the uncorrected crosslinking efficiency.
[c] Upper boundary Cα-Cα distance was calculated with the corrected crosslinking efficiency.

Table S3B. Crosslinking efficiency and calculated Cα-Cα distance of linker region

Integrin extracellular linker					
αIIb	β3	Crosslinking(%)	Corrected[a] crosslinking(%)	idealized[b] Cα-Cα distance(Å)	Upper boundary[c] Cα-Cα distance(Å)
L959	P688	100.0	100.0	3.60	10.60
E960	P688	98.0	98.0	3.73	10.79
R962	P691	56.0	22.0	6.42	17.93
A963	P691	34.0	30.5	7.82	17.13
I964	P691	57.5	37.5	6.32	16.48
I964	I693	89.5	84.5	4.27	12.06
I964	V696	42.0	42.0	7.31	16.05
P965	P691	96.0	87.5	3.86	11.78
P965	I693	97.0	94.0	3.79	11.16

[a] Corrected crosslinking efficiency was calculated as described in the supplement and Fig. S5; only ≥20% crosslinking in redox buffer was used for restraint generation.
[b] Idealized Cα-Cα distance was calculated with the uncorrected crosslinking efficiency.
[c] Upper boundary Cα-Cα distance was calculated with the corrected crosslinking efficiency.

Table S4. Format of Cα-Cα distance constraints for GPA transmembrane structure generation

res1	atom1	res2	atom2	upper	native
7	CA	31	CA	8.09	#G79 V80
7	CA	30	CA	8.31	#G79 G79
11	CA	34	CA	8.88	#G83 G83
11	CA	35	CA	11.00	#G83 V84
8	CA	31	CA	11.07	#V80 V80
8	CA	34	CA	11.74	#V80 G83
15	CA	38	CA	12.00	#T87 T87
16	CA	42	CA	12.16	#I88 I91
19	CA	42	CA	12.90	#I91 I91
16	CA	41	CA	13.03	#I88 L90
7	CA	34	CA	13.59	#G79 G83
18	CA	42	CA	13.90	#L90 I91
3	CA	27	CA	14.00	#L75 I76
3	CA	26	CA	14.04	#L75 L75
12	CA	35	CA	14.91	#V84 V84
16	CA	39	CA	14.91	#I88 I88
18	CA	41	CA	14.91	#L90 L90
4	CA	30	CA	15.31	#I76 G79
15	CA	39	CA	15.76	#T87 I88
4	CA	27	CA	16.08	#I76 I76
14	CA	37	CA	16.25	#G86 G86
10	CA	34	CA	16.45	#A82 G83
14	CA	38	CA	16.86	#G86 T87
19	CA	43	CA	16.92	#I91 S92
20	CA	45	CA	17.05	#S92 G94
7	CA	33	CA	17.07	#G79 A82
1	CA	26	CA	17.19	#I73 L75
30	CA	8	CA	8.09	#G79 V80
30	CA	7	CA	8.31	#G79 G79
34	CA	11	CA	8.88	#G83 G83
34	CA	12	CA	11.00	#G83 V84
31	CA	8	CA	11.07	#V80 V80
31	CA	11	CA	11.74	#V80 G83
38	CA	15	CA	12.00	#T87 T87
39	CA	19	CA	12.16	#I88 I91
42	CA	19	CA	12.90	#I91 I91
39	CA	18	CA	13.03	#I88 L90
30	CA	11	CA	13.59	#G79 G83
41	CA	19	CA	13.90	#L90 I91
26	CA	4	CA	14.00	#L75 I76
26	CA	3	CA	14.04	#L75 L75
35	CA	12	CA	14.91	#V84 V84
39	CA	16	CA	14.91	#I88 I88
41	CA	18	CA	14.91	#L90 L90
27	CA	7	CA	15.31	#I76 G79
38	CA	16	CA	15.76	#T87 I88
27	CA	4	CA	16.08	#I76 I76
37	CA	14	CA	16.25	#G86 G86
33	CA	11	CA	16.45	#A82 G83
37	CA	15	CA	16.86	#G86 T87
42	CA	20	CA	16.92	#I91 S92
43	CA	22	CA	17.05	#S92 G94
30	CA	10	CA	17.07	#G79 A82
24	CA	3	CA	17.19	#I73 L75

Supplemental Figure Legends

Figure S1. Disulfide Crosslinking with Redox Buffer and DTT Treatment

For constitutively crosslinked extracellular and exofacial residues, crosslinking was measured in redox buffer and after DTT treatment followed by Cu-phenanthroline as described in methods. The crosslinking patterns were plotted for GPA (A) and integrin (B).

Figure S2. Crosslinking in the Outer Membrane Leaflet Is Indistinguishable in Intact and Broken Cells

293T cells were transiently transfected with the indicated integrin cysteine mutants and metabolically labeled. Disulfide crosslinking was induced with Cu-phenanthroline in intact (-) or freeze-thawed broken cells (+). Integrin heterodimers were immunoprecipitated from the cell extract with mAb AP3 and subjected to nonreducing SDS-PAGE and autoradiography (upper panel) and phosphorimaging for quantitation of crosslinking efficiency (lower panel).

Figure S3. Kinetics of Crosslinking

Crosslinking was measured as described in Methods, except treatment with Cu-phenanthroline was for 0, 10, 30 or 60 mins.

Figure S4. Crosslinking between Nonidentical GPA TM Residues with Different Integrin Subunit Fusion Partners (Symmetry Test)

293T cells were co-transfected with αIIb/GPA71-131 G79 or V80 cysteine mutants and the indicated β3/GPA71-131 cysteine mutants. Alternatively, 293T cells were co-transfected with β3/GPA71-131 G79 or V80 cysteine mutant and the indicated αIIb/GPA71-131 cysteine mutants. Crosslinking efficiency was measured as described in Methods. The plots show that GPA TM crosslinking was independent of the integrin subunit fusion partner, demonstrating a lack of effect of the integrin fusion partner on the GPA structure, and suggesting that

symmetry was maintained in the GPA TM domain despite coupling to the asymmetric integrin ectodomain.

Figure S5. "Near Valley Correction" of Disulfide Crosslinking Efficiency

To compensate for inherent flexibility or structural perturbations introduced by cysteine mutation, the lowest crosslinking efficiency within 2 residues of crosslinked residues i or j was subtracted from the crosslinking efficiency of residue pair i and j as described in Methods, before restraint calculation.

Figure S6. Integrin and GPA Structure Generation

Procedure for structure generation of αIIbβ3 integrin TM, JM and cytoplasmic domains (A and B) and GPA TM domain (C and D). Residues of TM domains are in blue and marked by dashed lines. TM regions were divided into three overlapping, 12-residue segments (blue lines labeled N, M, or C for N-terminal, middle, C-terminal, respectively). The 12-residue segments were fixed as ideal α-helices with rigid sidechains and docked with the Rosetta global docking protocol in a modeled membrane slab (stage 1). The color codes in the schematic show in subsequent stages which residues were kept from the previous stage and which were grown. Arrows show the direction of chain growing. The 1,000 models in the first stage, which had already been selected to be among the lowest 10% in energy, were scored with the disulfide-based distance constraints, and 100 models with the lowest (or no) pairwise disulfide constraints violations were selected as the starting structures for chain growing. In chain growing stages 2-4, backbones of the first two or more residues of the starting structures were fixed, and structures of the remaining residues were built with the Rosetta fragment assembly protocol with energy scoring function for membrane protein and disulfide crosslinking-based distance constraints. Between each stage of building 5,000 to

10,000 models, 100 models with low disulfide restraint violation and low energy (not including disulfide restraint violation penalties) were selected for the next stage. At the final stage, the 10% low energy models were clustered based on root-mean-square deviation (RMSD) of C_α atoms for residues in the range, α_{IIb} I966-P998 and β_3 I693-E731 or GPA 73-95. Plots show the correlation of Rosetta all-atom energy (without disulfide restraint penalties) versus disulfide constraints violation scores in stages 1-4. The 100 models selected in successive stages are shown in magenta, and those in the largest cluster in the final stages are shown in red. The five most populous integrin (B) and GPA (D) clusters are summarized. Integrin results are shown both for clustering with 2.0 and 3.0 Å cutoffs, with the latter results in parentheses. The energies shown in A-D do not include disulfide restraint violation penalties.

