From Peptide to Non-Peptide. 1. The Elucidation of a Bioactive Conformation of the Arginine-Glycine-Aspartic Acid Recognition Sequence

Robert S. McDowell,*,‖ Thomas R. Gadek,*,‖ Peter L. Barker,†,‖ Daniel J. Burdick,‖ Kathyrn S. Chan,‖ Clifford L. Quan,‖ Nicholas Skelton,⊥ Martin Struble,‖ Eugene D. Thorsett,‡,‖ Maureen Tischler,‖ Jeffrey Y. K. Tom,‖ Thomas R. Webb,§,‖ and John P. Burnier‖

Contribution from the Department of Bioorganic Chemistry and Department of Protein Engineering, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080

Received October 15, 1993®

Abstract: An ensemble molecular dynamics method is used to map consensus conformations of the arginine-glycine-aspartic acid (RGD) sequence which are accessible to a set of potent, structurally diverse inhibitors of fibrinogen-glycoprotein IIbIIIa association. This procedure identifies a dominant low-energy RGD conformation that is consistent with the previously determined solution structure of a highly potent RGD-containing peptide. Enforcing an alternate, higher energy conformation is shown to eliminate activity. These observations strongly suggest that the consensus conformation identified is responsible for inhibiting fibrinogen-glycoprotein IIbIIIa binding, thus providing a structural rationale for the *de novo* design of potent non-peptidal inhibitors of platelet aggregation.

Introduction

The use of structure-based design to transform an amino acid epitope into a non-peptide lead compound remains one of the fundamental challenges of modern medicinal chemistry. In principle, combining the functional information obtained by protein mutagenesis or peptide synthesis with the structural information obtained using NMR, crystallography, and molecular modeling should enable the *de novo* design of biologically active, non-peptidal compounds with improved bioavailability and pharmacodynamics. In practice, the successful implementation of this paradigm has seldom been realized.[1-3] This approach requires as a first step the elucidation of a bioactive conformation of potent peptidal ligands, which is described herein. In a subsequent paper we describe the successful application of this information to the design of potent, non-peptidal inhibitors of platelet aggregation.

Because it is often impossible to directly determine the bound-state conformations of biologically active peptides, conformational analysis procedures are frequently used to map energetically accessible conformations of flexible molecules interacting with a common site. Given a collection of such molecules, the primary goal of these procedures is to identify consensus conformations that present a common spatial arrangement of shared binding determinants. Systematic conformational searching[4,5] provides a rigorous sampling of torsion space but is not feasible for systems with more than 6–7 rotatable bonds. Other methods such as distance geometry[6,7] and Monte Carlo sampling,[8,9] which are often coupled with energy minimization or molecular dynamics, provide alternative means of sampling the conformational space of more complex molecules. Unfortunately, all of these approaches generate voluminous data for each individual molecule, and the process of identifying commonalities between dissimilar molecules based on this data is often difficult, if not impossible.

The application of distance geometry to an ensemble of molecules represents an intriguing alternative to the methods described above.[10] In this approach, only conformations that superimpose related fragments from a collection of molecules are generated by the embedding procedure. As with any distance geometry procedure, the resulting structures usually require significant refinement before they can be considered energetically reasonable. Depending on the degree of strain present in the embedded structures, refinement by conventional minimization or dynamics methods can significantly disrupt the alignment.

We have implemented an ensemble molecular dynamics procedure that can be used either alone or as a refinement method in conjunction with ensemble distance geometry. In this procedure, a dynamics simulation of a collection of molecules is conducted in which corresponding groups of interest ("binding determinants") common to all molecules are tethered together by an additional constraining function added to the force field; nonbonded interactions between molecules are ignored. This procedure therefore samples the conformational manifold of a collection of molecules while constraining corresponding groups within those molecules to occupy similar locations in space. The resulting consensus conformations can be quenched by minimiza-

* To whom correspondence should be addressed.
† Current address: Affymax Research Institute, 4001 Miranda Ave., Palo Alto, CA 94304.
‡ Current address: Athena Neurosciences, 800F Gateway Blvd., South San Francisco, CA 94080.
§ Current address: Corvas, Inc., 3030 Science Park Drive, San Diego, CA 92121.
⊥ Department of Protein Engineering.
‖ Department of Bioorganic Chemistry.
® Abstract published in *Advance ACS Abstracts*, May 1, 1994.

(1) Hirschmann, R. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 127.
(2) Hirschmann, R.; Nicolaou, K. C.; Pietranico, S.; Salvino, J.; Leany, E. M.; Sprengeler, P. A.; Furst, G.; Smith, A. B. III; Strader, C. D.; Cascieri, M. A.; Candelore, M. R.; Donaldson, C.; Vale, W.; Maechler, L. *J. Am. Chem. Soc.* **1992**, *114*, 9217.
(3) Smith, A. B. III; Keenan, T. P.; Holcomb, R. C.; Sprengeler, P. A.; Guzman, M. C.; Wood, J. L.; Carroll, P. J.; Hirschmann, R. *J. Am. Chem. Soc.* **1992**, *114*, 10672.
(4) Smith, G. M.; Veber, D. F. *Biochem. Biophys. Res. Commun.* **1986**, *134*, 907.
(5) Dammkoehler, R. A.; Karasek, S. F.; Shands, E. F. B.; Marshall, G. R. *J. Comput.-Aided Mol. Design* **1989**, *3*, 3.
(6) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*; Research Studies Press: New York, 1988.
(7) Peishoff, C. E.; Dixon, J. S.; Kopple, K. D. *Biopolymers* **1990**, *30*, 45.
(8) Noguti, T.; Go, N. *Biopolymers* **1985**, *24*, 527.
(9) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
(10) Sheridan, R. P.; Nilakantan, R.; Dixon, J. S.; Venkataraghavan, R. *J. Med. Chem.* **1986**, *29*, 899.

