Biochemical and Biophysical Research Communications 256, 537–541 (1999)
Article ID bbrc.1999.0374, available online at http://www.idealibrary.com on IDEAL®

Combinatorial Chemistry Reveals a New Motif That Binds the Platelet Fibrinogen Receptor, gpIIbIIIa

David S. Thorpe,*,†,[1] Helen Yeoman,† A. W. Edith Chan,* Viktor Krchnak,*,[2] Michal Lebl,*,[3] and Stephen Felder†,[4]

*Department of Chemistry and †Department of Discovery Biology, Selectide Corporation, a subsidiary of Hoechst Marion Roussel, 1580 E. Hanley Boulevard, Tucson, Arizona 85737

Received February 9, 1999

Among cell adhesion molecules, the classic Arg-Gly-Asp (RGD) motif is the best studied. We used combinatorial chemical and affinity immunochemical methods to find a novel motif of unnatural peptide ligands for the fibrinogen receptor of platelets, gpIIbIIIa ($\alpha_{IIb}\beta_3$). The new D-amino acid motif, p(f/y)l, is unique among the ligands that bind the RGD pocket: It lacks the carboxylic acid group that is believed to coordinate with calcium in the MIDAS motif of the receptor. With an IC_{50} of 14 μM for the most potent compound, these linear p(f/y)l peptides had affinities similar to those of linear peptides containing RGD, and reversed sequences failed to compete with binding up to 1 mM. As the new motif was so different, molecular modeling was employed to suggest a model for molecular recognition. A reversed binding mechanism common for D-amino acid mimics of natural L-amino acid peptides offers an attractive hypothesis that suggests three points of contact similar to those made by the RGD-mimicking monoclonal antibody, OPG2. Interestingly, the model proposes that π-electrons in the new motif may substitute for the carboxylate group present in all other RGD-types of ligands. Although modeling linear peptides is subjective, the π-bonding model provides intriguing possibilities for medicinal chemistry after appropriate confirmatory studies.

Key Words: **combinatorial chemistry; integrin; gpIIbIIIa; $\alpha_{IIb}\beta_3$; MIDAS motif; π–cation bond; receptor; ligand; cell adhesion; peptide library.**

Combinatorial chemistry is a new tool for understanding pharmacological phenomena such as molecular recognition. At least seven integrins bind to the Arg-Gly-Asp (RGD) epitope (1–3). The guanidino group from arginine, and the carboxylic acid group from aspartate are the critical side chain features displayed by this motif, and the stereochemistry is strict only for the acidic function (4). Few protein motifs are better understood than the RGD ligand of cellular integrins. Despite extensive analoging efforts by medicinal chemists, there is little deviation from the pharmacophores in RGD. Combinatorial chemistry helps to explore the diversity of pharmacophores that bind a target molecule. Early combinatorial chemistry methods produced peptide libraries (reviewed in Ref. 5). In this report, the libraries take the form of thousands or millions of 100 μm polystyrene beads with a unique peptide displayed on each (6). The suitability of synthetic peptide libraries to explore the pharmacophoric space of the fibrinogen receptor was confirmed by screening libraries made with L-amino acids. Prior studies had found the expected RGD motif in libraries using solution phase assays of integrin binding (7), however, a higher throughput technique is on-bead staining (6). With this method, a soluble receptor is added to the beads displaying different peptides. Beads that bind the receptor are histochemically stained, recovered with manual pipettes, and the attached peptide is identified by Edman microsequencing. Peptide libraries using L or D amino acids were screened against the platelet fibrinogen receptor, gpIIbIIIa ($\alpha_{IIb}\beta_3$), a therapeutically important integrin. RGD and known dibasic analogs (8) were found in the L-amino acid libraries, and a new motif was discovered in the D-amino acid peptide library. The motif, pro (phe or tyr) leu, is pharmacophorically different from RGD, suggesting either unexplored interactions available within the fibrinogen-binding pocket of this model integrin, or a previously unexploited π-electron bonding mechanism.

[1] To whom correspondence should be addressed. E-mail: David.Thorpe@hmrag.com.