Figure S7. Center Models of the 5 Largest Clusters

Models and sequence are presented and colored as described in Figure 4. A and B. Central integrin models (from a total of 500 low energy structures) are shown both after clustering at 2.0 Å (A) and 3.0 Å (B) as described in the Fig. S6 legend. Clusters are after 4 stages, whereas the cluster structures shown in Fig. 4A are after a 5^{th} stage of growing extracellular linkers with redox buffer disulfide restraints. C. Central GPA models (from a total of 1000 low energy structures) are after clustering at 1.0 Å.

Figure S8. Comparison of TM Structures Generated with and without Disulfide-Based Distance Constraints

A. Superposition of GPA NMR structure (red) and GPA models generated without disulfide-based distance constraints (green, center models of top 5 clusters). B. Superposition of integrin TM structures generated with (red, center model of the largest cluster) or without

disulfide-based distance constraints (green, center models of top 5 clusters). Structures without distance restraints were generated as described in Figure S6, except structures to seed the next stage were selected based only on Rosetta energy.

Figure S9. Superposition of the 500 Low-Energy $\alpha_{IIb}\beta_3$ Models on the TM Domains

This represents the ensemble after 4 stages, whereas the cluster structures shown in Fig. 4A are after a 5th stage of growing the extracellular linkers with redox buffer disulfide restraints.

Figure S10. Sequence Alignment of Integrin TM and Juxtamembrane Regions

All human integrin α and β subunit sequences are aligned. The boundary of TM and juxtamembrane regions is marked with a dashed line. Green stars indicate residues showing disulfide crosslinking peaks in αIIb and β3. The conserved small amino acids (G, A or S) in both α and β TM domains are marked in yellow and correspond to the Gly in α_{IIb} and β_3 TM interfaces. The Lys or Arg conserved at the boundary of TM and juxtamembrane region for β integrins are marked in cyan.

Figure S11. Ramachandran Plots for Residue α_{IIb} Gly-991

A. All 500 low energy models. B. The cluster 1 structural ensemble. The figure was generated by the RAMPAGE server (Lovell et al., 2003). In B, the position of the final cluster center structure is shown in red.

Figure S12. Disulfide Crosslinking Efficiency within Regions Defined in Integrin Ectodomain Crystal Structures Near the Linker to the TM Domains and Correlation with Distance

Disulfide crosslinking efficiency in redox buffer is plotted against distance in $\alpha_{IIb}\beta_3$ (Zhu et al., 2008) and $\alpha_V\beta_3$ (Xiong et al., 2002) crystal structures. The specific residues tested in $\alpha_{IIb}\beta_3$ or homologous residues in $\alpha_V\beta_3$ are indicated. Note that a disulfide was introduced by

mutation to cysteine of L959 and P688 in the $\alpha_{IIb}\beta_3$ crystal structure and these residues must by definition be close in the $\alpha_{IIb}\beta_3$ structure; however, the homologous I955 and P688 residues in $\alpha_V\beta_3$ are similarly close. To calculate distances in $\alpha_V\beta_3$, the α_{IIb} and α_V sequences were aligned; the sequence in the α_V structure extends more C-terminally than in α_{IIb}. The upper bound distance used in restraints for building ectodomain models is shown as a dashed line.

Figure S13. Lack of Relationship among Three NMR $\alpha_{IIb}\beta_3$ Cytoplasmic Domain Structures and the Disulfide/Rosetta Structure

NMR structures with the indicated PDB ID codes (blue) were superimposed on the juxtamembrane/cytoplasmic portion of the Disulfide/Rosetta structure (red) using Cα atoms of residues α_{IIb} 989-997 and β_3 716-737. Structures are from (Vinogradova et al., 2002) (A) and (Weljie et al., 2002) (B and C).

Supplemental References

Barth, P., Schonbrun, J., and Baker, D. (2007). Toward high-resolution prediction and design of transmembrane helical protein structures. Proc. Natl. Acad. Sci. U.S.A *104*, 15682-15687.

Barth, P., Wallner, B., and D., B. (2009). Prediction of membrane protein structures with complex topologies using limited constraints. Proc. Natl. Acad. Sci. U.S.A *106*, 1409-1414.

Bonneau, R., Strauss, C.E., Rohl, C.A., Chivian, D., Bradley, P., Malmstrom, L., Robertson, T., and Baker, D. (2002). De novo prediction of three-dimensional structures for major protein families. J Mol Biol *322*, 65-78.

Bradley, P., and Baker, D. (2006). Improved beta-protein structure prediction by multilevel optimization of nonlocal strand pairings and local backbone conformation. Proteins *65*, 922-929.

Gottschalk, K.E. (2005). A coiled-coil structure of the αIIbβ3 integrin transmembrane and cytoplasmic domains in its resting state. Structure *13*, 703-712.

Karplus, K., Karchin, R., Draper, J., Casper, J., Mandel-Gutfreund, Y., Diekhans, M., and Hughey, R. (2003). Combining local-structure, fold-recognition, and new fold methods for protein structure prediction. Proteins *53 Suppl 6*, 491-496.

Lovell, S.C., Davis, I.W., Arendall, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: phi, psi and Cβ deviation. Proteins *50*, 437-450.

Luo, B.-H., Springer, T.A., and Takagi, J. (2004). A specific interface between integrin transmembrane helices and affinity for ligand. PLoS Biol. *2*, 776-786.

Partridge, A.W., Liu, S., Kim, S., Bowie, J.U., and Ginsberg, M.H. (2005). Transmembrane domain packing stabilizes integrin aIIbb3 in the low affinity state. J. Biol. Chem. *280*, 7294-7200.

Peerschke, E.I. (1995). Regulation of platelet aggregation by post-fibrinogen binding events. Thromb. Haemostas. *73*, 862-867.

Rohl, C.A., Strauss, C.E., Misura, K.M., and Baker, D. (2004). Protein structure prediction using Rosetta. Methods in enzymology *383*, 66-93.

Smith, S.O., Song, D., Shekar, S., Groesbeek, M., Ziliox, M., and Aimoto, S. (2001). Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers. Biochemistry *40*, 6553-6558.

Takagi, J., Petre, B.M., Walz, T., and Springer, T.A. (2002). Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. Cell *110*, 599-611.

Vinogradova, O., Velyvis, A., Velyviene, A., Hu, B., Haas, T.A., Plow, E.F., and Qin, J. (2002). A structural mechanism of integrin $a_{IIb}b_3$ "inside-out" activation as regulated by its cytoplasmic face. Cell *110*, 587-597.

Weljie, A.M., Hwang, P.M., and Vogel, H.J. (2002). Solution structures of the cytoplasmic tail complex from platelet αIIb- and β3-subunits. Proc. Natl. Acad. Sci. USA *99*, 5878-5883.

Xiong, J.P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S.L., and Arnaout, M.A. (2002). Crystal structure of the extracellular segment of integrin αVβ3 in complex with an Arg-Gly-Asp ligand. Science (New York, N.Y *296*, 151-155.