0002-7863/94/1516-5069$04.50/0

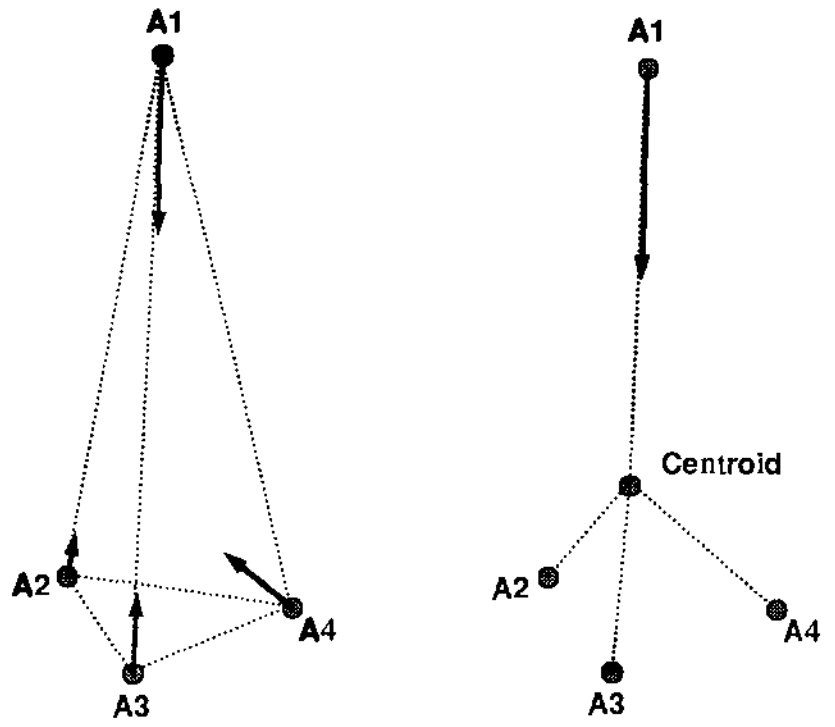


Figure 1. A comparison of the net gradients produced by a pairwise tethering approach (left) and a centroid-based approach (right). A1–A4 represent corresponding groups from four molecules that are to be aligned. The tethers are shown as dotted lines; the net gradients by solid arrows. In this example, groups 2–4 are within the desired distance tolerance (r^0). Centroid tethering functions similar to a simplex, applying the majority of the force to the group farthest from the mean location of the constituent groups. By contrast, pairwise tethering causes the net force to be distributed more equally on all constituent groups.

tion, or sampled over a long trajectory to indicate the degree of "rigidity" of a particular alignment.

A quadratic square-well potential of the following form is used as a tethering function:

$$E(r_{ij}) = 0.0 \qquad \text{if } r_{ij} \leq r^0_{ij}$$

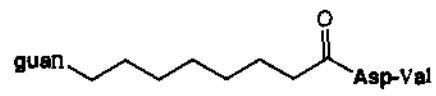
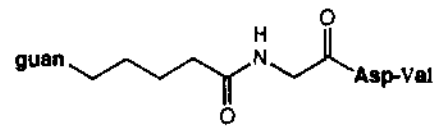
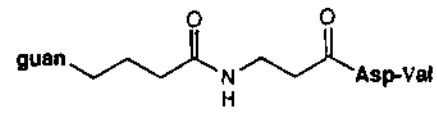
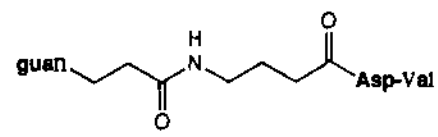
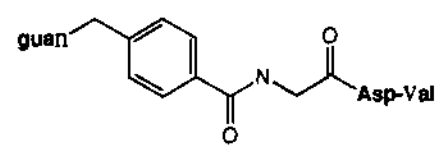
$$E(r_{ij}) = 0.5k_{ij}(r_{ij} - r^0_{ij})^2 \qquad \text{if } r_{ij} > r^0_{ij}$$

The parameter r_{ij} represents the distance between a group on molecule i and the centroid j calculated as the mean location of the corresponding groups from the ensemble of molecules; r^0_{ij} is a hypothetical maximum distance that is allowed. Centroid positions are recalculated after each step of dynamics based on the instantaneous coordinates of the constituent atoms. Although pairwise constraints between corresponding atoms in the ensemble could alternatively be applied, the two tethering approaches yield different net forces on the constituent atoms, as illustrated in Figure 1. Centroid-based tethering offers distinct advantages: for an ensemble consisting of N molecules, the centroid method requires N tethers per group instead of the $N(N-1)/2$ tethers per group required by the pairwise method. It is also possible to assess the ease with which a trial compound can adopt conformations that are accessible by a collection of other known molecules by independently adjusting the relative degree to which a constituent molecule is used to calculate centroid positions and the stringency with which it is tethered to those positions.

The magnitude of the force constant k_{ij} determines the degree to which the tethering potential contributes to the overall forcefield; setting k_{ij} equal to zero results in an unconstrained simulation. Nonaligned structures can be "randomized" using large values of r^0_{ij} and small values of k_{ij}; these values can be decreased and increased, respectively, during a simulation to induce alignment. Structures previously aligned using ensemble distance geometry can be refined starting with intermediate values of r^0_{ij} and k_{ij}, gradually "tightening" the constraints once the ensemble energy has equilibrated.

This method was developed as part of a strategy in which peptide analogs were used to explore potential bioactive conformation(s) of the arginine-glycine-aspartic acid (RGD) sequence. The RGD sequence is a critical recognition feature that mediates the binding

Table 1. Structure–Activity Trends of Linear RGD Analogs.[a]

molecule	structure	ELISA IC_{50} (μM)
1		0.010
2		0.033
3		0.700
4		1.100
5		0.009

[a] The guanidine moiety is represented using the abbreviation guan.

of a number of adhesive proteins, including fibrinogen, to the platelet glycoprotein IIbIIIa (GPIIbIIIa) and other integrins.[11] RGD-containing peptides can block GPIIbIIIa/fibrinogen binding, are effective inhibitors of platelet aggregation, and therefore represent leads in the development of agents for the treatment of arterial thrombotic diseases.[12]

We have recently used NMR studies to examine the conformational profiles of three related RGD-containing peptides: (cyclo)S-Acetyl[1]-D-Tyr[2]-Arg[3]-Gly[4]-Asp[5]-Cys[6]-OH and the two epimeric sulfoxides derived by oxidation of the thioether.[13] One of the sulfoxides, G4120, displays a remarkably rigid structure in water at neutral pH as evidenced by extreme coupling constants and unusual side chain–side chain NOE signals.[13] In ELISA and platelet aggregation assays, G4120 displays IC_{50} values of 0.002 and 0.150 μM, respectively,[14] making it one of the most potent RGD-based inhibitors of platelet aggregation reported. The solution structure of G4120 is stabilized by a bifurcated hydrogen bond between the carbonyl oxygen of acetyl[1] and the amide hydrogens of Gly[4] and Cys[6], which results in a type-II' β-turn from acetyl[1] to Gly[4] and a "cupped" presentation of the RGD sequence. Because the less potent sulfide and epimeric sulfoxide lacked many of the diagnostic features observed in the NMR spectra of G4120, we proposed that the solution structure of G4120 was related to its bound conformation.[13]

