Combining Pharmacophore Search, Automated Docking, and Molecular Dynamics Simulations as a Novel Strategy for Flexible Docking. Proof of Concept: Docking of Arginine−Glycine−Aspartic Acid-like Compounds into the $\alpha_v\beta_3$ Binding Site

Nicolas Moitessier,*,†,§ Christophe Henry,† Bernard Maigret,‡ and Yves Chapleur†

Groupe SUCRES and Equipe de Dynamique des Assemblages Membranaires, Unité Mixte 7565 CNRS, Université Henri Poincaré-Nancy 1, B.P. 239, F-54506 Nancy-Vandoeuvre, France

Received December 22, 2003

A novel and highly efficient flexible docking approach is presented where the conformations (internal degrees of freedom) and orientations (external degrees of freedom) of the ligands are successively considered. This hybrid method takes advantage of the synergistic effects of structure-based and ligand-based drug design techniques. Preliminary antagonist-derived pharmacophore determination provides the postulated bioactive conformation. Subsequent docking of this pharmacophore to the receptor crystal structure results in a postulated pharmacophore/receptor binding mode. Pharmacophore-oriented docking of antagonists is subsequently achieved by matching ligand interacting groups with pharmacophore points. Molecular dynamics in water refines the proposed complexes. To validate the method, arginine−glycine−aspartic acid (RGD) containing peptides, pseudopeptides, and RGD-like antagonists were docked to the crystal structure of $\alpha_v\beta_3$ holoprotein and apoprotein. The proposed directed docking was found to be more accurate, faster, and less biased with respect to the protein structure (holo and apoprotein) than DOCK, Autodock, and FlexX docking methods. The successful docking of an antagonist recently cocrystallized with the receptor to both apo and holoprotein is particularly appealing. The results summarized in this report illustrated the efficiency of our light CoMFA/rigid body docking hybrid method.

Introduction

Ligand−protein interactions have been recognized as central phenomena in most biological processes ranging from catalysis to signaling.[1] Theoretical and computational methods are widely used to analyze, investigate, and predict ligand−macromolecule interactions. Computational approaches also allow the design of new drugs for a variety of diseases without recourse to long and expensive experimental efforts.[2−6] Consequently, medicinal chemists increasingly adopt in silico high-throughput docking-based screening as a fundamental discovery tool.[7−9]

The earliest approaches to the so-called docking problem treated ligands and macromolecules as rigid body objects. The programs attempted to define the most energetically favorable associations, which would in turn form putative stable complexes. Experimental evidence highlighted the importance of the flexibility of both docking partners in the binding process, thus revealing the limitations of the rigid body docking model. As computational performance increased, rigid docking methods have evolved into flexible docking methods that explore large conformational spaces. Since the pioneering work from the Goodford[10−12] and Kuntz[13−16] groups, numerous powerful packages of software have been disclosed and reviewed.[2,3] Considerable efforts have been devoted to the improvement of the search engines and of the accuracy of the scoring functions. The second-generation procedures eventually showed significant improvements in binding predictions.[2,3] However, additional constraints (such as solvation and application to virtual screening) have emerged, forcing methodology improvements.

The conformational analysis tool should effectively explore the scoring function landscape within reasonable CPU time. In fact, the systematic conformational search was rapidly abandoned in favor of stochastic methods (including Monte Carlo, simulated annealing, and genetic algorithms) and other strategies such as incremental construction or distance geometry.[2] The scoring functions should top-rank the experimentally observed binding modes and/or approximate the binding free energy.[17] A major drawback of current docking methods is the oversimplification of the binding process (e.g., rigid protein, single conformation of the ligand, treatment of crystallographic water molecules). Low predictive accuracy of the scoring functions stems from these computationally convenient assumptions.[18−23] In 1994, Leach reported the first example of docking to partially flexible binding sites.[24] More recent efforts have been directed toward the docking to conformational ensembles.[25−27] Other major actors in the ligand docking to proteins are the water molecules that could play a key role in the binding event. A possible strategy that would include both the solvent effects and the flexibility of both partners is the use of simulated annealing

* To whom correspondence should be addressed. Phone: 514-398-8543. Fax: 514-398-3797. E-mail: nicolas.moitessier@mcgill.ca.
† Groupe SUCRES.
§ Present address: Department of Chemistry, McGill University, Sherbrooke Street W., Montréal, Québec, H3A 2K6, Canada.
‡ Equipe de Dynamique des Assemblages Membranaires.

10.1021/jm0311386 CCC: $27.50
Published on Web 07/20/2004

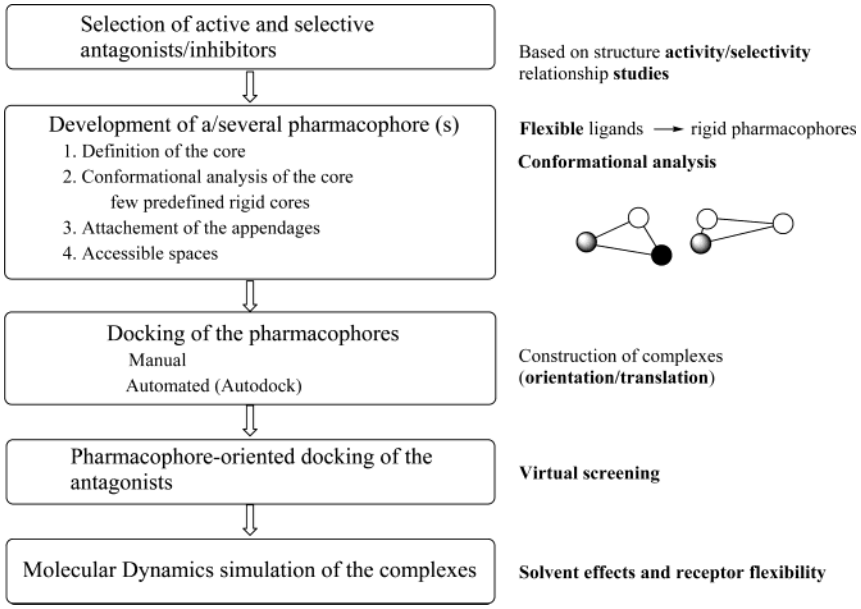


Figure 1. Flow chart of the full protocol.

molecular dynamics simulations with an accurate force field.[28] Unfortunately, such calculations are tractable only for a few starting configurations of the macromolecule−ligand system because of the computational expense.