Zhu, J., Luo, B.H., Xiao, T., Zhang, C., Nishida, N., and Springer, T.A. (2008). Structure of a Complete Integrin Ectodomain in a Physiologic Resting State and Activation and Deactivation by Applied Forces. Mol. cell *32*, 849-861.

Supplemental Figures

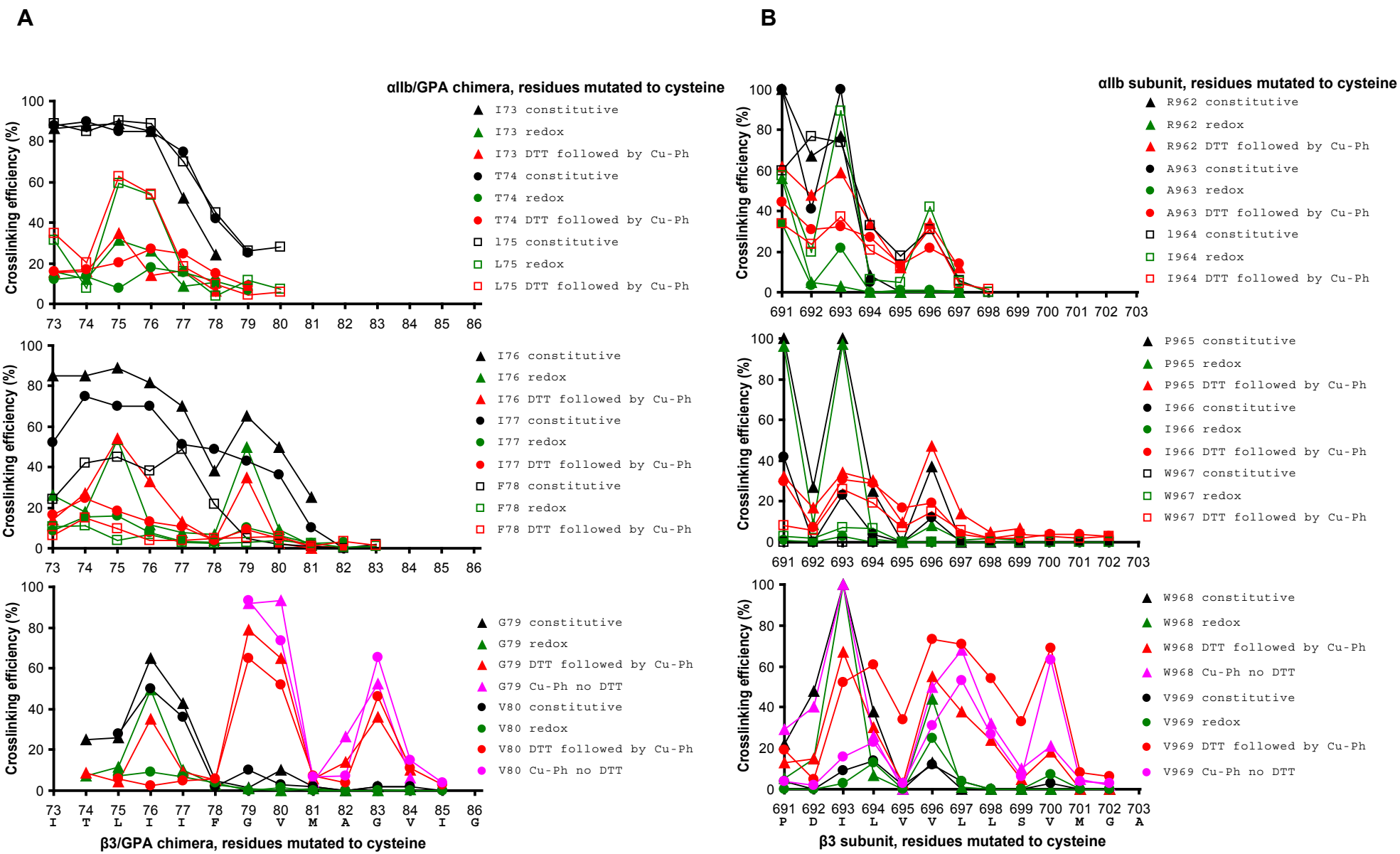
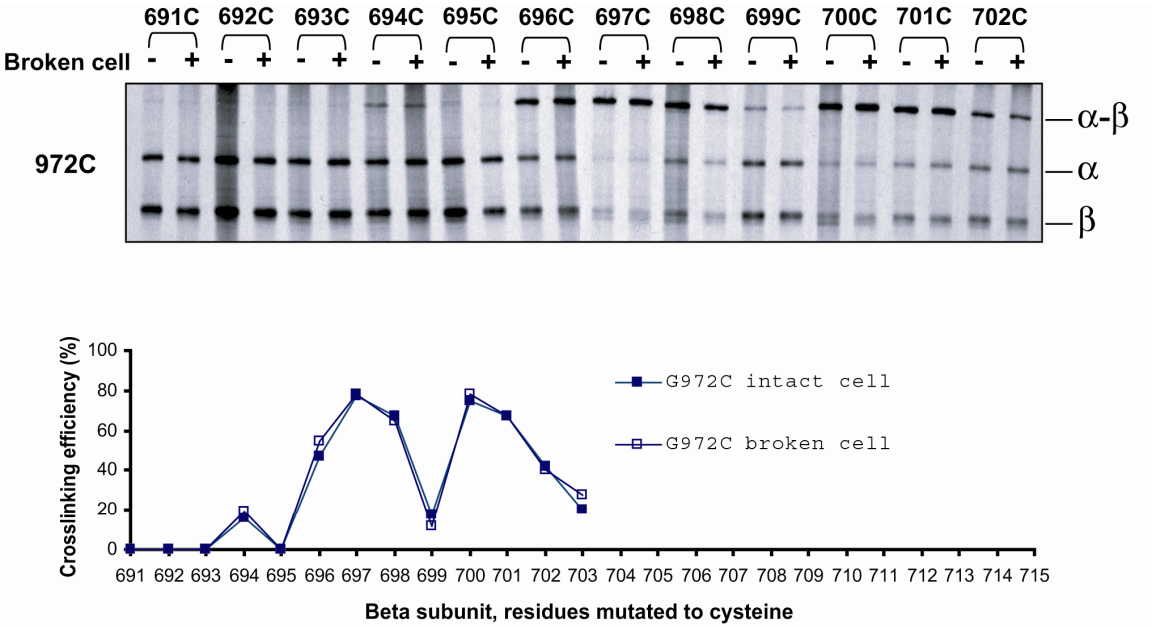


Figure S1


691C 692C 693C 694C 695C 696C 697C 698C 699C 700C 701C 702C
Broken cell - + - + - + - + - + - + - + - + - + - + - + - +
972C
α-β
α
β

Crosslinking efficiency (%)
100
80
60
40
20
0
G972C intact cell
G972C broken cell
691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715
Beta subunit, residues mutated to cysteine