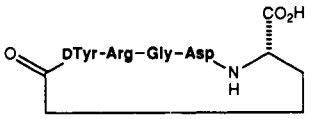
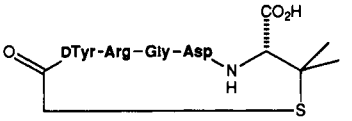
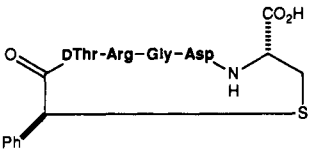
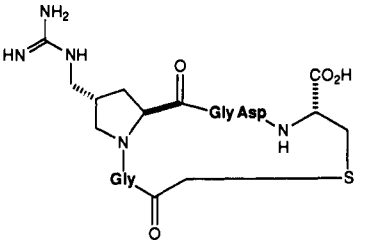
We describe here a complimentary exploration of the structural requirements for RGD-mediated antagonism using ensemble molecular dynamics studies to identify consensus conformations of bioactive linear peptidomimetics and cyclic peptides. These studies, combined with the earlier NMR work, support a self-consistent structural model that has enabled the *de novo* design of a potent non-peptidal agent.[15]

Materials and Methods

Peptide/Peptidomimetic Analogs. The chemical and conformational requirements for RGD binding were initially examined by constructing a series of analogs of the linear peptide GRGDV (Table 1) in which the

(11) Ruoslahti, E.; Pierschbacher, M. D. *Science* **1987**, *238*, 491.
(12) Nichols, A. J.; Ruffolo, R. R., Jr.; Huffman, W. F.; Poste, G.; Samanen, J. *Trends Pharm. Sci.* **1992**, *13*, 413.
(13) McDowell, R. S.; Gadek, T. R. *J. Am. Chem. Soc.* **1992**, *114*, 9245.
(14) 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.; Burnier, J. P. *J. Med. Chem.* **1992**, *35*, 2040.
(15) McDowell, R. S.; Blackburn, B. K.; Gadek, T. R.; McGee, L. R.; Rawson, T.; Reynolds, M. E.; Robarge, K. D.; Somers, T. C.; Thorsett, E. D.; Tischler, M.; Webb, R. R. III; Venuti, M. C. *J. Am. Chem. Soc.*, following paper in this issue.

Table 2. Activities of Cyclic RGD Peptides

molecule	structure	IC_{50} (μM) ELISA	IC_{50} (μM) PRP
6		0.003	0.125
7		0.007	0.200
8		0.002	0.300
9		0.002	0.420

Arg-Gly fragment is replaced with linkages that vary the relative placement of rigid and flexible regions. These molecules are expected to accommodate different spatial presentations of the charged moieties and thus serve collectively as "probes" for mapping potential binding topographies of the RGD sequence.

In previous studies of cyclic RGD peptides, we had observed that the chirality with which groups were attached to the cyclization "linkage" profoundly affects activity.[14] Because the chemical composition of these groups was less consequential than the chirality with which they were attached, we reasoned that the cyclic peptides that retained activity must be able to present a common conformation of the RGD sequence. Some of these cyclic peptides are listed in Table 2. Molecules **6–8** utilize variations of the cyclic peptide "linkage" to affect the conformation of the RGD epitope, while molecule **9** constrains the backbone ϕ and side chain χ_1 angles of the arginine via cyclization.

Cyclic Hexapeptide Analog. As a control to test the relevance of the "cupped" conformation observed for G4120, we embedded the RGD sequence in a cyclic hexapeptide (cyclo)-D-Tyr^1-Arg^2-Gly^3-Asp^4-Phe^5-Gly^6 (**10**) which was designed to enforce an extended RGD conformation. Cyclic hexapeptides of the form (cyclo)-X-X-Gly-X-X-Gly often adopt extended structures consisting of β-turns initiated at Gly^3 and Gly^6, as evidenced by NMR and crystallographic studies.[16]

Peptide Synthesis. Molecules **1–9** were prepared utilizing the same Boc protection schemes and polystyrene solid support used to synthesize G4120 and related analogs.[14]

The protected linear precursor to **10**, Tyr^1-Arg^2-Gly^3-Asp^4-Phe^5-Gly^6, was prepared using standard Boc protected amino acids and BOP couplings. The Gly^6 residue was attached to a modified polystyrene resin via a thioester linkage. On removal of the Boc group from Boc-Tyr^1-Arg^2-Gly^3-Asp^4-Phe^5-Gly^6-S-resin, the side-chain-protected peptide spontaneously cyclized and cleaved from the resin.

Assays. Compounds were initially evaluated for potency using an ELISA assay which measures the inhibition of the association of solubilized GPIIbIIIa with fibrinogen coated on a microtiter plate.[14] Potent compounds (IC_{50} < 0.1 μM) were further evaluated in a platelet aggregation assay using ADP-stimulated human platelets in platelet-rich plasma (PRP).

Ensemble Molecular Dynamics. Potential consensus conformations of the active linear (**5**, Table 1) and cyclic (**6–9**, Table 2) analogs were mapped using the ensemble molecular dynamics procedure previously described. The activity of linear analog **1** indicated that the backbone amides of the arginine residue are not critical for GPIIbIIIa/fibrinogen binding; **1** and **2** were not included in the ensemble because their flexibility would do little to constrain potential RGD conformations. Earlier studies demonstrated the primary importance of the charged arginine and aspartic acid side chains in binding and suggested important secondary roles for the hydrophobic side chains flanking these residues.[14] The primary objective of this study was to define consensus conformations of the RGD backbone that could provide similar presentations of these charged side chains. We therefore chose the guanidine group of the arginine, the carboxylate of the aspartic acid, and the α- and β-carbons of the RGD sequence as the essential "determinants" for alignment. It was further assumed that binding to GPIIbIIIa would require an approximate alignment of the side chains of the flanking hydrophobic residues. Based on this analysis, the atoms used for tethering are listed in Table 3, along with the final values of $k_{i,j}$ and $r^0_{i,j}$ which were used for the aligned trajectories.