The independent docking studies on a series of serine protease and on $\alpha_V\beta_3$ integrin have illustrated the sensitivity of docking approaches to the protein structure.[21,29−32] Two docking studies were carried out with the crystallized $\alpha_V\beta_3$ apoprotein structure. Binding models have been developed by means of molecular dynamics (MD) simulation[29] and the Autodock docking program.[30] These two studies on the apoprotein predicted a better interaction of the arginine side chain mimics of the antagonists with Asp150. The experimentally solved structure revealed that the interaction with Asp218 was favored over interaction with Asp150.[31]

Many docking methods were evaluated for their ability to reproduce protein−ligand complexes. Cross-docking, defined as docking of a series of ligands to a known protein structure, is more indicative of the accuracy of the methods. The ideal method should predict the binding mode of any ligand without any biases from the receptor structure (docking to holoprotein and apoprotein or homology models should provide similar binding modes).

Herein, we present a new strategy for docking flexible ligands to flexible proteins. This technique can be decomposed into four steps: (1) pharmacophore definition, (2) docking of this pharmacophore, (3) oriented docking of the ligands, and (4) MD simulation. The MD refinement, which relies on a classical molecular mechanics force field, enables the receptor to move in an aqueous medium. We illustrate this novel strategy with arginine−glycine−aspartic acid (RGD) like molecules, which are known integrin antagonists. Since the discovery of the fundamental role of the $\alpha_v\beta_3$ integrin in a large number of physiological disorders, its RGD-containing ligands have received considerable attention in contemporary pharmaceutical research.[33,34] In the present work we have generated a three-point pharmacophore of $\alpha_v\beta_3$ antagonists and then further docked it to the binding site of the crystalline structure of the unbound receptor.[35] The defined complex was the basis for directed docking and MD study of a representative antagonist. The high accuracy of this approach led us to extend the scope of the method to include virtual screening. This method was found to be less sensitive to the protein flexibility than Autodock, DOCK, and FlexX methods. This validation added further credence to the method.

Results and Discussion

Method. The full protocol is presented in Figure 1. Selective, active, and structurally diverse antagonists are first selected and then used to construct a three-point pharmacophore. The pharmacophore captures both the common geometric and electronic features and the bioactive conformation of these antagonists. Subsequent automated or manual docking of this pharmacophore provides the suitable binding modes. This predicted interaction pattern provides a basis for oriented rigid docking by simply matching the antagonist interacting groups with the pharmacophore points. Since the binding mode is predefined, the protocol restricts the docking and scoring processes to effective binding modes and precludes the use of a CPU time intensive conformational search. This method can therefore be expanded to virtual screening of libraries of compounds. In the present approach, the pharmacophore development accounts for the ligand flexibility while the docking is rigid. Further MD simulation accounts for the receptor flexibility and solvent effects.

Target Selection. $\alpha_v\beta_3$ integrin was chosen as a validation target. This study has combined our recent interest in the design and preparation of $\alpha_v\beta_3$ integrin antagonists with our interest in docking methods.[36,37] In addition, the low accuracy in predictions of the docking studies of antagonists to the unbound receptor structure was also an impetus for this work. The previously reported docking studies on the apoprotein predicted a better interaction of the arginine side chain mimics of the antagonists with Asp150 while interaction with Asp218 was observed.[31] During the preparation of

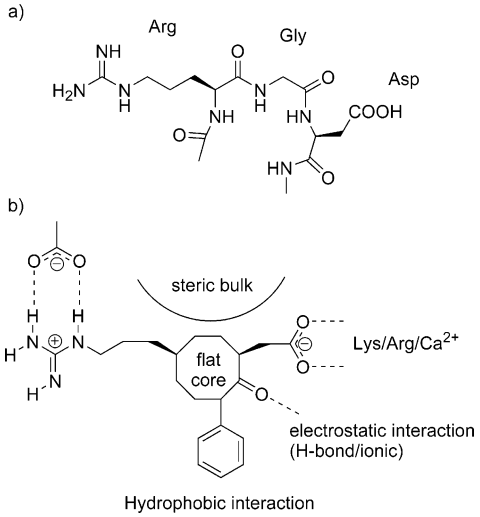


Figure 2. (a) Active sequence and (b) proposed binding mode.

this manuscript, Kessler and co-workers have published another docking study using the structure of the receptor cocrystallized with a cyclopeptide (holoprotein).[32] The observed induced fit was critical for the success of this last docking study. Ideally, the accuracy of docking method should not rely on the available structure. Finally, providing synthetic antagonists of $\alpha_v\beta_3$, a cell surface receptor called vitronectin receptor, may offer an effective therapy for treating diseases.[33,34] Although potent antagonists have been rapidly disclosed, further development of selective ligands toward this particular integrin was not straightforward because of the structural similarities among the integrin superfamily members. A careful study of the structure−activity relationship studies led us to propose a tentative two-dimensional picture of the pharmacophore and its binding interactions (Figure 2).[33] The antagonists share two common structural features that are determinants for receptor recognition: a carboxylate and a guanidinium-like moiety (Figure 2). The two charged groups interact with complementary residues or cation in the binding site by mimicking the arginine and aspartic acid side chains of the adhesive RGD sequence.[33] In addition, a hydrophobic group was often introduced next to the carboxylate group resulting in enhanced potencies.

Data Set Selection. The putative pattern proposed in Figure 2 was further exploited to develop a three-point pharmacophore. The investigated compounds have been chosen for their selectivity toward the $\alpha_v\beta_3$ integrin over the $\alpha_{IIb}\beta_3$ integrin, a competitive receptor involved in platelet aggregation. To establish as general a model as possible, peptidic and nonpeptidic candidates presenting apparent diversities were incorporated in the training set (Charts 1 and 2).