Figure S2

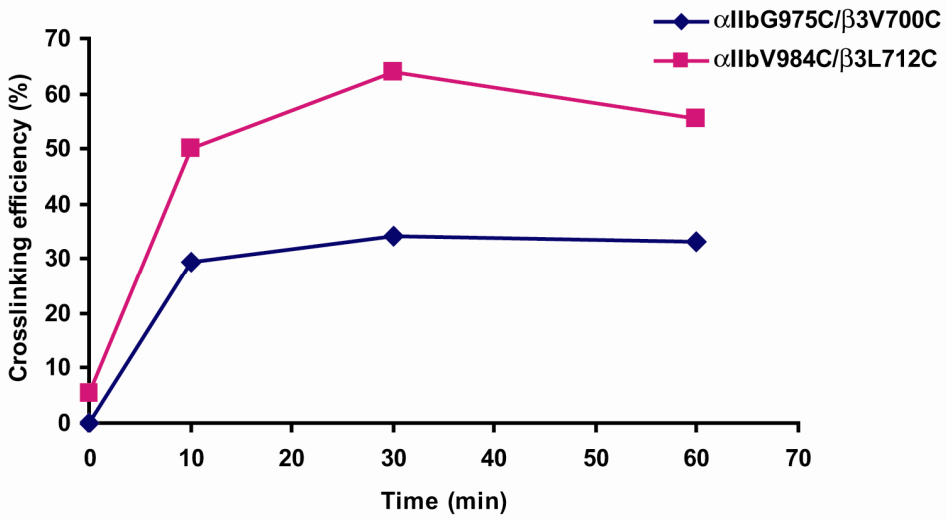

αIIbG975C/β3V700C
αIIbV984C/β3L712C
Crosslinking efficiency (%)
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
Time (min)

Figure S3

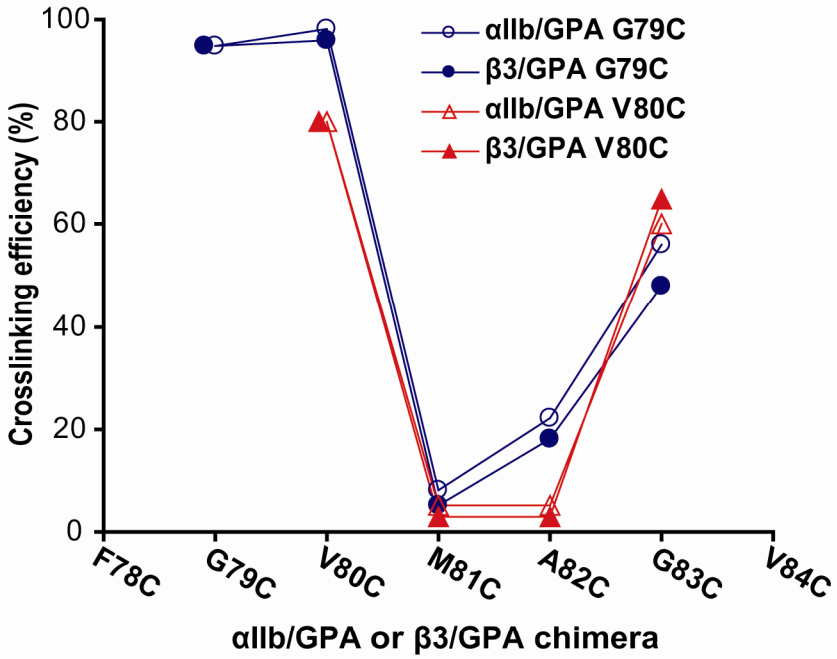

αIIb/GPA G79C
β3/GPA G79C
αIIb/GPA V80C
β3/GPA V80C
Crosslinking efficiency (%)
100
80
60
40
20
0
F78C
G79C
V80C
M81C
A82C
G83C
V84C
αIIb/GPA or β3/GPA chimera

Figure S4

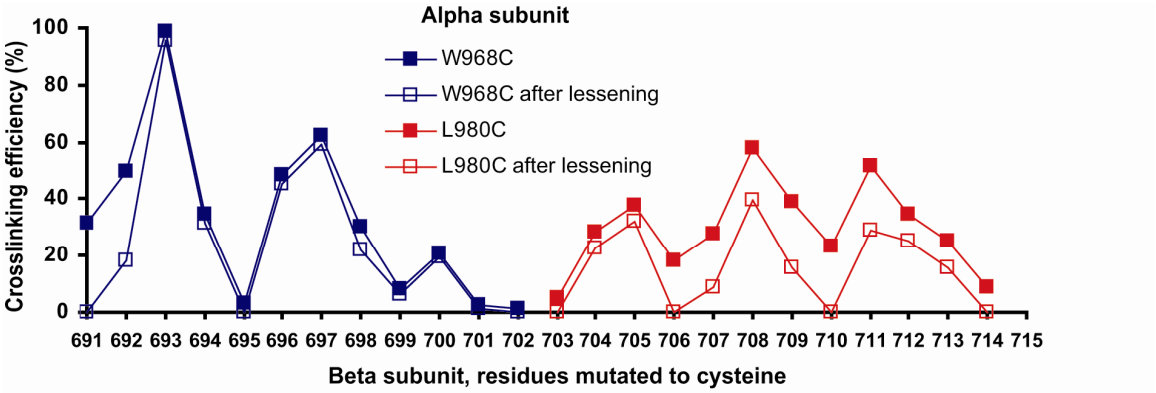
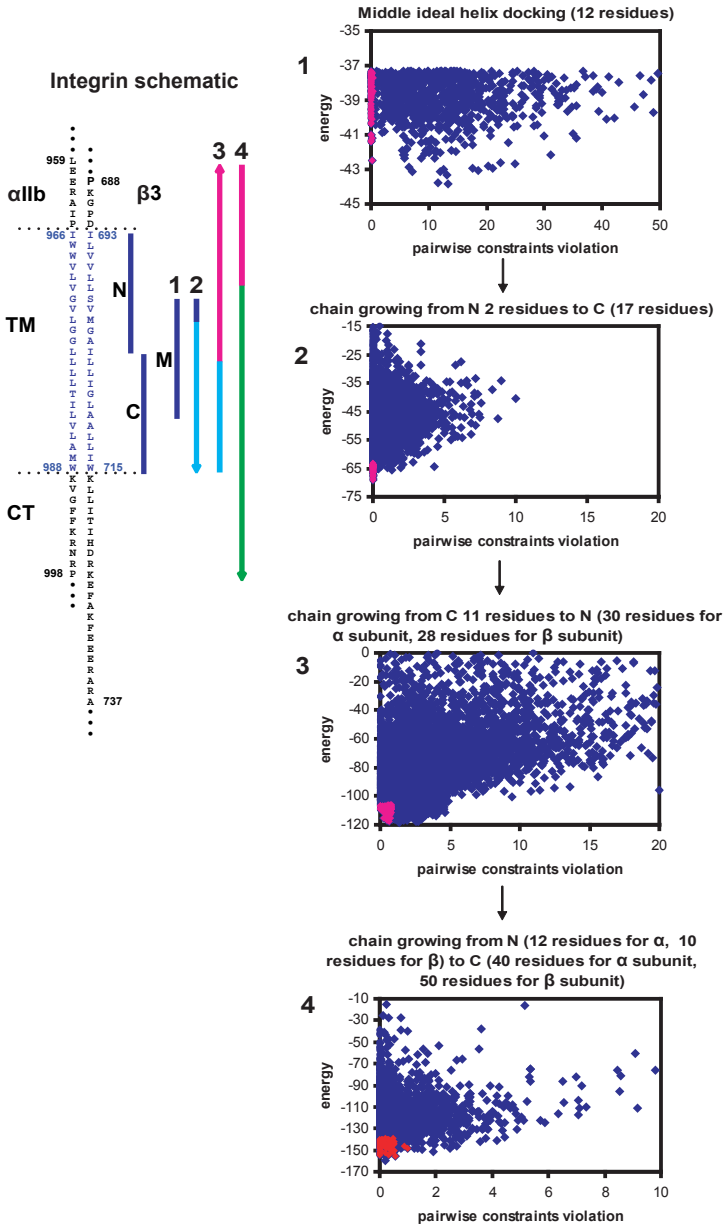
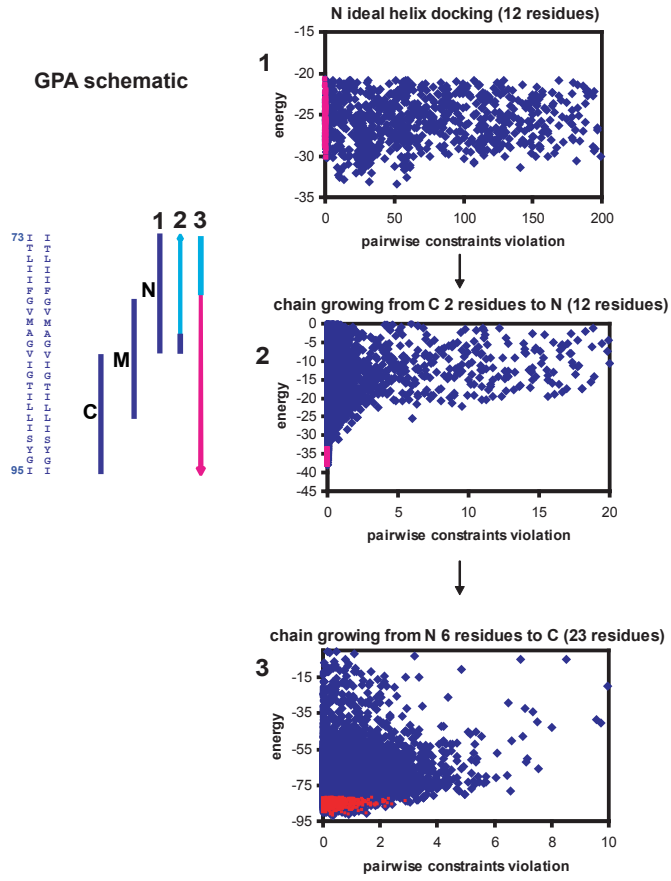