The timecourse of a dynamics cycle is illustrated in Figure 2. Initial geometries for each cycle were generated using the distance geometry program DGEOM[17] and refined using conjugate gradients minimization. Starting with an initial high-temperature "randomization" in which the molecules were essentially unconstrained, the temperature of the ensemble was gradually lowered to 300 K, while $k_{i,j}$ and $r^0_{i,j}$ were adjusted to align the corresponding groups. Structures were subsequently sampled each picosecond for 100 ps in the presence of stringent constraints. A total of 10 such cycles were calculated. To assess the energetic impact of the tethering potentials, parallel simulations were conducted in the absence of constraints.

The all-atom AMBER force field[18,19] was used for all energy calculations, employing an infinite cutoff for nonbonded interactions and a linear dielectric ($\epsilon = 4.0r$) to partially compensate for the lack of explicit solvent.[18] Atomic charges on the termini of charged groups (Arg guanidine, Asp, Cys, and Pen carboxylates) were scaled by 0.25 relative to the default values to avoid artifacts due to excessive charge–charge interactions. Calculations were performed on a Silicon Graphics 4D/380 computer using a modified version of the Discover program (Biosym, San Diego).

NMR Measurements on Cyclic Hexapeptide. The spin systems of the individual amino acids were identified by inspection of the 1D spectrum

(16) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Prot. Chem.* **1985**, *37*, 1.

(17) Blaney, J. M.; Crippen, G. M.; Dearing, A.; Dixon, J. S. DGEOM, QCPE Program No. 590.

(18) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

(19) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

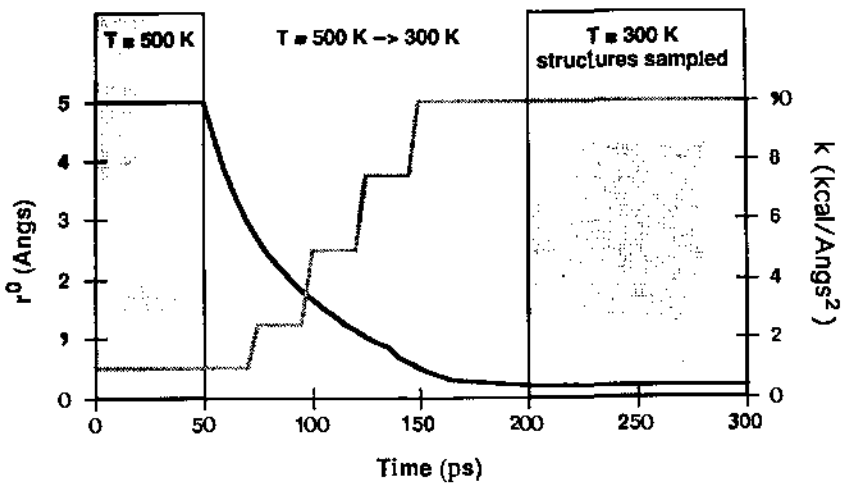


Figure 2. Graph illustrating the timecourse of a cycle of ensemble dynamics. An initial 50-ps interval is calculated at high temperature (500 K) with large constraining distances (5 Å; black line) and low tethering force constants (0.5 kcal-Å^{-2}; grey line), thereby "randomizing" the ensemble. Over a subsequent 150-ps interval, the simulation temperature is lowered to 300 K, the constraining distances are gradually lowered to 0.2 Å, and the tethering force constant is increased to 25 kcal-Å^{-2}, thereby inducing an alignment of the structures. Following equilibration, structures are sampled for 100 ps under constant temperature and constraint conditions.

in conjunction with the ^{1}H–^{1}H COSY[20,21] spectrum. The Gly^6 signals were distinguished from the Gly^3 signals by the presence of medium range crosspeaks with the aromatic hydrogens of Tyr^1 and Phe^5 in a ROESY[22,23] experiment. Similarly, the β-methylenes of Tyr^1 were distinguished from those of Phe^5 by the presence of a strong crosspeak with the aromatic H$^\delta$ in the ROESY spectrum. A table of relevant chemical shifts and coupling constants is available as supplementary material.

Conformations consistent with the φ angles determined from $^3J_{\alpha\beta}$ coupling constants[24,25] and interproton distances determined from ^{1}H–^{1}H dipolar interactions observed in the ROESY spectrum were calculated using DGEOM and refined using energy minimization as previously described.[13] Crosspeak intensities were classified into three catagories (strong, moderate, and weak) based on visual inspection of the contour plots, noting the area of the lowest contour as well as the number of contours observed.

Results

Peptide/Peptidomimetic Analogs. ELISA values for the linear peptidomimetic analogs are listed in Table 1. The series of linear analogs **1–4** illustrate the requirements for local conformational flexibility in the linkage connecting the guanidine group to the Asp residue. Although these analogs are all quite flexible, the dramatic reduction in activity observed for molecules **3** and **4** suggests that the alternative placement of the amide group renders the required geometry for RGD binding energetically unfavorable. Molecule **5** presents a rigidified "Arg" side chain.

ELISA and platelet aggregation values for the cyclic peptides are shown in Table 2. Molecules **6–9** are all effective inhibitors of both GPIIbIIIa/fibrinogen binding and platelet aggregation and are therefore assumed to accommodate a similar presentation of the RGD epitope.

Ensemble Molecular Dynamics. Each of the 10 ensemble dynamics cycles produced 100 aligned structures sampled over a 100-ps interval. The mean potential energies of these alignments, exclusive of the tethering potential, are listed in increasing order in Table 4. Also listed for comparison are the mean energies of five unconstrained simulations, each of which was started from a different minimized distance geometry structure. Four of the consensus alignments (A–D) actually had lower energies than any of the unconstrained simulations, while three of the alignments (H–J) had higher overall energies. For each of the alignments,

Table 3. Sets of Atoms Used for Ensemble Dynamics Constraints and Final Values of Tethering Parameters for Aligned Structures[a]

	molecule						
atom	**5**	6	7	**8**	**9**	k_{ij} (kcal-A^{-2})	r^0_{ij} (Å)
DTyr^2 CA		x	x	x	x	5.0	**1.5**
DTyr^2 CB		x	x	x		5.0	**1.5**
Arg^3 CA		x	x	x	x[b]	25.0	**0.2**
Arg^3 CB		x	x	x		25.0	**0.2**
Arg^3 CZ	x	x	x	x	x	25.0	**0.2**
Gly^4 CA	x	x	x	x	x	25.0	**0.2**
Asp^5 CA	x	x	x	x	x	25.0	**0.2**
Asp^5 CB	x	x	x	x	x	25.0	**0.2**
Asp^5 CG	x	x	x	x	x	25.0	**0.2**
Cys^6 CA	x	x	x	x	x	5.0	**1.5**
Cys^6 CB	x	x	x	x	x	5.0	**1.5**