Cyclo(RGDfV) **1**[38,39] and its methylated analogue **2**[40] from Kessler's group were selected. They both feature the RGDfV sequence, albeit in two different conformations. The pseudopeptide **3** from DuPont-Merck[41] and the dimeric RGD sequence **4**[42] were also included in this study (Chart 1). Four nonpeptidic antagonists were also chosen and are presented in Chart 2.[33] They are based on different scaffolds. Antagonists **5** and **7** share a benzodiazepine template, while compounds **6** and **8** were constructed around a urea and an indazole core, respectively. In regards to the aromatic or hydrophobic group,

Chart 1. Selected Peptidic and Pseudopeptidic Antagonists

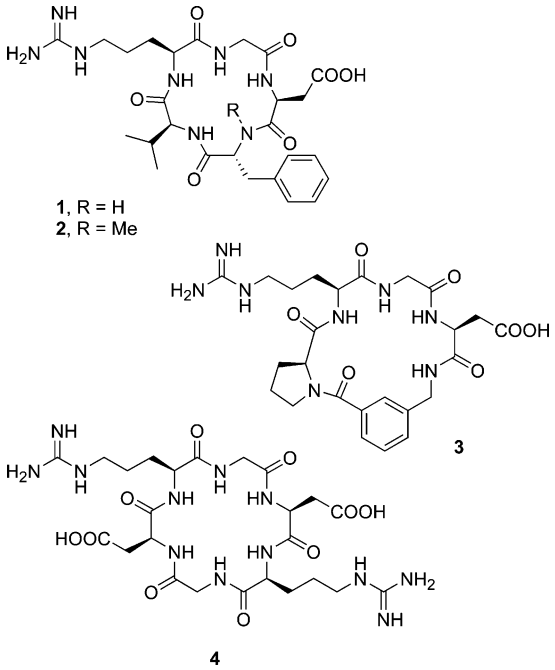
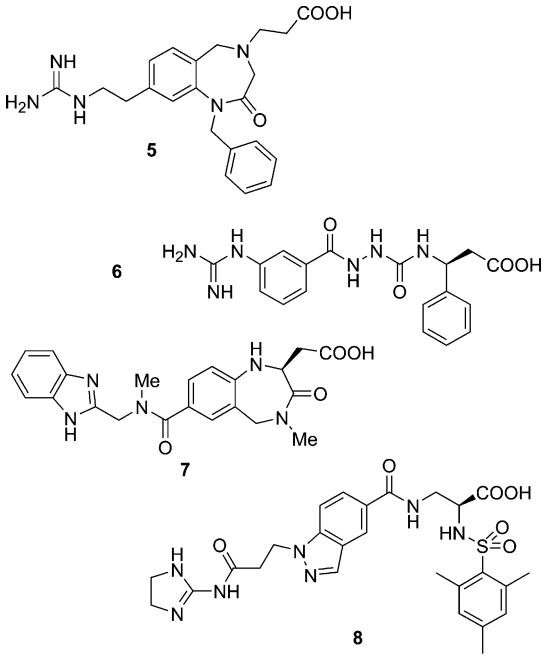


Chart 2. Selected Nonpeptidic Antagonists



4 and **7** were outliers and were selected to increase the diversity of the training set.

Three-Point Pharmacophore Construction. The search for the common bioactive shape adopted by a series of receptor antagonists is a challenging task because many conformations coexist in solution. The difficulty of conformational analysis increases with highly polar compounds. Few approaches are available including simulations (e.g., simulated annealing or Monte Carlo) and evolutionary approaches (e.g., genetic algorithms, evolutionary programming), which relied on energy computation and geometric rule-based methods (e.g., systematic search). Although the systematic search is often time-consuming for highly flexible compounds, it considers all the geometrically accessible conforma-

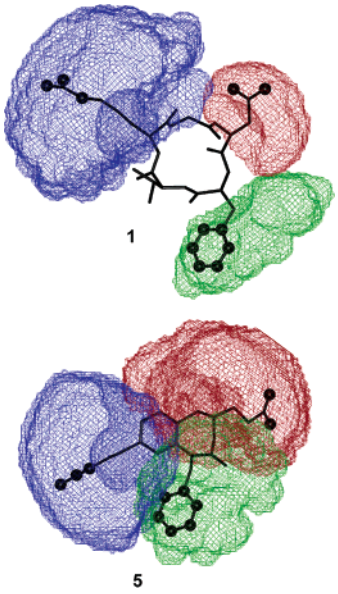


Figure 3. Geometrically accessible volumes for the pharmacophoric groups as exemplified by antagonists **1** and **5**.

tions (the convergence criterion). Exploring the conformational hypersurface proves to be more difficult when using molecular dynamics or other techniques. Although our previous work relying on a simulated annealing and MD combination proved to be valuable in providing a pattern common to a series of $\alpha_{IIb}\beta_3$ antagonists, it suffered from disadvantages directly related to the use of molecular dynamics.[36] An exhaustive survey of solution conformations requires long molecular dynamics trajectories and provides a large amount of data that are often difficult to analyze. In contrast, geometric rule-based methods are easy to use. Additionally, computing the potential energy of any conformation would require evaluating the solvent contribution (the integrin antagonists are charged molecules that cannot be accurately studied in a vacuum).

In this context, we proposed a pharmacophore based on geometric considerations rather than energy evaluations. Steric fields for each pharmacophoric groups were derived by systematic conformational search. Prior selection of crucial groups for activity was necessary (ionic+, ionic−, and hydrophobic groups in the present work). Although this situation can be viewed as a drawback of the method, a large body of SAR data supported this choice.[33] It is noteworthy that it also requires user interference when choosing the core or aligning the sets of volumes (vide infra). Recently, a four-point pharmacophore study of integrin antagonists also based on geometric rules was proposed by Pickett and co-workers using the ChemDiverse suite of programs.[43] This excellent report revealed the potential of pharmacophore search in integrin antagonist design.

In the present work, the central cores were kept frozen and the side chain dihedral angles were systematically rotated. This resulted in sets of volumes for each investigated compound. Except for **4** and **7**, which did not feature any hydrophobic group, three volumes were delineated. Two examples of this mapping are presented in Figure 3.

Stepwise manual alignment of both central cores and sets of volumes followed by Boolean computation of overlapping volumes led to a common geometrically accessible region for each group. It is worth noting that different chronological orders of superposition were tested, leading to roughly similar patterns. The resulting model is presented in Figure 4. The blue region indicates the area where a positively charge is expected. The red contour demarcates the volume where a negatively charged moiety is predicted to participate in the binding. Finally, the green surface surrounds a region where a hydrophobic group flanking the acid would be favorable to the activity. The average interpoint distances are shown and can be modulated owing to the flexibility of the ionic+ side chain.

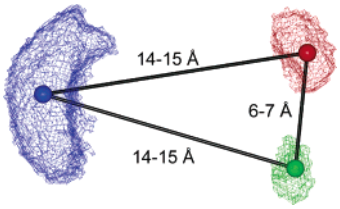


Figure 4. Three-point pharmacophore.

The relative spatial positions of the negatively charged and hydrophobic groups are well defined. This provides an interesting description of the geometrical distribution occurring within the receptor binding site. The developed model also confirms that strict requirements should be satisfied to provide active antagonists. Another interesting feature of this model is the large size of the blue volume corresponding to the positively charged pharmacophoric group. This large volume may stem from a necessary flexibility for activity. For example, few complementary side chains (Asp or Glu) may be located in the corresponding binding pocket that would constitute a relatively large and negatively charged region and allow the guanidinium to move while keeping some ionic interactions.