[2] Present address: SIDDCO, Inc., 9000 S. Rita Rd., Bldg. 40, Tucson, AZ, 85747.

[3] Present address: Spyder Instruments, Inc., P.O. Box 910485, San Diego, CA 92191.

[4] Present address: NeoTech, P.O. Box 64326, Tucson, AZ 85728.

Abbreviations used: AP, alkaline phosphatase; BCIP, 5-bromo-4-chloroindolyl-phosphate; GAM, goat antimouse antibody; HBS, Hepes buffered saline; mAb, monoclonal antibody; MIDAS, metal ion-dependent adhesion site; OPD, o-phenylenediamine dihydrochloride.

0006-291X/99 $30.00

EXPERIMENTAL PROCEDURES

Materials. Common reagents were from Aldrich (Milwaukee, WI), Fisher (Tustin, CA), Pierce (Rockford, IL) or Sigma (St. Louis, MO). Amino-Tentagel and RAM-Tentagel were from Rapp Polymere (Tubingen, Germany). Rink resin and Nα-Fmoc protected amino acids were from Advanced Chemtech (Louisville, KY) or Novabiochem (San Diego, CA). Several materials were supplied by Genentech, and these included soluble gpIIbIIIa, G4120, a potent, cyclic RCD peptide developed by Genentech (9). In addition, a monoclonal antibody, 4B12, also was prepared and supplied by Genentech (9). Goat antimouse (GAM)[1]-alkaline phosphatase was from American Qualex, San Clemente, CA. BCIP (5-bromo-4-chloroindolyl-phosphate) and OPD (*o*-phenylenediamine dihydrochloride) were from Pierce. The Hepes buffered saline (HBS) had the following composition: 0.01% gelatin, 150 mM NaCl, 20 mM Hepes, pH 7.5 and 5 mM $CaCl_2$.

Methods. Libraries were made with 19 L- or D-amino acids (single letter L-amino acid code, A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y) using the split-mix synthesis approach and probed by modified methods previously reported (6). The first L-amino acid library had a structure of YXXDV, bound to TentaGel-130 beads through a previously described linker (7) to the carboxyl termini of the randomized peptides. Two large generic 5-mer libraries were made that displayed L-amino acid peptides or D-amino acid peptides that were linked to TentaGel through the dipeptide linker sequence, KY (XXXXX-KY-TentaGel). About 2000 beads were tested in the first small library, and one million beads were screened from each of the generic 5 mer libraries. All free peptides were made as their acid amides using Rink resin or RAM-TentaGel. Their biological activities were confirmed in an ELISA that monitors the binding of soluble gpIIbIIIa to fibrinogen immobilized in 96-well plates as already described (9). Briefly, plates were coated with fibrinogen, blocked, and the binding of soluble gpIIbIIIa was monitored by ELISA reaction using the gpIIbIIIa-specific mAb 4B12 and GAM-horse radish peroxidase. The optical density at 492 nm was measured in duplicate wells after the reaction with OPD and hydrogen peroxide was terminated with 0.3 M sulfuric acid. Data were fitted by a nonlinear least squares method (10) and are representative of three experiments.

Beads were stained using 30 nM each of soluble gpIIbIIIa, mAb 4B12, and goat-anti-mouse immunoglobulin conjugated with alkaline phosphatase (GAM-AP) in HBS. The binding of gpIIbIIIa to beads was detected by the staining of beads after reaction of alkaline phosphatase with BCIP to generate an insoluble, blue colored product deposited onto the beads. The beads were spread out in petri dishes and examined with dissecting microscopes. Blue colored beads were collected with hand pipettes and stripped with 0.5 M NaOH, then water, then dimethyl formamide. After washing the beads with water and blocking with HBS, the specific binding population of beads was identified by probing again, but in the presence of 6 μM G4120, a cyclic RGD analog of high affinity (K_d = 2 nM) from Genentech. Colorless beads following alkaline phosphatase reaction represented the competed population. After a series of repeated staining and competition, isolated beads were subjected to Edman microsequencing. Based on these sequencing results, peptides were made by solid phase peptide synthesis. The molecular weights of these compounds were confirmed by MS, and the purity by HPLC was greater than 95%.