[a] Tethered atoms are indicated with reference to a standard cyclic peptide framework (cyclo)S-Acetyl^1-DTyr^2-Arg^3-Gly^4-Asp^5-Cys^6-OH. Corresponding atoms were used when appropriate (e.g., DThr^2 CA in **8** and DGly^2 CA in **9** for DTyr^2 CA). [b] Due to differences in the cyclic peptide backbone, this atom was tethered to the centroid of the other three using a less stringent constaint: k_{ij} = 5.0 kcal-A^{-2} and r^0_{ij} = 1.5 Å.

Table 4. Potential and Constraint Energies for Ten Ensemble Dynamics Simulations and Five Unconstrained Simulations, along with Mean RMS Superimpositions of the RGD α- and β-Carbons to the Solution Structure of G4120[a]

simulation	potential energy (kcal/mol)	constraint energy (kcal/mol)	RMS fit to G4120 (Å)
	Ensemble Dynamics		
A	250.8 (10.0)	11.4 (2.3)	0.7
B	254.9 (10.9)	10.8 (2.4)	0.6
C	257.4 (10.6)	10.2 (2.4)	0.6
D	261.4 (11.4)	10.2 (2.6)	0.5
E	271.7 (10.0)	10.4 (2.3)	0.7
F	271.8 (10.3)	12.0 (2.8)	0.7
G	272.7 (9.9)	10.6 (2.6)	1.1
H	282.1 (8.8)	10.2 (2.6)	0.7
I	282.5 (13.2)	10.7 (2.3)	1.2
J	285.1 (9.8)	11.2 (2.3)	1.9
	Unconstrained Simulations		
1	266.7 (9.7)		1.7
2	273.5 (9.7)		1.7
3	273.7 (10.9)		1.6
4	275.8 (10.3)		1.9
5	278.8 (10.3)		1.8

[a] Simulations are listed in increasing order of mean potential energy, excluding tethering contribution. Constraint energies reflect the final values of k_{ij} and r^0_{ij} from Table 3. Standard deviations are indicated in parentheses, based on 100 structures sampled. RMS fits were calculated using analogs **6–8**.

the tethering function added approximately 4% to the net potential energy (Table 4). Following conjugate gradient minimization, the contribution of the tethering potential was less than 1.0 kcal for each of the alignments. The aligned conformations are therefore comparable in energy to unconstrained structures and are maintained without requiring the use of an overly restrictive tethering potential.

Despite starting from different conformations, alignments A–F and H converged to the same fold of the RGD backbone, with an average pairwise RMS deviation of 0.5 Å for the RGD α- and β-carbons. Because multiple arginine and aspartic acid side-chain rotamers were observed in these alignments, no unique consensus location was determined for either the guanidine group or the Asp carboxylate. This fold is remarkably similar to the backbone conformation observed in the solution structure of G4120, as indicated by the RMS deviations shown in Table 4. A representative set of structures minimized from alignment A is shown in Figure 3 along with the NMR structure of G4120. Representative structures minimized from the higher energy alignments G, I, and J are illustrated in Figure 4.

The average pairwise RMS deviations of the corresponding backbone atoms in the unconstrained simulations ranged from

(20) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229.
(21) Bax, A.; Freeman, R. *J. Magn. Reson.* **1981**, *44*, 542.
(22) Griesinger, C.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 261.
(23) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811.
(24) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11.
(25) Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741.