Docking of the Pharmacophore. An X-ray crystal structure of an extracellular part of the $\alpha_v\beta_3$ integrin was disclosed and was further exploited in our docking studies.[35] Visual inspection of the postulated binding site revealed that two calcium atoms could interact with the negatively charged point of the pharmacophore, while three carboxylate-containing sites (two sites on α_v, which are Asp 150 and Asp 218; one site with two aspartic acids, which are Asp 219 on α_v and Asp 217 on β_3) might bind to the ionic+ point. Thus, three alternative pharmacophore−receptor complexes were manually built (Figure 5). Automated docking of this pharmacophore with Autodock[44,45] was also carried out. Interestingly, all the 50 runs performed with this suite of programs converged toward a single orientation. This obtained orientation corresponded to one of the three binding modes previously proposed. This orientation positioned the acidic moiety in proximity to the calcium cation and induced interactions between Asp218 and ionic+ and between Tyr122 and the hydrophobic vertex of the triangle (Figure 5). To get more than one orientation, we decreased the number of generations (poorer convergence) and performed another 50 runs. In all the cases, the ionic− pharmacophore interacting site was positioned in the free coordination site of the calcium cation. Second, Autodock[44,45] targeted the three aspartic acids although interactions with Asp218 led to the best scores. This nicely validated the initially

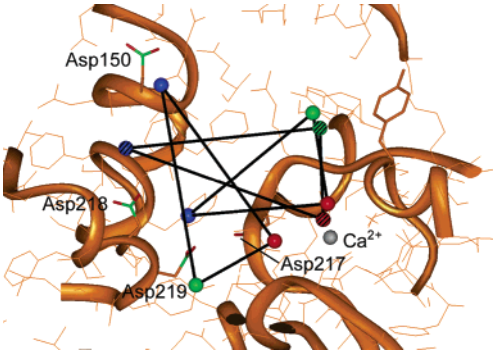


Figure 5. Pharmacophore docking: three chosen pharmacophore binding modes. Autodock proposed binding mode is hashed.

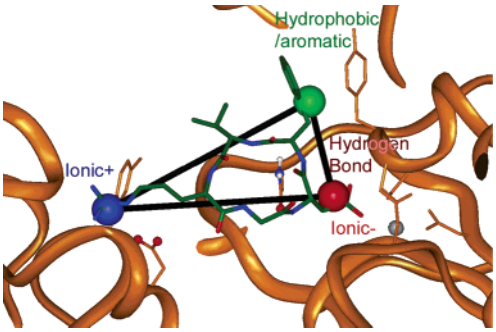


Figure 6. Modeled structure of cyclo(RGDfV)−$\alpha_v\beta_3$ complex.

proposed three orientations for the pharmacophore. Moreover, any binding mode of this three-point pharmacophore was more than 0.5 kcal·mol^{-1} higher in energy than the top-ranking one. This value appears low because it represents the score of only three dummy atoms.

Molecular Dynamics Simulation. To further evaluate the complementarity of the ligands and the receptor and to account for the induced fit, molecular dynamics simulation in explicit water was carried out. Cyclo(RGDfV) **1** was chosen as a representative ligand for this study. This synthetic ligand features the naturally occurring RGD sequence in a conformationally constrained form. The three complexes (as defined by manual docking) were solvated and then energy-minimized. These slightly optimized complexes were further refined by iterative MD and minimization stages until stable energy values were reached. This preliminary optimization of the first complex led initially to strong ionic interactions of the ligand with Asp 150 and the first calcium ion. However, a jump of the ligand was rapidly observed from the first to the second calcium atom, thus leading to a new arrangement. This jump also broke the eastern ionic bond destabilizing the complex. The second complex complied with the binding requirements extracted from SAR data (Figures 2 and 5), while the third kept strains even after the minimization stage. It is worth noting that this second binding mode matched the highest scoring pharmacophore binding with Autodock. This second fully optimized complex was finally subjected to MD simulation in explicit water (100 ps, 300 K) and shows a great stability. A picture of this complex is shown in Figure 6.

The interactions predicted from SAR data (Figure 2) occurred. First, the ionic interaction between the ligand guanidinium and Asp218 (on α_v) is, as expected, a side-

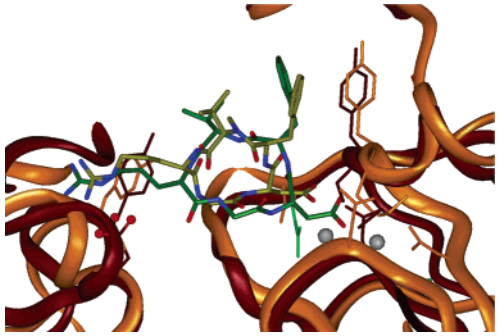


Figure 7. Overlay of modeled (**1**) and crystallized (**2**) ligand−$\alpha_v\beta_3$ complexes.

on binding. Second, the aspartic acid side chain fits in a short narrow pocket filled by a calcium cation. The two ionic interactions were expected because there were strict requirements in the design of $\alpha_v\beta_3$ antagonists. In contrast, the steric bulk at the Gly position and the hydrophobic $\pi-\pi$ stacking interaction were less predictable and were observed in our model. Even more gratifying was the prediction of the hydrogen bond interaction between the amide bond flanking the RGD sequence and the receptor. In addition, since the interacting side chain (Asn215 on β_3) contains both hydrogen bond donor and acceptor, this allows either hydrogen bond donor or acceptor on the ligand. Interestingly, after the simulation, the three-point pharmacophore binding mode still matches well with the proposed binding mode. Indeed, the triangle constituted by the three centers of the three accessibility volumes superposes the three crucial moieties of the docked ligand (Figure 6).

Recently, a structure of the extracellular portion of $\alpha_v\beta_3$ cocrystallized with cyclo(RGDfV(Me)) was elucidated.[46] It is worth mentioning that we visualized the reported X-ray crystal structure of the complex after the whole study was done, thus avoiding any user interference. Figure 7 shows an overlay of experimental and theoretical complexes. The main difference comes from the distorted metal-ion-dependent adhesion site (MIDAS region) (Mn^{2+} in the crystal structure and Ca^{2+} in the modeled structure). The additional methyl group of **2** compared to **1** was found to modify the conformation of the core. This explains the large deviation (rmsd > 5 Å) of the valine amino acid side chain compared to the good superposition of the ionic portion of the two peptides (rmsd = 2.5 Å). The good agreement between both structures validated the developed pharmacophore building and rigid docking approach.