Molecular modeling was done after retrieval of the X-ray structural coordinates of the RGD mimicking mAb, OPG2, 1opg.pdb (11) from the Brookhaven protein database. The binding loop (amino acids 100-110) was taken from the structure and used as a template for modeling. The modeling package used was SYBYL 6.3, from Tripos, Inc. (St. Louis, MO). All of our peptides were first modeled from the X-ray structure and then subjected to energy minimization using the Tripos force field with the adopted-basis Newton-Raphson method. Finally, all of the peptides were overlapped or fitted with the OPG2 binding loop for visual comparison.

TABLE 1

RGD Family of Ligands Found by Screening L-Amino Acid Peptide Libraries

Structure	IC_{50} (μM)
YRGDV-NH_2	1
YRRDV-NH_2	28
YKRDV-NH_2	79
YFRDV-NH_2	130
YWRDV-NH_2	250
RGDFVKY-NH_2	6
RGDFPKY-NH_2	6
RGDWAKY-NH_2	6
RGDFWKY-NH_2	25
FKKFLKY-NH_2	320

RESULTS AND DISCUSSION

This report describes the first successful project on a therapeutic target using what has become known as the "Selectide Process" (6). To validate the strategy of on-bead staining, a library of limited complexity was screened. The library had the structure YXXDV using 19 L-amino acids in the randomized positions denoted by "X." Sequential staining and competition steps led to the isolation of beads displaying the RGD motif (Table 1). The technique of on-bead staining proved to be very sensitive and effective. Not only was RGD found, but spacing variants (e.g., YWRDV) of 100- to 200-fold lower affinities were identified. The screen also found dibasic peptides of intermediate affinity. These results were as expected, given what is known about integrins and their ligands. Screening of a random pentapeptide library also found the expected RGD motif (Table 1), however, all of the RGD peptides had arginine at the amino termini. Perhaps the RGD in more internal sequences could not bind the receptor. The RGD binding site is apparently in a cleft at least 11 Angstroms deep (12), and this would correspond to the length of a tetrapeptide linking RGD to the TentaGel bead.

Once validated, the on-bead screening technique was applied to a library of unnatural amino acids. Repeated staining and competition of a pentapeptide library of 18 D-amino acids and Glycine identified a novel motif, pro (phe or tyr) leu and abbreviated as p(f/y)l. Inhibition of the binding of soluble gpIIbIIIa to immobilized fibrinogen is shown in Fig. 1. The most potent compound had an IC_{50} of 14 μM, and the interaction was specific because peptides of the same composition, but reversed sequence, failed to inhibit binding at concentrations up to 1 mM.

An advantage of combinatorial chemistry is the rapid collection of structural and functional information. The structure-activity relationship of the D-amino acid peptides is summarized in Table 2 showing a 23-fold range of affinities for the new motif. In all

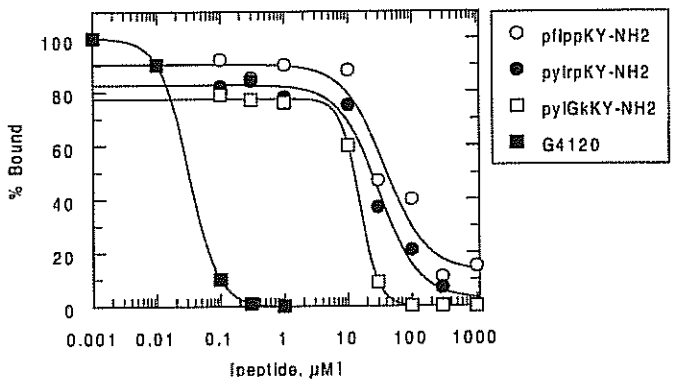

FIG. 1. Inhibition of the binding of soluble gpIIbIIIa to immobilized fibrinogen by soluble peptides with the new motif. Reversed sequence control peptides did not inhibit binding at concentrations up to 1 mM.