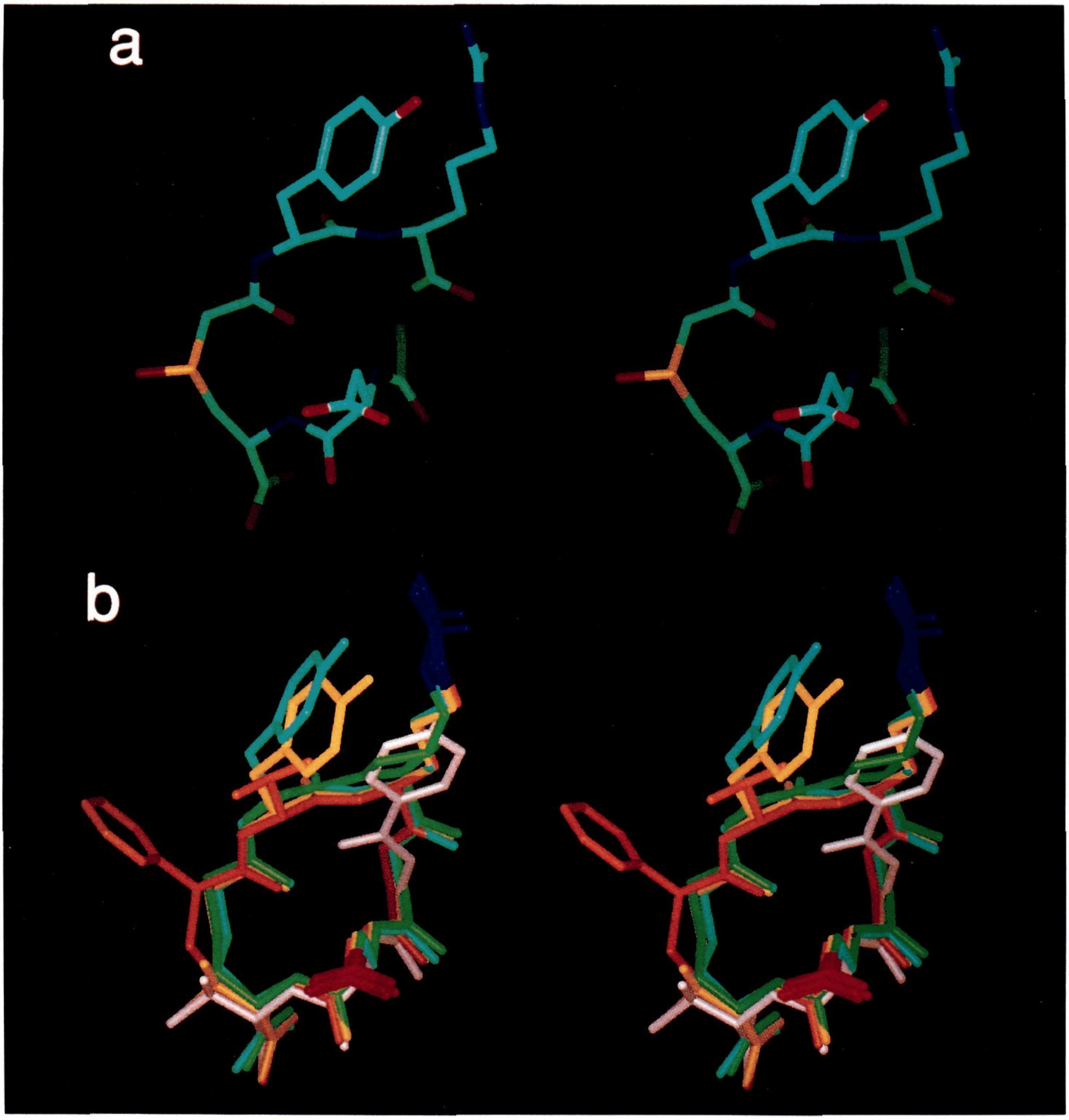


Figure 3. Stereoviews of the NMR structure of G4120 (a) and a representative set of structures minimized from alignment A (b). Alignments B–F and H converged to a similar structure but have less energetically favorable conformations of the cyclic "linkage". Analogs are color-coded according to the following scheme: **5**, white; **6**, cyan; **7**, yellow; **8**, orange; and **9**, green. The "Arg" guanidine and "Asp" carboxylate are colored blue and red, respectively. The NMR structure of G4120 and alignment A feature a similar conformation of the RGD sequence, as evidenced by the stereo plots and RMS fits noted in Table 4.

1.6 to 1.9 Å (Table 4), indicating that no single molecule in the ensemble was rigid enough to uniquely determine the backbone fold of the RGD sequence. Instead, each molecule selected a different subset of the conformational space available to the linear RGD sequence. The consensus alignments illustrated in Figures 3 and 4 therefore represent the intersections of those subsets. Although it is clear that multiple "intersections" are geometrically possible, virtually all of the low-energy alignments (A–D) reproduced the solution conformation of G4120. Three of the higher energy alignments (E, F, and H) had the same backbone fold of the RGD sequence but possessed less favorable conformations of the cyclic "linkage".

Cyclic Hexapeptide Structure. In the ELISA, **10** has an IC_{50} of 2.1 μM and is significantly less active than its linear counterpart GRGDF (IC_{50} = 0.082 μM). The NMR spectra of **10** indicate the presence of two conformations (approximately 4:1 ratio). Diagnostic NMR data for the major conformation are summarized in Table 5. The coupling constants, ROE signals, and upfield shifts of the Arg and Asp side-chain methylene hydrogens indicate that the structure of **10** in solution is dominated by the conformation shown in Figure 5, in which D-Tyr and Arg, respectively, occupy the i+1 and i+2 positions of a type-II′ β-turn, while Asp and Phe occupy the corresponding positions in a type-I turn. Although **10** and G4120 possess a similar β-turn in the

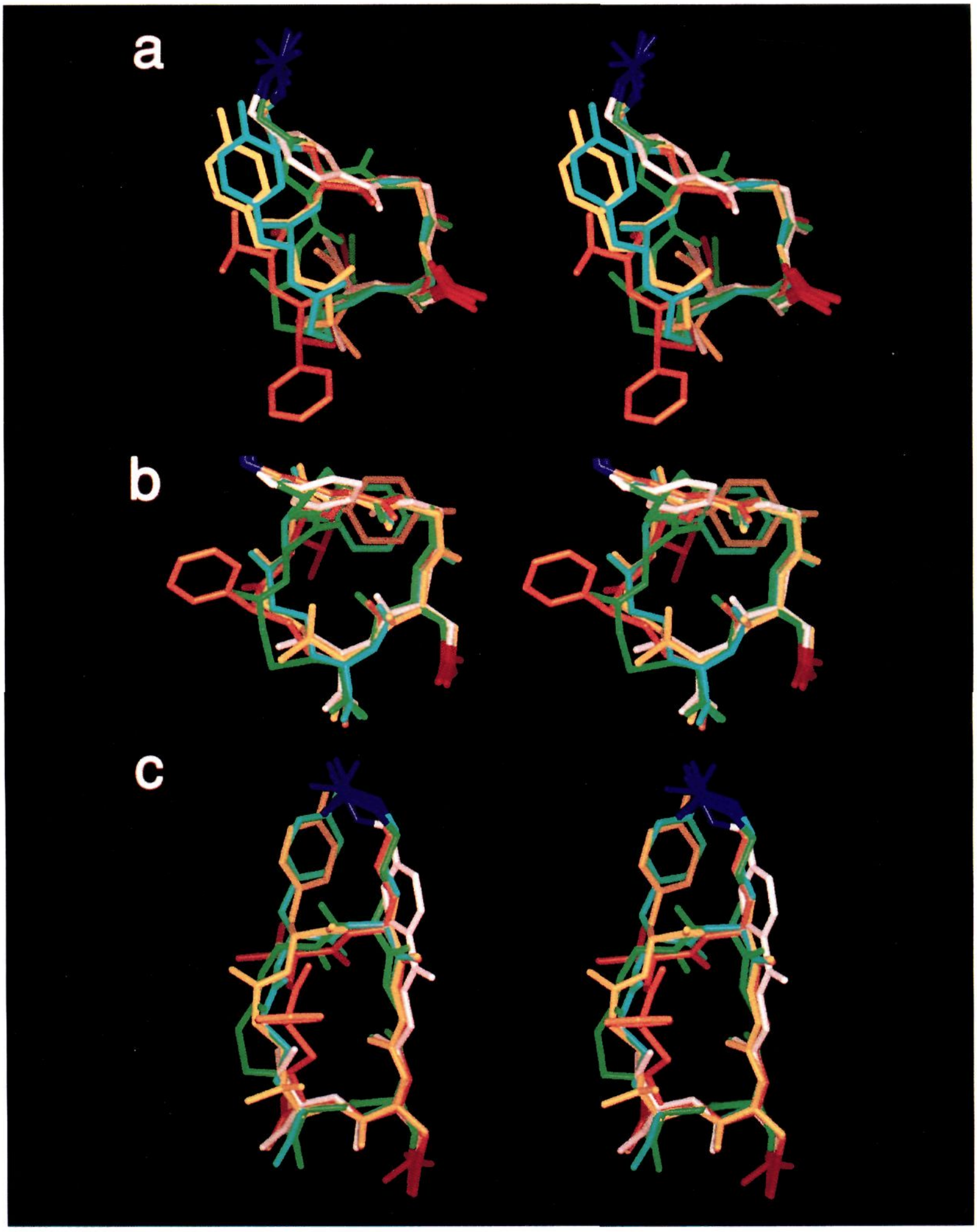


Figure 4. Higher energy alignments featuring alternate RGD conformations: G (a); I (b), and J (c). Coloring scheme as in Figure 3.