This binding mode was top-ranking with the Autodock scoring function. This is of utmost interest because MD could process a single orientation and translation combination of the pharmacophore.

Directed Docking/Virtual Screening. This MD study revealed the high potential of the developed protocol and encouraged us to extend its scope to include virtual screening. The proposed binding mode of the pharmacophore (Figures 5 and 6) can be further exploited for virtual screening. Matching of antagonist interacting groups with the three pharmacophore points would restrict the number of poses for further scoring. This approach directs the docking to the postulated binding mode without recourse to a time-consuming conformational search. Most of the CPU time spent

Table 1. Comparative Study Showing the Scores

	score							
	apoprotein (PDB code: 1JV2)				holoprotein (PDB code: 1L5G)			
	protocol 20 runs	Autodock 50 runs	DOCK 20 runs	FlexX 50 runs	protocol 20 runs	Autodock 50 runs	DOCK 20 runs	FlexX 50 runs
1	3	2	1	0	3	2	3	1
2	3	3	2	0	3	2	3	0
3	3	1	3	0	0	3	3	3
4	0	0	0	0	3	2	3	1
5	2	0	2	0	3	2	0	1
6	2	1	2	0	2	3	2	1
7	3	0	3	3	3	3	3	3
8	2	3	2	3	0	0	0	0
total	18	10	11	6	17	17	17	10

during a docking run is used for docking or scoring irrelevant binding modes. The efficacy of this protocol relative to the rigid docking methods comes from the ligand flexibility accounted for in the pharmacophore generation. The pharmacophore generation and docking have to be done once with a limited number of antagonists, while the directed docking can be applied to libraries of molecules.

To carry out this directed docking-based virtual screening, the DOCK matching algorithm was used. DOCK uses spheres to model the binding site and then matches the sphere centers with atoms of the ligands. The docked three-point pharmacophore was used as a set of three sphere centers. This procedure was applied to the training set (Charts 1 and 2) and revealed a great accuracy. For the sake of comparison, Autodock,[44,45] DOCK,[13−16] and FlexX[47,48] methods were chosen (Table 1). DOCK, Autodock, and FlexX use different methods to dock the compound to the binding site and are therefore appropriate for comparison purposes. In addition, they all contain a minimization stage for optimizing the proposed binding mode. Autodock uses a genetic algorithm for optimizing the translation, rotation, and conformation of the ligands. The FlexX program uses a rapid incremental construction algorithm to assemble the ligand in the binding pocket, while the DOCK algorithm performs matching of points located within the binding site and ligand atoms. Autodock3.0 was previously found to effectively dock RGD-like antagonists to the holoprotein integrin[32] and was also found to be the most efficient in docking studies of matrix metalloproteinase (MMP) inhibitors.[49] FlexX can be combined with many scoring functions. Thus, the sampling procedure rather than the scoring function will be the crucial element. In this study, G_Score[50] (part of CScore) was found to be the most accurate for scoring the ligand binding modes. The comparative study data are summarized in Table 1. The pharmacophore was previously built assuming that these active molecules bind in a similar conformation/orientation/translation mode. This comparative study also assumes that each ligand (**1**−**8**) binds as compound **2** does.

A score (0−3) was attributed to the proposed binding modes according to the following rules:

(score 3) correct orientation (as defined in the reported crystal structure, interaction with Asp218, Ca^{2+}/Mn^{2+}, and Tyr122),

(score 2) both ionic interactions (with Asp218 and Ca^{2+}/Mn^{2+}),

(score 1) one ionic interaction (with either Asp218 or Ca^{2+}/Mn^{2+}) and one hydrophobic interaction (with Tyr122),

(score 0) wrong binding mode.

The docking of compound **2** (of known binding mode) was the most appealing (Figure 8). For instance, Autodock, FlexX, and DOCK failed to propose the correct binding mode to both the apoprotein and the holoprotein (Table 1). DOCK and Autodock failed to predict the hydrophobic interaction with Tyr122 in the holoprotein, while FlexX predicted wrong binding modes in both the holo and apoprotein. The docking of the set of known antagonists is in good agreement with the results from the Kessler and Merck groups. In many cases, interaction with Asp150 was preferred over interaction with Asp218. Moreover, the proposed binding mode did not allow optimal binding with the calcium ion. The slight induced fit observed for the bound structure properly positions the Tyr122 and Asp150 side chains for optimal binding with compound **2**. This receptor structure allowed more efficient docking of the whole series (Table 1). The efficacy of our protocol was similar with the apoprotein and the holoprotein, while the efficacy of the other procedures was significantly affected by the structure of the protein. It is clear from these data that the proposed protocol is more accurate and less sensitive to the protein adjustments. This is of prime importance when considering that the macromolecule structures used for virtual screening are held rigid and cannot fit to the screened ligands. The poor efficiency of FlexX is directly related to its incremental construction approach. In many case, FlexX correctly positioned the arginine side chain or the carboxylate but was not able to predict the second ionic interaction. In the crystal structure, both ionic interactions occur through distorted (poor alignment) conformations while FlexX proposed a perfect alignment for each ionic interaction.

Another consideration of importance in virtual screening methods is the time required for docking large libraries of compounds. A minimum of 10 runs were required to reach the convergence with the present protocol, DOCK, and Autodock methods. FlexX required 50 runs to converge. The times required by each method measured for 10 runs are summarized in Table 2. As can be seen in Table 2, the directed docking protocol was about 10 times faster than the other three methods while being more accurate. The accuracy and required time are clear indicators of the efficacy of a tool for virtual screening.