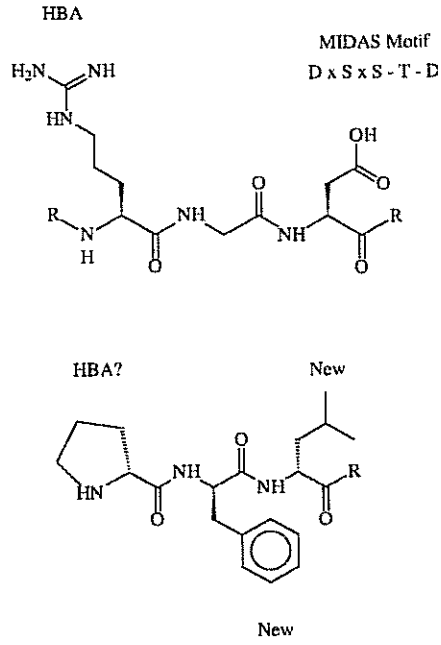

FIG. 2. Comparison of the two motifs, and a proposed pharmacophoric map for gpIIbIIIa. RGD is shown at the *top*, pfl is shown at the *bottom*. The chemical structures for the ligand motifs are shown along with their proposed contacts to the integrin. *HBA* denotes hydrogen bond acceptor-rich pocket or anionic surface; *MIDAS* represents a calcium chelating segment of the integrin, beginning with Asp119 (13, 16, 26), that is proposed to bind the acidic group of RGD ligands through a divalent salt bridge; *New* marks a pocket(s) for a hydrophobic or aromatic contact for the new ligand motif; *R* stands for additional amino acids.

cases, the amino terminus was D-proline. D-Phenylalanine and D-tyrosine at the second position differ only by a hydroxyl group. Only occasional, homologous substitutions occur in the third position of the p(f/y)l motif. Importantly, no acidic groups were present, and all peptides were synthesized as acid amides. Moreover, these peptides were attached to the beads through their carboxyl termini during screening. Thus, the acidic group in all known RGD-mimicking ligands of gpIIbIIIa, natural and synthetic, is absent from this family of ligands. The L-amino acid linker segment, KY, contributes to affinity, however, there are two cases where the contribution is two-fold or less (Table 2). There is no trend of individual amino acids, or classes of amino acids, in the fourth or fifth randomized positions of the D-peptide library. This emphasizes that the critical component of the new motif is the tripeptide sequence p(f/y)l.

TABLE 2

Structure–Activity Relationship of Peptides with the New Motif Found by Screening a D-Amino Acid Peptide Library

Sequence	IC_{50} (μM)
1. pylGkKY-NH_2	14 {100}[a]
2. pylrpKY-NH_2	35 {125}
3. pylqqKY-NH_2	140
4. pylqkKY-NH_2	56
5. pylwtKY-NH_2	320
6. pylssKY-NH_2	100
7. pymGaKY-NH_2	79 {125}
8. pflmnKY-NH_2	50 {110}
9. pflppKY-NH_2	45 {315}
10. pfmnvKY-NH_2	250 {320}
11. pfiwGKY-NH_2	270
12. pflvpKY-NH_2	220
13. pflplKY-NH_2	110

[a] IC_{50} of peptides without the KY segment from the linker are shown within braces. A library of structure XXXXX-KY-TentaGel was made using 18 D-amino acids and glycine and probed as described under Methods, but with the addition of 0.75 M NaCl to HBS. In the single-letter amino acid code, the D-isomer is indicated by the lowercase letter, e.g., p represents D-proline.