Alpha subunit
W968C
W968C after lessening
L980C
L980C after lessening
Crosslinking efficiency (%)
100
80
60
40
20
0
691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715
Beta subunit, residues mutated to cysteine

Figure S5

A



C



B

Select 10% lowest energy models (500) for clustering, rmsd cutoff: 2.0 Å (3.0 Å)

Cluster No.	Cluster size	Average energy	Average pairwise constraints violation
1	52 (224)	-145.82±4.61 (-144.26±4.11)	0.26±0.23 (0.26±0.27)
2	17 (79)	-144.93±3.82 (-143.64±3.51)	0.13±0.18 (0.27±0.29)
3	13 (23)	-142.74±3.99 (-143.68±3.68)	0.30±0.17 (0.44±0.56)
4	12 (18)	-143.83±2.78 (-144.34±3.93)	0.25±0.17 (0.25±0.21)
5	9 (17)	-142.53±2.75 (-144.46±3.21)	0.36±0.44 (0.31±0.24)

D

Select 10% lowest energy models (1000) for clustering, rmsd cutoff: 1.0 Å

Cluster No.	Cluster size	Average energy	Average pairwise constraints violation	rmsd of center model to NMR structure (Å)
1	652	-84.70±1.91	0.46±0.45	0.71
2	78	-84.66±1.73	0.28±0.47	0.99
3	48	-83.81±1.16	0.79±0.79	1.43
4	38	-84.37±1.95	0.65±0.65	1.55
5	34	-84.04±1.45	0.20±0.28	0.92

Figure S6

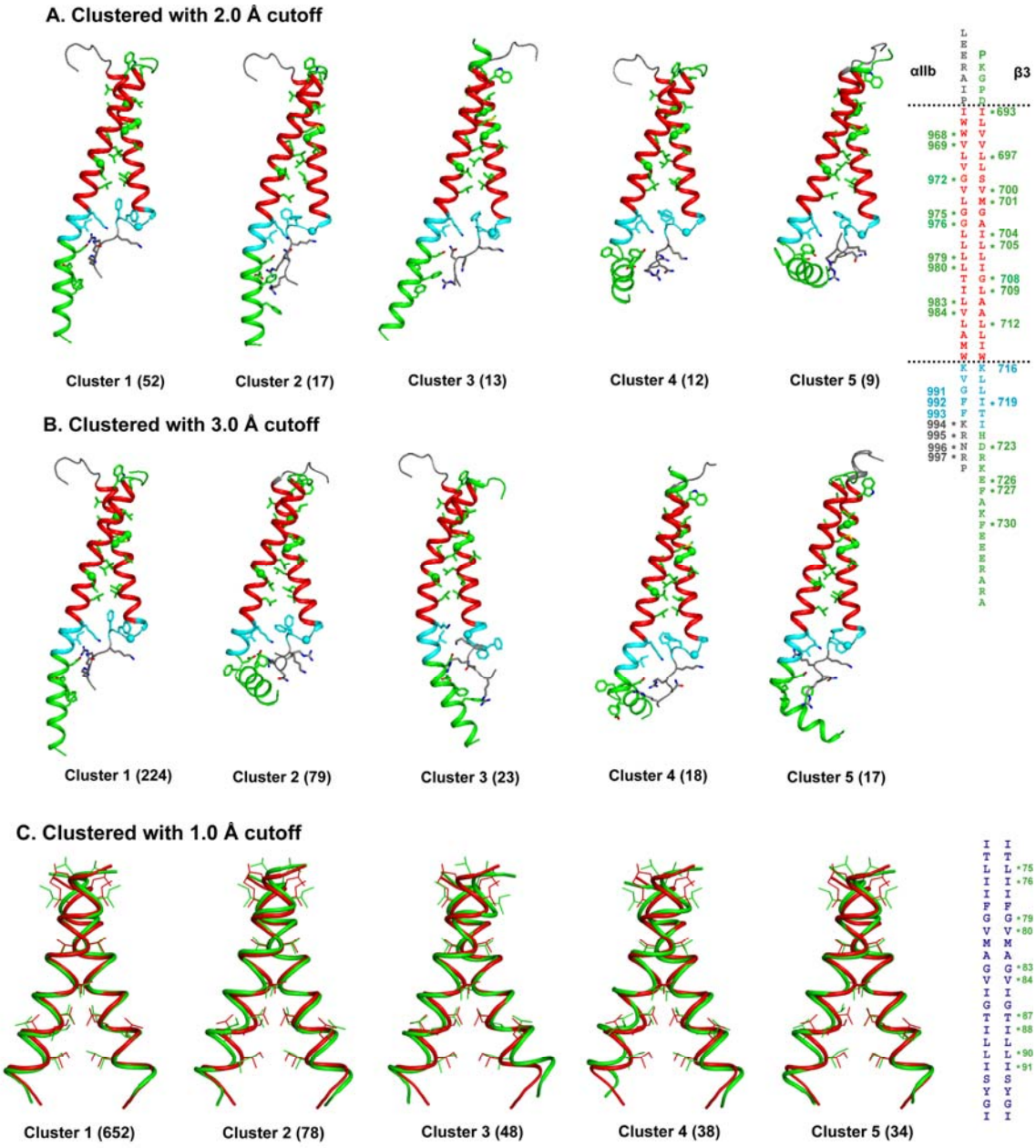

A. Clustered with 2.0 Å cutoff
αIIb
β3
Cluster 1 (52)
Cluster 2 (17)
Cluster 3 (13)
Cluster 4 (12)
Cluster 5 (9)
B. Clustered with 3.0 Å cutoff
Cluster 1 (224)
Cluster 2 (79)
Cluster 3 (23)
Cluster 4 (18)
Cluster 5 (17)
C. Clustered with 1.0 Å cutoff
Cluster 1 (652)
Cluster 2 (78)
Cluster 3 (48)
Cluster 4 (38)
Cluster 5 (34)