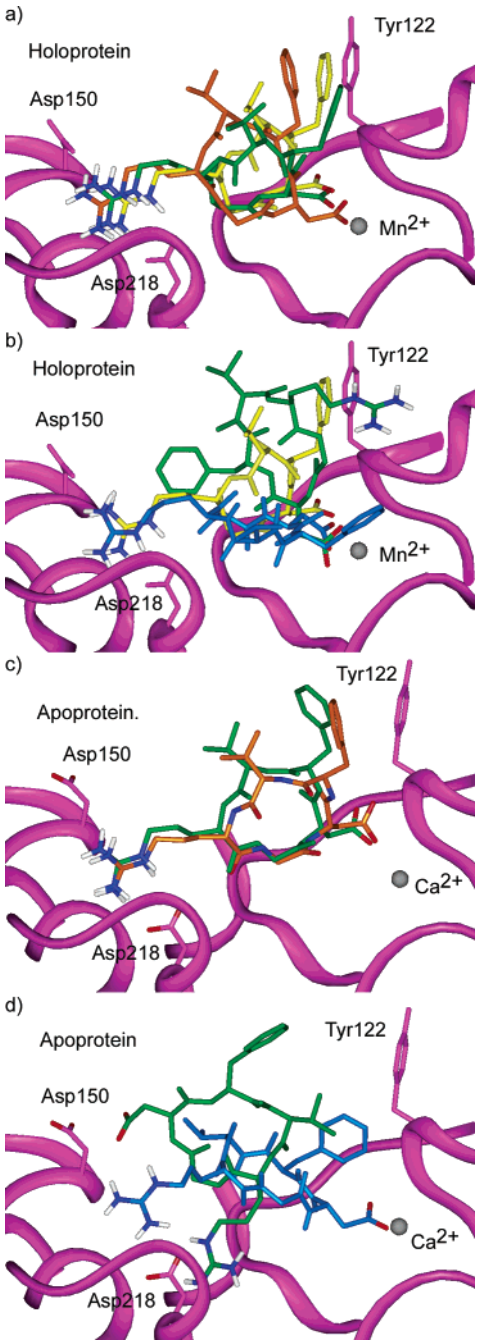


Figure 8. Docking of compound **2** (a) to the holoprotein using our protocol (orange) and DOCK (green) compared to the crystal structure (yellow), (b) to the holoprotein using FlexX (green) and Autodock (blue) compared to the crystal structure (yellow), (c) to the apoprotein using our protocol (orange) and DOCK (green), and (d) to the apoprotein using FlexX (green) and Autodock (blue).

Table 2. Comparative Study Showing Times Required for Docking of Compounds

	protocol	Autodock	DOCK	FlexX
time[a]	293	3298	1936	2058

[a] Average time calculated for the whole set (s/(compound/10 runs); 50 runs were performed for the convergence check). See Experimental Section for further details on the procedures.

It is clear that our protocol is an accurate and effective way of docking compounds to flexible proteins. This efficacy comes from the synergistic effects of a hybrid ligand-based drug design (pharmacophore determination) and structure-based drug design (docking) method. Although the usual docking methods exploit the available structure of the targeted protein, the pharmacophore development accounts for the flexibility of the receptor giving clues on the binging conformations. The pharmacophore docking is not affected by the modification of the steric environment between structures of a different source. Autodock, DOCK, and FlexX docking results are much more affected by the structures used for docking.

Conclusion

With the ever-increasing computational power, rigid body docking could be applied to libraries of potential ligands.[7−9] However, this particular approach implies the docking of a single conformation or docking of a library of predefined conformations. In the former case, the proposed conformation must be the bioactive one, while in the latter case, many conformations should be screened. Another strategy is to generate a large number of conformations. This approach is also computationally demanding and cannot be easily transferred to virtual screening. We reasoned that docking only the suspected bioactive conformation (defined by preliminary pharmacophore search) might be much less CPU-time-consuming. Thus, a potential pharmacophore might be constructed and then docked as a rigid body. If needed, extensive MD simulations can further refine the elucidated structure of the complex. This pharmacophore should therefore include the conformational degrees of freedom to account for the ligand flexibility, and MD simulation would account for the solvation and entropy contributions and for the receptor flexibility. This would ultimately lead to a novel, flexible QSAR/docking hybrid method. Indeed, during the preparation of this manuscript, Joseph-McCarthy and co-workers reported a similar method, which proved to be valuable.[51]

The method was applied to RGD-like compounds, which act as $\alpha_v\beta_3$ integrin antagonists. A three-point pharmacophore was developed and docked in the X-ray crystal structure of this receptor. Visual inspection and Autodock docking study proposed three alternative binding modes. At this stage, matching antagonists to bound pharmacophore led to model complexes. Further MD simulation refined the structures. The quick pharmacophore construction was exclusively based on geometric considerations and enabled us to account for the ligand flexibility. This approach was approved to be valuable for quick determination of the relative arrangement of the pharmacophoric groups. The subsequent rigid docking of the pharmacophore produced a template for further docking of any antagonist as exemplified with **1**. Autodock proved to be a great tool because the highest scoring binding mode of the pharmacophore was later found to be the experimentally observed one. However, Autodock along with the DOCK and FlexX programs was less effective than the developed protocol in properly docking compound **2** and most of the training set to both the apo and holoprotein. This was a further validation of the proposed method and a proof of its great accuracy in docking antagonists to receptor regardless of the exploited structure (the unbound and bound structures in this work). The time

needed is also critical when processing a large number of molecules. Our method was also found to be faster than the tested existing methods. Virtual screening of large libraries could now be envisioned on the basis of the proposed pharmacophore binding mode.[52]

Taken together, our pharmacophoric determination and docking experiment further our understanding of the structural requirement for binding to $\alpha_v\beta_3$ integrin. Such information would be useful for the design of new potential $\alpha_v\beta_3$ integrin antagonists.

Experimental Section

General. Molecular modeling was performed on Silicon Graphics Indigo 2 and Fuel workstations equipped with 250 MHz R10000 processors or 700 MHz R16000 running the IRIX operating system (version 6.5). The molecules were manipulated using InsightII version 2000[53] (Accelrys) and modeled using the InsightII/Discover package with CVFF as a force field (MD simulation). Structures for docking studies (Autodock, DOCK, and FlexX) were generated from Sybyl version 6.9.1[54] (Tripos Inc.). Heavy calculations were performed at CINES (Montpellier) on an O3000 SGI equipped with 512 R14 000 processors.

Pharmacophore Construction. The development of the pharmacophore is based on the pattern deduced from the SAR studies (Figure 2). The construction follows a stepwise protocol: (1) definition of the central core and analysis of its conformations, (2) installation of the side chains featuring the crucial moieties, and (3) study of the geometrically accessible spaces of these postulated interacting groups.

The peptides were built taking advantage of the reported structural data as constraints. For example, the main chain of **4** was built by adjusting the Φ and Ψ dihedral angles to the experimentally observed values (from NMR spectroscopy).[42] This procedure was reiterated for the other three peptides or pseudopeptide based on reported experimental data (**1**,[38,39,55] **2**,[40] **3**). We took particular attention to the intramolecular hydrogen bonds occurring for each structure. A total of 2000 steps of molecular dynamics simulation ($T = 300$ K) with these constraints followed by minimization using steepest descents and then conjugate gradient led to optimized structures. The side chains were next installed onto these central cores. Again, minimization was performed and led to the structures ready for the next steps.