No analog made to date has used pharmacophores so different from those of RGD. The novelty of the unnatural peptide motif is striking, and this caused us to consider hypothetical mechanisms of binding. We propose two models for binding besides allosteric mechanisms. Figure 2 shows the structural features of the old and new motifs and the relationships they may have to receptor contacts. In ligands containing RGD (Fig. 2, top), the arginyl side chain binds a hydrogen bond acceptor(s), or an anionic surface. The carboxylate group of aspartate coordinates with a calcium ion associated with a D x S x S - T - D divalent cation binding structure, or to a pair of acidic residues of the integrin (13–16). In the proposed model (Fig. 2, bottom), the D-prolyl residue binds the pocket used by arginine in RGD. This would account for why the cyclic RGD peptide G4120 competes with the binding of soluble gpIIbIIIa to the beads displaying p(f/y)l sequences. Also, medicinal chemistry shows that secondary amines can substitute for the guanidino group of RGD (17). The

TABLE 3

Reversed Binding Mode Alignment Used in Molecular Modeling

OPG2 (103–107)	PFYR--YDGG
Reverse Seq 1	YKkGlyp
Reverse Seq 9	YKpplfp

Note. The OPG2 sequence is shown amino to carboxyl from left to right, and the p(f/y)l peptide sequences are shown with their amino to carboxyl from right to left and numbered according to Table 2. Amino acids are indicated by the single-letter codes; lowercase letters represent D-amino acids. Bold letters are residues that are proposed to make important contacts with gpIIbIIIa.

aromatic and leucyl residues of the new motif are proposed to bind to a new contact site(s) in or adjacent to the fibrinogen-binding pocket. It is unclear whether this new site(s) is necessary for binding to fibrinogen, however, at least one of the receptor contacts that binds the new motif must overlap with those used by both G4120 and fibrinogen.

Additional support for this model comes from studies in medicinal chemistry and with RGD-mimicking antibodies (11, 18, 19). These monoclonal antibodies have the essential sequences **RYD** or **RGYFD** in the antigen-combining site. Crystallographic studies of the Fab from OPG2, which has **RYD**, show the tyrosyl residue pointing in a direction opposite to the side chains of arginine and aspartate (11). In addition, the potential complementary surface modeled around the **RYD** of OPG2 predicts a cleft opposite to the pockets for Arg and Asp that is occupied by Tyr. The RGD binding domain of gpIIbIIIa also is flanked in the primary structure by hydrophobic and aromatic residues (20). Thus, the aromatic residues and leucine of p(f/y)l may bind a hydrophobic cleft adjacent to contacts used by RGD. Panning of recombinant phage display libraries have shown a preference for aromatic residues close to RGD (8, 21). Finally, medicinal chemistry shows that hydrophobic and aromatic tethers used to make constrained cyclic analogs of RGD peptides increase their affinity (17). Such moieties are opposite to the guanidino and acidic functionalities, and these tethers may be binding the aromatic or hydrophobic contact(s) that may bind the p(f/y)l ligands.

An alternative model of binding for the new motif is based upon the recent elucidation of π-electron binding mechanisms in biological microenvironments. The cation–π interaction is now well accepted as an important noncovalent binding force in biological systems (22–24). As the aspartyl groups of RGD and RYD are believed to coordinate with a calcium ion, we have modeled D-Tyr or D-Phe with the aspartyl group of RYD in the loop of OPG2. This second model also invokes the reverse binding mode often seen with D-amino acid analogs of L-amino acid ligands (24, and references therein). Table 3 shows the alignment that is involved in this case of reverse binding. Reverse binding is where the D-amino acid analog binds the receptor with