Figure S7

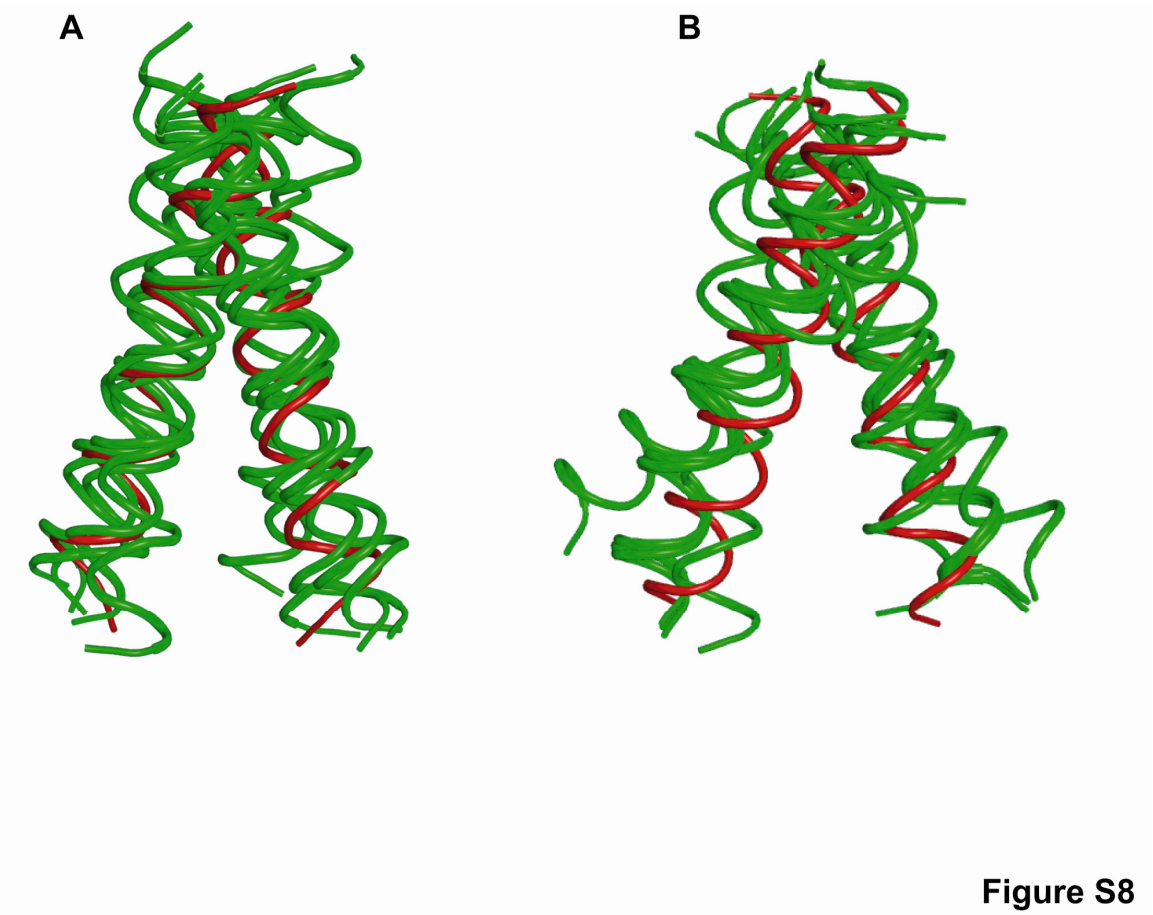


Figure S8

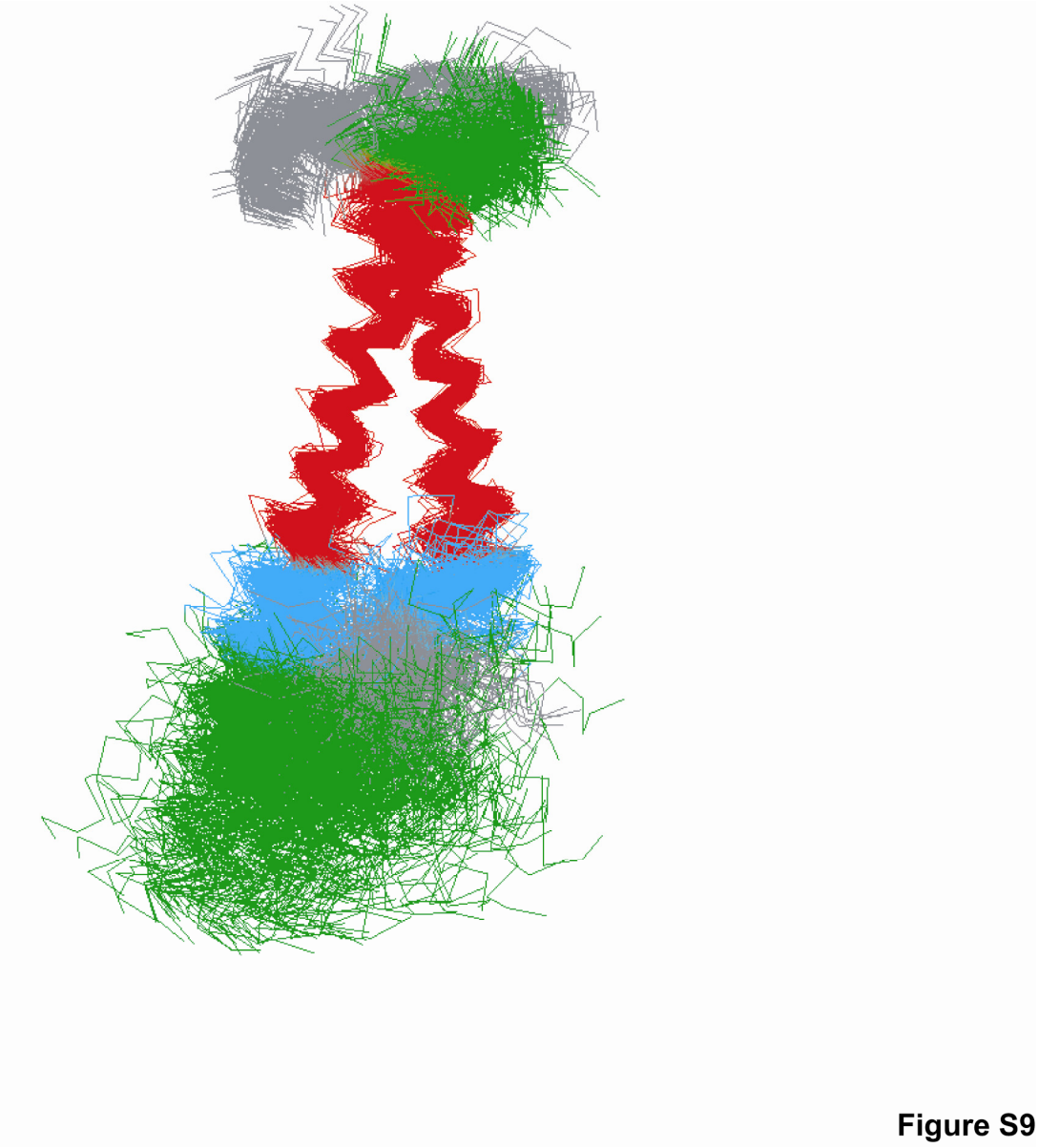


Figure S9

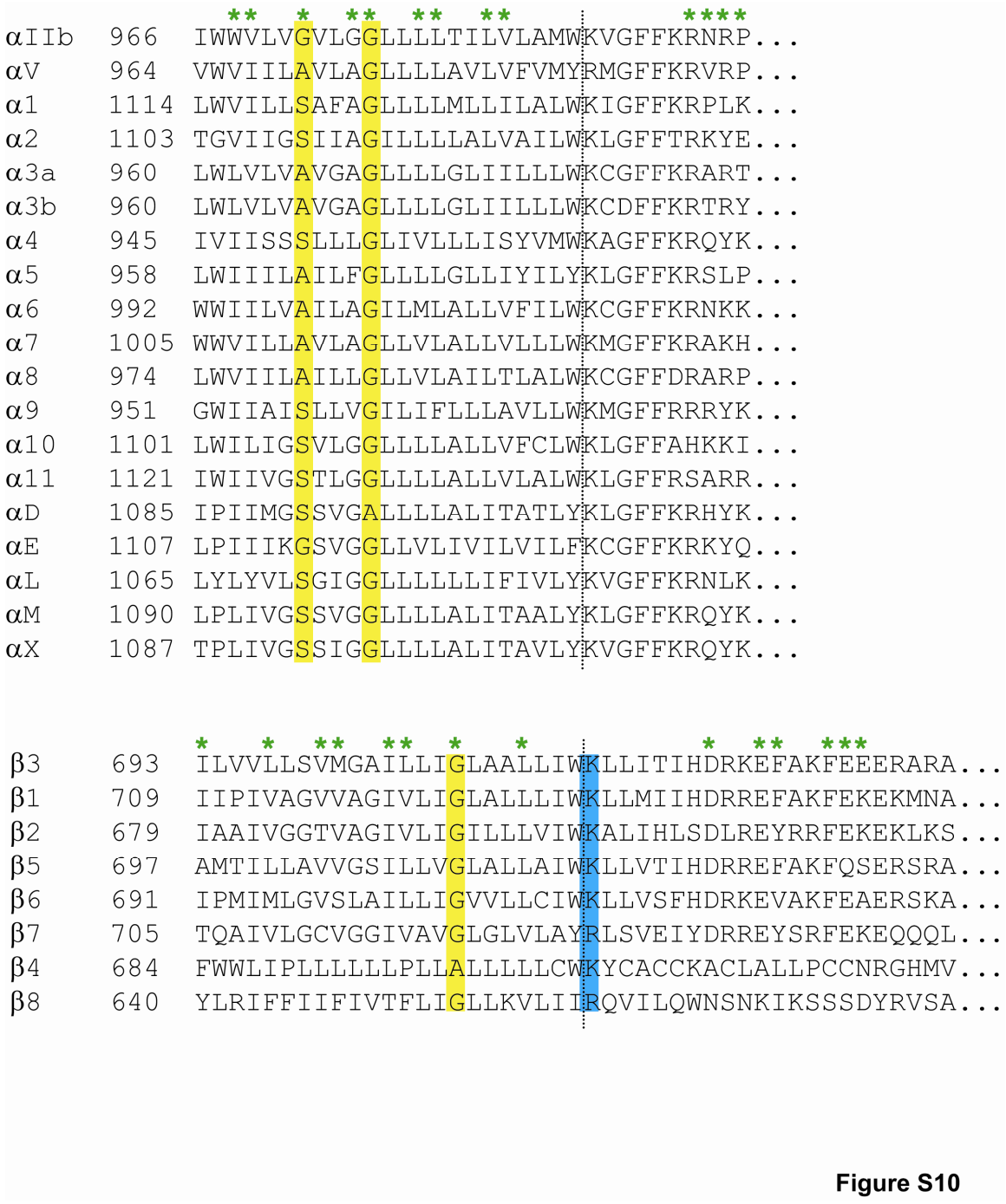

αIIb 966 IWWVLVGVLGGLLLLTILVLAMWKVGFFKRNRP...
αV 964 VWVIILAVLAGLLLLAVLVFVMYRMGFFKRVRP...
α1 1114 LWVILLSAFAGLLLLMLLILALWKIGFFKRPLK...
α2 1103 TGVIIGSIIAGILLLLALVAILWKLGFFTRKYE...
α3a 960 LWLVLVAVGAGLLLLGLIILLLWKCGFFKRART...
α3b 960 LWLVLVAVGAGLLLLGLIILLLWKCDFFKRTRY...
α4 945 IVIISSSLLLGLIVLLLISYVMWKAGFFKRQYK...
α5 958 LWIIILAILFGLLLLGLLIYILYKLGFFKRSLP...
α6 992 WWIILVAILAGILMLALLVFILWKCGFFKRNKK...
α7 1005 WWVILLAVLAGLLVLALLVLLLWKMGFFKRAKH...
α8 974 LWVIILAILLGLLVLAILTLALWKCGFFDRARP...
α9 951 GWIIAISLLVGILIFLLLAVLLWKMGFFRRRYK...
α10 1101 LWILIGSVLGGLLLLALLVFCLWKLGFFAHKKI...