A simulated annealing conformational search generated two different conformations of the bicyclic scaffolds of **5** and **7** due to a flip of the seven-membered rings. These core conformations were next considered separately. A complete exploration of the geometrical space for this class of nonpeptidic antagonists was thus achieved.

Geometrically accessible volumes were generated using a systematic conformational search available in the Search-_Compare module interfaced with InsightII, keeping the central core fixed. The volumes based on van der Waals radii were calculated by rotating all the rotatable bonds with an increment of 30° and a window of 360° except when symmetry rules apply (for example, a window of 180° was used for the carboxylate groups). Boolean computations merged the complementary volumes for core conformations of **5** and **7** (see above). At this stage, we had in hand eight sets of volumes. Manual alignment with concurrent optimization of the superposition of both the cores and the sets of volumes followed by Boolean computation of the overlapping volumes led to a final single set of three volumes. The centers of these three volumes defined a three-point pharmacophore.

Docking. The Cartesian coordinates of the extracellular segment of the free $\alpha_v\beta_3$ integrin and of the compound **2**$-\alpha_v\beta_3$ complex were retrieved from the Brookhaven Protein Databank (PDB).[35,46] PDB codes are the following: 1JV2, 1L5G. One Ca^{2+}, which was shown to play a role in the adhesion to RGD peptides,[56] was added in a known metal-ion-dependent adhesion site of the free binding site as defined by Arnaout and co-workers.[35] The hydrogen atoms were added and visually inspected.

Automated docking study with Autodock[44,45] led to a single binding mode of the three-point pharmacophore. Autodock is a fully automated docking suite of programs that employs a Lamarckian genetic algorithm (LGA) as a search engine and a LUDI-type scoring function.[57] Three-dimensional energy scoring grids of 0.375 Å resolution and 30 Å × 30 Å × 30 Å dimensions were computed. A total of 50 runs with a maximum of 250 000 energy evaluations and 100 runs with 100 000 energy evaluations were performed. The default parameters for the LGA and Sollis and Wet local search were used.[44,45] The three points of the pharmacophore were substituted by N, O, and A atom types with charges of +1, −1, and 0, respectively.

Alternatively, the pharmacophore was manually docked in the suspected binding site of the receptor by maximizing the electrostatic interactions. The calcium mentioned above was assumed to participate in the binding. Thus, the negatively charged center of the pharmacophore was positioned in a free coordination site of the cation. The presence of three geometrically accessible carboxylate moieties in the binding site resulted in three alternative binding sites for the positively charged moiety and consequently to three postulated binding orientations.

Molecular Dynamics Simulation. Antagonist **1** was matched to the three pharmacophore−receptor complexes resulting in three **1**−receptor complexes, which were then refined by a relaxation process in a stepwise manner. A water solvent layer of 15 Å around the complex was added, and the system was allowed to relax following the procedure described below. The protein and the ligand were held fixed, the solvent water molecules being free to move. A preliminary minimization was performed to remove close atom contacts by 100 000 cycles of minimization using conjugate gradients. At this point, the main chain of the protein was fixed and the complex was subjected to a further 100 000 cycles of minimization using conjugate gradients. The last 100 000 steps were reiterated with the whole system free to move until an energy gradient of less than 0.01 kcal/mol was achieved.

The complexes were next subjected to molecular dynamics simulation (300 K). However, the instability of the systems required repetitive intermediate minimization steps. Thus, seven loops constituted by 20 000 MD steps of 0.1 fs followed by 100 000 steps of minimization were performed until the energy converged to a stable value. One of the three resulting systems appeared to feature expected interactions with the receptor (as defined in Figure 2) and was further subjected to MD simulation. A 100 ps trajectory (steps of 1.0 fs) was recorded, and the final conformation was minimized.

Docking of the Data Set. DOCK 4.0. A Connolly surface was generated and used to generate a set of 50 spheres. Three-dimensional energy scoring and bump filter grids of 0.3 Å resolution and 30 Å × 30 Å × 30 Å dimensions were computed. The automated matching default parameters were used. A bump filter was used, and the flexible ligand orientation and conformation were relaxed (300 steps) and energy-scored.

FlexX1.9. Default parameters for the incremental construction were used, and serial scoring with CScore was performed. Fifty runs were carried out for each ligand−receptor complex. The binding site was generated by keeping all residues within 7 Å from Asp150, Asp218, Tyr122, and metal atoms.

Autodock3.0. The previously computed scoring grids were used. A maximum of 2 000 000 energy evaluations was used.

Acknowledgment. Financial support from Région Lorraine, Institut Nancéien de Chimie Moléculaire, and Association pour la Recherche contre le Cancer is gratefully acknowledged, and CINES is acknowledged for the computing facilities. Eric Therrien (Université de Montréal) is also greatly acknowledged for his assistance.

References

(1) *Protein−Ligand Interactions: From Molecular Recognition to Drug Design*; Böhm, H.-J., Schneider, G., Eds.; Wiley-VCH: Weinheim, Germany, 2003.

(2) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 409−443.

(3) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein−small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.

(4) Klebe, G. Recent developments in structure-based design. *J. Mol. Med.* **2000**, *78*, 269−281.

(5) Joseph-McCarthy, D. Computational approaches to structure-based ligand design. *Pharmacol. Ther.* **1999**, *84*, 179−191.

(6) Kubinyi, H. Structure-based design of enzyme inhibitors and receptor ligands. *Curr. Opin. Drug. Discovery Dev.* **1998**, *1*, 4−15.

(7) Good, A. C.; Krystek, S. R.; Mason, J. S. High-throughput and virtual screening: core lead discovery technologies move towards integration. *Drug Discovery Today* **2000**, *5*, S61−S69.

(8) Klebe, G., Ed. Virtual screening: an alternative or complement to HTS? *Perspect. Drug Discovery Des.* **2000**, *20*, 1−295.

(9) Bajorath, J. Integration of virtual and high-throughput screening. *Nat. Rev. Drug Discovery* **2002**, *1*, 882−894.

(10) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.

(11) Boobbyer, D. N. A.; Goodford, P. J.; McWhinnie, P. M.; Wade, R. C. New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. *J. Med. Chem.* **1989**, *32*, 1083−1094.

(12) Goodford, P. J. Drug design by the method of receptor fit. *J. Med. Chem.* **1984**, *27*, 557−564.