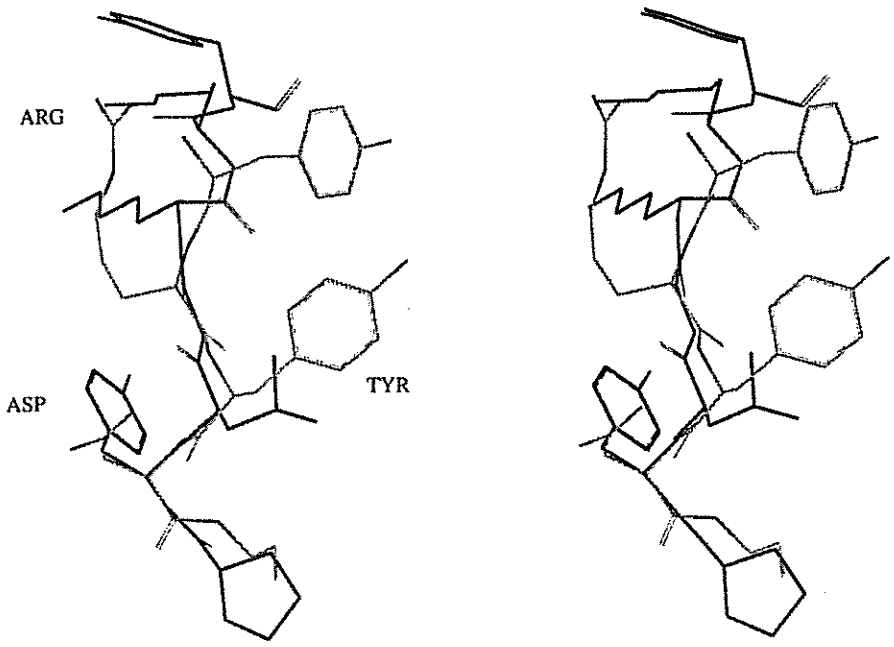

FIG. 3. A relaxed stereodiagram illustrating the three dimensional overlap of peptide pylGkKY-NH_2 with the RYD loop of OPG2. The side chain functionalities of D-Tyr and Lys of the peptide are overlapping well with the side chains of Asp105 and Arg103, respectively. In addition, the side chain of D-Leu overlaps with that of Tyr104. The OPG2 loop is shown as a gray atomic wire tracing proceeding from the N-terminal residue to the carboxyl terminal residue top to bottom, and the pylGkKY-NH_2 is shown in black with its main chain atoms antiparallel to OPG2.

an antiparallel orientation compared to the main chain atoms of a natural L-amino acid ligand. One gap is introduced into the alignment, but the result of molecular modeling illustrated in Fig. 3 gives added strength to this hypothesis. Templated modeling of the D-amino acid peptides led to overlapping D-Tyr or D-Phe with the aspartyl residue of the OPG2 loop. This was considered plausible given cation–π bonding schemes. In addition, the D-Leu side chain can occupy a similar space as that of Tyr104 of the OPG2 mAb, thereby picking up the required hydrophobic contact. The modest structure-activity relationship shown in Table 2 also shows a tendency to loose potency when the carboxyl terminal KY residues were omitted. In addition to maintaining the hydrophobic and cation–π bonds, the L-Lys residue of the p(f/y)l-KY peptides can be positioned to cover the same region of space as the arginyl side chain of OPG2, thereby mimicking the function of the arginyl side chain. Of the two p(f/y)l peptides lacking KY but retaining modest affinity, one of them has D-Asn in the position occupied by arginyl in RYD, and this could make useful hydrogen bonds. Thus, several regions of overlap proposed by this second model explain why the small RGD analog G4120 competed the binding of the receptor to the beads displaying p(f/y)l peptides and why these peptides inhibited the binding of gpIIbIIIa to immobilized fibrinogen. Finally, because of its presence in all of the D-amino acid ligands, the amino terminal proline is likely to make an additional contact with the receptor of unknown character. Obviously, confirmatory studies are needed to validate either of the models proposed in this report.

The discovery of this new motif shows that the pharmacophoric interactions around RGD are more complex than widely believed and provides new ligand probes for studying integrins. A tremendous effort in medicinal chemistry has gone into understanding the RGD system, yet there is still a lot to learn. As illustrated by the model integrin gpIIbIIIa, combinatorial chemistry is likely to play a greater role in the future understanding of molecular recognition.

ACKNOWLEDGMENTS

The authors thank John Burnier of Genentech (South San Francisco) for providing gpIIbIIIa, 4B12, and G4120 and our colleagues at Selectide—Aleksandra Weichsel, Ron Ferguson, Martha Ackerman-Berrier, Bill Wire, Ron Jung, Christine Cuzzocrea, Dennis Bonilla, Brian Byers, Farid Abdul-Latif, Zuzana Leblova, Kolia Sepetov, and Olga Issakova—for their expert and dedicated work on this project.