Table 5. Diagnostic NMR Data for c-D-Tyr[1]-Arg[2]-Gly[3]-Asp[4]-Phe[5]-Gly[6]

^3J(HN-H$^\alpha$), Hz		NOE crosspeaks			chemical shifts, ppm		
D-Tyr[1]	4.2	D-Tyr[1]	NH – Gly[6]	H$^\alpha$,H$^{\alpha'}$	Arg[2]	H$^\gamma$	0.88
Arg[2]	8.4	Arg[2]	H$^\alpha$ – Gly[3]	NH	Arg[2]	H$^{\gamma'}$	0.75
Gly[3]	8.8	Arg[2]	NH – D-Tyr[1]	H$^\beta$,H$^{\beta'}$	Asp[4]	H$^\beta$	2.35[a]
	<2	Arg[2]	H$^\gamma$ – D-Tyr[1]	H$^\epsilon$			
Asp[4]	4.9	Asp[4]	H$^\beta$ – Phe[5]	H$^\delta$			
Phe[5]	9.0						
Gly[6]	n.d.						

[a] Methylene hydrogens for Asp[4] were not resolved.

D-Tyr-Arg region, the extended RGD conformation observed in the dominant structure of **10** is clearly not productive for fibrinogen-GPIIbIIIa binding. Based on constrained minimizations of **10** using the RGD α- and β-carbon coordinates of G4120 as a template, the extended RGD conformation is favored over the cupped conformation by approximately 4 kcal. Insufficient data were available to determine a structure for the minor conformation.

Discussion

Molecular dynamics provides a useful means of exploring the conformational space of flexible molecules where numerous low-energy pathways exist for conformational interconversion. Despite random starting structures, all of the ensemble dynamics simulations were able to align the flexible RGD analogs at minimal

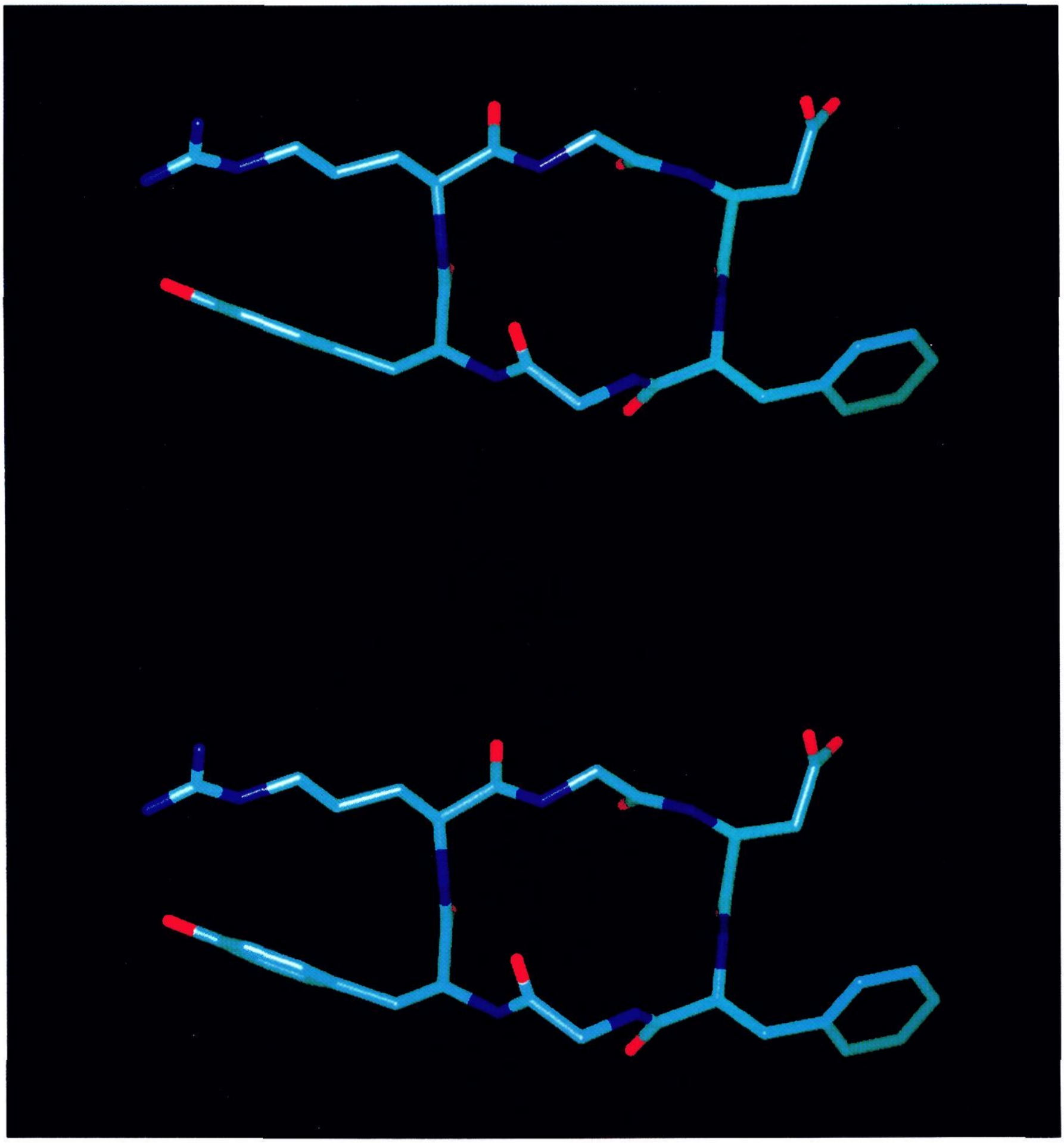


Figure 5. Sample structure for the primary solution conformation of (cyclo)-D-Tyr[1]-Arg[2]-Gly[3]-Asp[4]-Phe[5]-Gly[6] (**10**) suggested by the ROE signals, coupling constants, and upfield chemical shifts listed in Table 5.