α11 1121 IWIIVGSTLGGLLLLALLVLALWKLGFFRSARR...
αD 1085 IPIIMGSSVGALLLLALITATLYKLGFFKRHYK...
αE 1107 LPIIIKGSVGGLLVLIVILVILFKCGFFKRKYQ...
αL 1065 LYLYVLSGIGGLLLLLLIFIVLYKVGFFKRNLK...
αM 1090 LPLIVGSSVGGLLLLALITAALYKLGFFKRQYK...
αX 1087 TPLIVGSSIGGLLLLALITAVLYKVGFFKRQYK...
β3 693 ILVVLLSVMGAILLIGLAALLIWKLLITIHDRKEFAKFEEERARA...
β1 709 IIPIVAGVVAGIVLIGLALLLIWKLLMIIHDRREFAKFEKEKMNA...
β2 679 IAAIVGGTVAGIVLIGILLLVIWKALIHLSDLREYRRFEKEKLKS...
β5 697 AMTILLAVVGSILLVGLALLAIWKLLVTIHDRREFAKFQSERSRA...
β6 691 IPMIMLGVSLAILLIGVVLLCIWKLLVSFHDRKEVAKFEAERSKA...
β7 705 TQAIVLGCVGGIVAVGLGLVLAYRLSVEIYDRREYSRFEKEQQQL...
β4 684 FWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMV...
β8 640 YLRIFFIIFIVTFLIGLLKVLIIRQVILQWNSNKIKSSSDYRVSA...