REFERENCES

1. Ruoslahti, E. J. (1991) *Clin. Invest.* **87,** 1–5.
2. Busk, M., Pytela, R., and Sheppard, D. (1992) *J. Biol. Chem.* **267,** 5790–5796.
3. Nishimura, S. L., and Pytela, R. (1993) *Mol. Biol. Cell* **4**(Suppl.), 285 (Abstr.).
4. Pierschbacher, M. D., and Ruoslahti, E. (1987) *J. Biol. Chem.* **262,** 17294–17298.
5. Furka, A. (1995) *Drug Dev. Res.* **36,** 1–12.
6. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., and Knapp, R. J. (1991) *Nature* **358,** 82–84.
7. Salmon, S. E., Lam, K. S., Lebl, M., Kandola, A., Khatiri, P. S., Wade, S., Patek, M., Kocis, P., Krchnak, V., Thorpe, D., and Felder, S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11708–11712.
8. Koivunen, E., Wang, B., and Ruoslahti, E. (1994) *J. Cell Biol.* **124,** 373–380.
9. Barker, P. L., Bullens, S., Bunting, S., Burdick, D. J., Chan, K. S., Deisher, T., Eigenbrot, C., Gadek, T. R., Gantzos, R., Lipari, M. T., Muir, C. D., Napier, M. A., Pitti, R. M., Padua, A., Quan, C., Stanley, M., Struble, M., Tom, J. Y. K., and Burnier, J. P. (1992) *J. Med. Chem.* **35,** 2040–2048.
10. Leatherbarrow, R. J. (1992) GraFit Version 3.0, Erithacus Software Ltd., Staines, UK.
11. Kodandapani, R., Veerapandian, B., Kunicki, T. J., and Ely, K. R. (1995) *J. Biol. Chem.* **270,** 2268–2273.
12. Beer, J. H., Springer, K. T., and Coller, B. S. (1992) *Blood* **79,** 117–128.
13. Bajt, M. L., and Loftus, J. C. (1994) *J. Biol. Chem.* **269,** 20913–20919.
14. Pasqualini, R., Koivunen, E., and Ruoslahti, E. (1995) *J. Cell Biol.* **130,** 1189–1196.
15. Kobayashi, H., Kitabayashi, K., Matsumoto, K., and Hirokawa, H. (1991) *Virology* **185,** 901–903.
16. Tozer, E. C., Liddington, R. C., Sutcliffe, M. J., Smeeton, A. H., and Loftus, J. C. (1996) *J. Biol. Chem.* **271,** 21978–21984.
17. Duggan, M. E., Naylor-Olsen, A. M., Perkins, J. J., Anderson, P. S., Chang, C. T.-C., Cook, J. J., Gould, R. J., Ihle, N. C., Hartman, G. D., Lynch, J. J., Lynch, R. J., Manno, P. D., Schaffer, L. W., and Smith, R. L. (1995) *J. Med. Chem.* **38,** 3332–3341.
18. Tomiyama, Y., Brojer, E., Rugger, Z. M., Shattil, S. J., Smiltneck, J., Gorski, J., Kumar, A., Kieber-Emmons, T., and Kunicki, T. J. (1992) *J. Biol. Chem.* **267,** 18085–18092.
19. Jarrin, A., Andrieux, A., Chapel, A., Buchou, T., and Marguerie, G. (1994) *FEBS Lett.* **354,** 169–172.
20. Loftus, J. C., Smith, J. W., and Ginsberg, M. H. (1994) *J. Biol. Chem.* **269,** 25235–25238.
21. O'Neil, K. T., Hoess, R. H., Jackson, S. A., and Ramachandran, N. S. (1992) *Proteins* **14,** 509–515.
22. Dougherty, D. A. (1996) *Science* **271,** 163–168.
23. Mecozzi, S., West, A. P., Jr., and Dougherty, D. A. (1996) *J. Am. Chem. Soc.* **118,** 2307–2308.
24. Mecozzi, S., West, A. P., Jr., and Dougherty, D. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10566–10571.
25. Wermuth, J., Goodman, S. L., Jonczyk, A., and Kessler, H. (1997) *JACS* **119,** 1328–1335 (and references therein).
26. Lee, J-O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995) *Cell* **80,** 631–638.