energetic expense by gradually increasing the stringency of the tethering potential. Given the inherent limitations of molecular dynamics for traversing high-energy transitions (e.g., the "ring flip" of a fused ring system), we are exploring extensions of the method that might prove more appropriate when less flexible molecules are included. In principle, an identical tethering strategy could be combined with simulated annealing[26] instead of dynamics as the conformational "search engine". Replacing gradient-based procedures such as dynamics or simulated annealing with stochastic methods such as Monte Carlo sampling[8,9] would allow alignments to be generated on the basis of conformation-dependent properties (e.g., volume overlap, dipole moment) that are not directly accessible using simple atom-based tethering.

As with any conformational analysis approach, the validity of ensemble dynamics results depends on the balance between the completeness with which conformational space is sampled and the convergence of the resulting structural alignments. Although only a high-resolution systematic search about all torsional degrees of freedom can guarantee absolute completeness, sampling is enhanced if diverse starting points are used for multiple simulations. In this study, 10 sets of random, unconstrained distance geometry conformations were used as the basis for independent ensemble dynamics cycles. Sampling could possibly have been improved by generating a larger number of potential starting conformations and using cluster analysis methods to maximize

(26) Nigels, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *239*, 129.

the variation between conformations selected for subsequent alignment.

The convergence properties of ensemble dynamics are determined by the compounds and structural groups selected for alignment. Using a structurally diverse collection of molecules reduces the potential number of energetically accessible alignments to be found, thereby increasing the probability that a low-energy alignment is biologically relevant. Many spurious alignments of the RGD epitope would obviously have been generated had the ensemble not included the cyclic peptides **6–9**. Numerous artifacts could also emerge if tethering groups are not chosen based on rigorous structure–activity analysis: including too few groups can lead to nonsensical alignments, while enforcing stringent constraints for nonessential groups can eliminate potentially important conformations from being sampled. Based on earlier data, it was clear which shared groups were essential for binding and presentation of the RGD backbone. These groups were assigned "tight" tethering potentials, while shared functional groups implicated in a secondary role were assigned less restrictive potentials.

Variations in the composition of the cyclic peptide linkage used to generate G4120 resulted in structurally diverse molecules (**6–9**) of comparably potency as inhibitors of platelet aggregation. Although none of these analogs displayed a unique conformation of the RGD sequence in unconstrained molecular dynamics simulations, applying restraints based on the ensemble of molecules produced a dominant consensus conformation of the RGD backbone that is energetically comparable to the structures sampled in the unconstrained simulation. Compellingly, this conformation is very similar to the conformation determined by NMR for G4120, a potent rigid cyclic peptide. Higher energy alignments featured an extended RGD conformation; incorporating the RGD sequence into a framework (**10**) that favors such a conformation results in greatly reduced activity. These findings lead us to propose that the cupped presentation of the RGD sequence is required for high potency. This structural model has provided us with a "blueprint" for the design of a non-peptidal agent. The results of that design effort are presented in the next paper in this series.

Experimental Section

Linear and cyclic analogs **1–9** were prepared by Boc and/or Fmoc protocols described previously.[14]

Synthesis of (cyclo)-D-Tyr[1]-Arg[2]-Gly[3]-Asp[4]-Phe[5]-Gly[6] (10). *S*-Tritylthioglycolic acid (3 mM) was coupled to Phe-MBHA resin (1 mM) with BOP (3 mM) in DMA. The trityl group was removed by treatment of the resin with TFA (2% in CH_2Cl_2) at room temperature. The deprotection was monitored for the absence of a yellow color upon treatment of the resin (10–20 beads) with TFA (1–2 drops). BocGly (3 mM) was activated with DIPC (3 mM) and agitated with the thioglycolatePhe resin for 2 h at room temperature in the presence of *N*-methylmorphine, NMM (3 mM). After removal of the Boc group with TFA (50% in CH_2Cl_2), standard BOP (3 mM) couplings of Boc protected amino acids (3 mM) were used to prepare Boc-D-Tyr[1]-Arg[2]-Gly[3]-Asp[4]-Phe[5]-Gly[6]-thioglycolatePhe resin. Deprotection of the N-terminus (50% TFA in CH_2Cl_2), followed by neutralization with NMM in DMA enabled an intramolecular cyclization between the amino group of the D-Tyr[1] and the carbonyl of the Gly[6] residue. The crude side-chain-protected cyclic peptide was obtained as a solution in DMA after 16 h at room temperature. The DMA was removed under vacuum and the side chains were deprotected with HF using anisole and ethyl methyl sulfide as scavengers.

All compounds were purified by reverse-phase HPLC and characterized by high-resolution mass spectrometry and amino acid analysis. Characterization data for compound **8** has previously been reported.[14] Data for compounds **1–7**, **9**, and **10** are found in supplementary material.

NMR spectra were recorded on a Bruker AMX 500 spectrometer at 293 K.

NMR Conformational Studies of cyclo-D-Tyr[1]-Arg[2]-Gly[3]-Asp[4]-Phe[5]-Gly[6] (10). The peptide (6 mg, 0.01 mM) was dissolved in DMSO-d_6 (0.1 mL, 99.9% ^2H) and diluted with ^2H_2O (0.9 mL, 99.9% ^2H) buffered to pH 7 with sodium phosphate (10 mM). ROESY spectrum: relaxation delay 2 s, mixing time 200 μs, spin-lock field 4000 Hz.

Fibrinogen/GPIIbIIIa Solid-Phase ELISA. The details of this assay have been described previously.[14]

Acknowledgment. High-resolution mass spectra were obtained by Kathy O'Connell. The Fibrinogen/GPIIbIIIa ELISA was performed by Craig D. Muir and Chip Alee. Amino acid analyses were performed by Allan Padua.

Supplementary Material Available: ROESY and COSY spectra of **10** and tables of analytical data (HRMS and amino acid analysis) for compounds **1–7**, **9** and **10** (4 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.