Figure S10

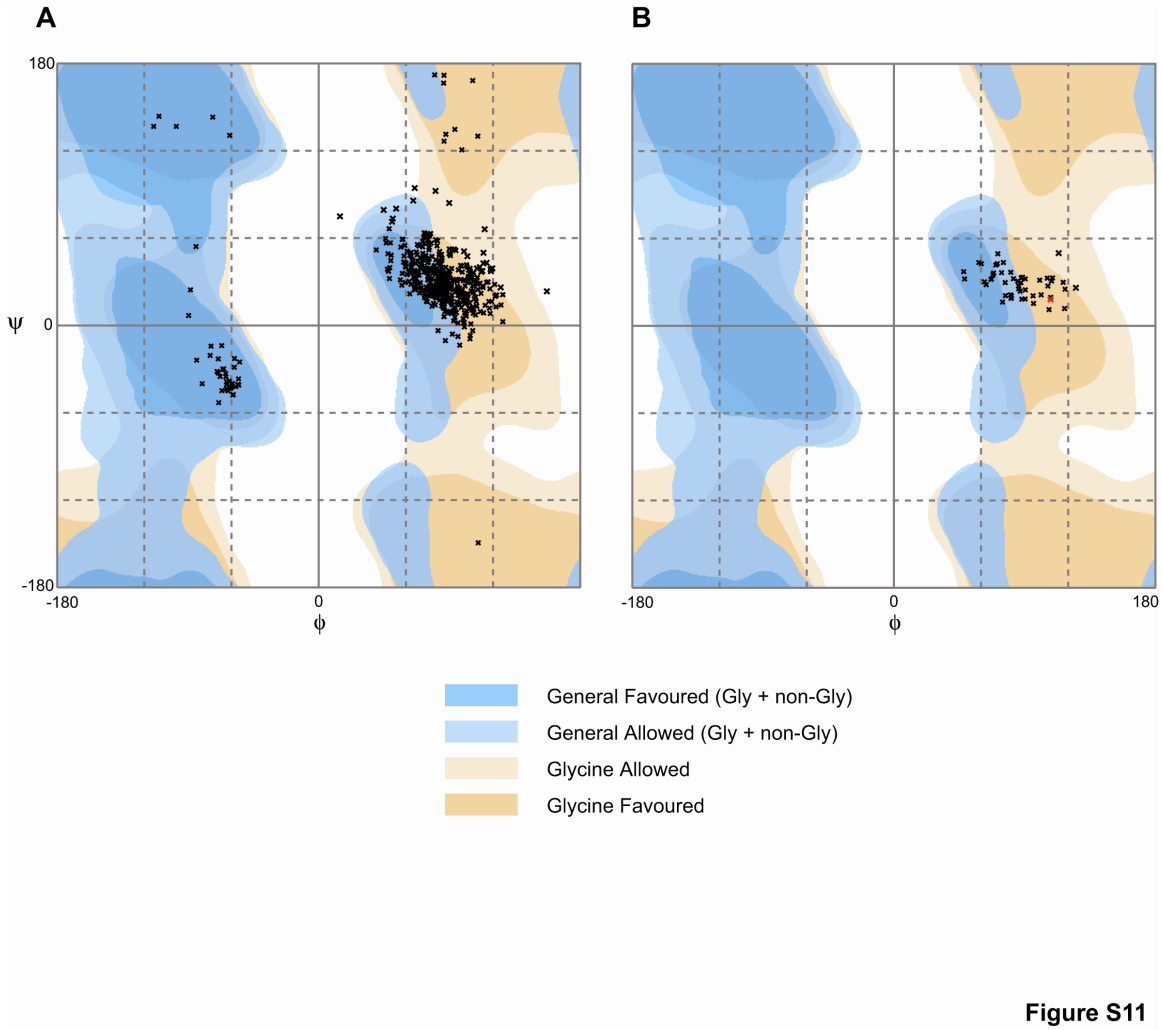

A
B
180
0
-180
ψ
-180
0
ϕ
-180
0
180
ϕ
General Favoured (Gly + non-Gly)
General Allowed (Gly + non-Gly)
Glycine Allowed
Glycine Favoured

Figure S11

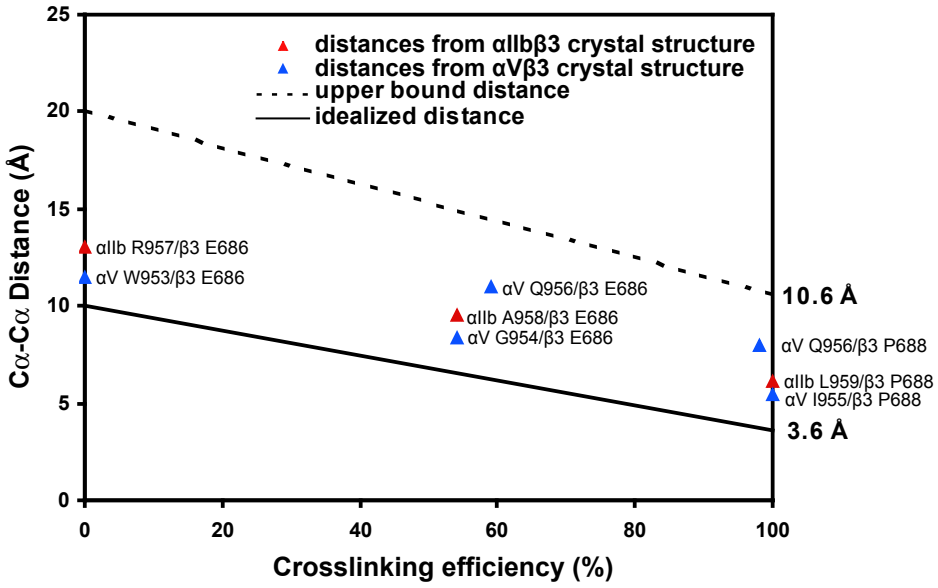

distances from αIIbβ3 crystal structure
distances from αVβ3 crystal structure
upper bound distance
idealized distance
Cα-Cα Distance (Å)
Crosslinking efficiency (%)
0
5
10
15
20
25
0
20
40
60
80
100
αIIb R957/β3 E686
αV W953/β3 E686
αV Q956/β3 E686
αIIb A958/β3 E686
αV G954/β3 E686
10.6 Å
αV Q956/β3 P688
αIIb L959/β3 P688
αV I955/β3 P688
3.6 Å

Figure S12

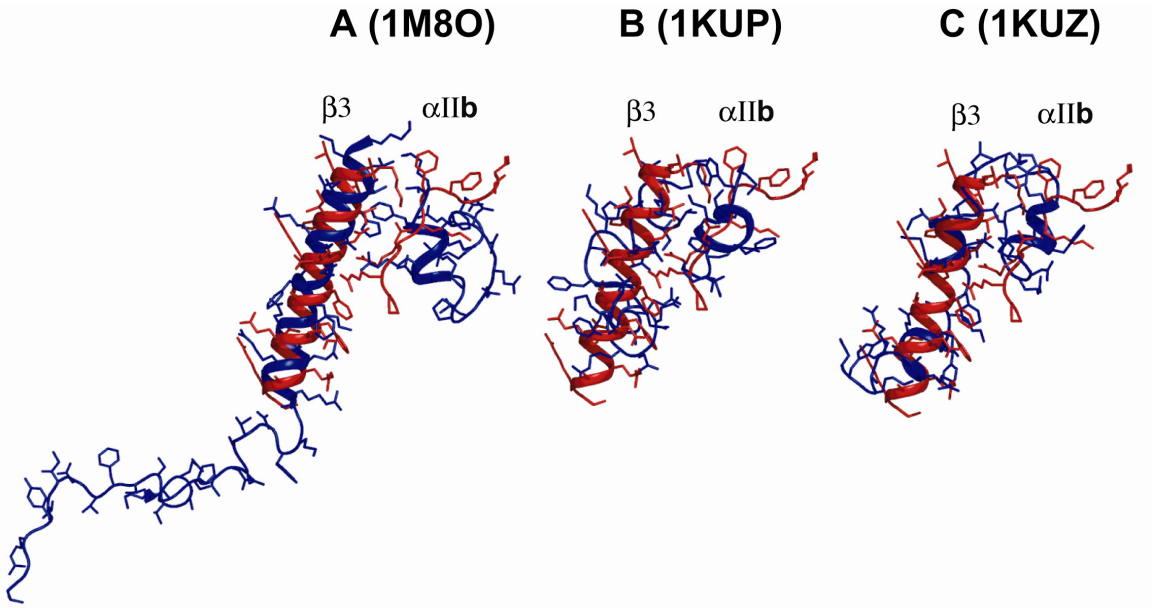


Figure S13