(13) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule−ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269−288.

(14) DesJarlais, R.; Sheridan, R. P.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. Docking flexible ligands to macromolecular receptors by molecular shape. *J. Med. Chem.* **1986**, *29*, 2149−2153.

(15) DesJarlais, R.; Sheridan, R. P.; Seibel, G. L.; Dixon, J. S.; Kuntz, I. D. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.* **1988**, *31*, 722−729.

(16) Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, *257*, 1078−1082.

(17) Tame, J. R. H. Scoring functions: a view from the bench. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 99−108.

(18) Böhm, H. J.; Klebe, G. What can we learn from molecular recognition in protein−ligand complexes for the design of new drugs? *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588−2614.

(19) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644−2676.

(20) Bursavich, M. G.; Rich, D. H. Designing non-peptide peptidomimetics in the 21st century: inhibitors targeting conformational ensembles. *J. Med. Chem.* **2002**, *45*, 541−558.

(21) Murray, C. W.; Baxter, C. A.; Frenkel, A. D. The sensitivity of the results of molecular docking to induced fit effects: Application to thrombin, thermolysin and neuramidase. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 547−562.

(22) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FlexX Incremental Construction Algorithm for Preotein−Ligand Docking. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 228−241.

(23) McGovern, S. L.; Shoichet, B. K. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J. Med. Chem.* **2003**, *46*, 2895−2907.

(24) Leach, A. R. Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.* **1994**, *235*, 345−356.

(25) Knegtel, R. M. A.; Kuntz, I. D.; Oshiro, C. M. Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **1997**, *266*, 424−440.

(26) Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T. FlexE: Efficient molecular docking considering protein structure variations. *J. Mol. Biol.* **2001**, *308*, 377−395.

(27) Osterberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in Autodock. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 34−40.

(28) Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.

(29) Feuston, B. P.; Cilberson, J. C.; Duggan, M. E.; Hartman, G. D.; Leu, C.-T.; Rodan, S. B. Binding model for nonpeptide antagonists of $\alpha_v\beta_3$ integrin. *J. Med. Chem.* **2002**, *45*, 5640−5648.

(30) Gottschalk, K.-E.; Gunther, R.; Kessler, H. A three-state mechanism of integrin activation and signal transduction for integrin $\alpha_v\beta_3$. *ChemBioChem* **2002**, *5*, 470−473.

(31) Gottschalk, K.-E.; Kessler, H. The structures of integrins and integrin−ligand complexes: implications for drug design and signal transduction. *Angew. Chem., Int. Ed.* **2002**, *41*, 3767−3774.

(32) Marinelli, L.; Lavecchia, A.; Gottschalk, K.-E.; Novellino, E.; Kessler, H. Docking studies on $\alpha_V\beta_3$ integrin ligands: pharmacophore refinement and implications for drug design. *J. Med Chem.* **2003**, *46*, 4393−4404.

(33) Henry, C.; Moitessier, N.; Chapleur, Y. Vitronectin receptor−$\alpha_v\beta_3$ integrin−antagonists: chemical and structural requirements for activity and selectivity. *Mini-Rev. Med. Chem.* **2002**, *2*, 531−542.

(34) Miller, W. H.; Keenan, R. M.; Willette, R. N.; Lark, M. W. Identification and in vivo efficacy of small-molecule antagonists of integrin $\alpha_v\beta_3$ (the vitronectin receptor). *Drug Discovery Today* **2000**, *5*, 397−408.

(35) Xiong, J.-P.; Stehle, T.; Diefenbach, B.; Zhang, R.; Dunker, R.; Scott, D. L.; Joachimiak, A.; Goodman, S. L.; Arnaout, M. A. Crystal structure of the extracellular segment of integrin $\alpha_v\beta_3$. *Science* **2001**, *294*, 339−345.

(36) Minoux, H.; Moitessier, N.; Chapleur, Y.; Maigret, B. Elucidation of a common structure of selective fibrinogen receptor antagonists. *J. Comput-Aided Mol. Des.* **1998**, *12*, 533−542.

(37) Moitessier, N.; Dufour, S.; Chrétien, F.; Thiery, J.-P.; Maigret, B.; Chapleur, Y. Design, synthesis and preliminary biological evaluation of a focused combinatorial library of stereodiverse carbohydrate-scaffold-based peptidomimetics. *Bioorg. Med. Chem.* **2001**, *9*, 511−523.

(38) Müller, G.; Gurrath, M.; Kessler, H. Dynamic forcing, an aid for the determination of the activity and selectivity of RGD (Arg-Gly-Asp) peptides. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 341−343.

(39) Haubner, R.; Gratias, R.; Diefenbach, B.; Goodman, S. L.; Jonczyk, A.; Kessler, H. Structural and functional aspects of RGD-containing cyclic pentapeptides as highly potent and selective integrin $\alpha_v\beta_3$ antagonists. *J. Am. Chem. Soc.* **1996**, *118*, 7461−7472.

(40) Dechantstreiter, M. A.; Planker, E.; Matha, B.; Lohof, E.; Holzemann, G.; Jonczyk, A.; Goodman, S. L.; Kessler, H. N-Methylated cyclic RGD peptides as highly active and selective $\alpha_v\beta_3$ integrin antagonists. *J. Med. Chem.* **1999**, *42*, 3033−3040.

(41) Bach, A. C., II; Espina, J. R.; Jackson, S. A.; Stouten, P. F. W.; Duke, J. L.; Mousa, S. A.; DeGrado, W. F. Type II′ to type I β-turn swap changes specificity for integrins. *J. Am. Chem. Soc.* **1996**, *118*, 293−294.

(42) Burgess, K.; Lim, D.; Mousa, S. A. Synthesis and solution conformation of cyclo(RGDRGD): a cyclic peptide with selectivity for the $\alpha_v\beta_3$ receptor. *J. Med. Chem.* **1996**, *39*, 4520−4526.

(43) Pickett, S. D.; McLay, I. M.; Clark, D. E. Enhancing the hit-to-lead properties of lead optimization libraries. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 263−272.

(44) Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins: Struct., Funct., Genet.* **1990**, *8*, 195−202.

(45) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(46) Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. Crystal structure of the extracellular segment of integrin $\alpha_v\beta_3$ in complex with an Arg-Gly-Asp ligand. *Science* **2002**, *296*, 151−155.

(47) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(48) Rarey, M.; Kramer, B.; Lengauer, T. Multiple automatic base selection: Protein−ligand docking based on incremental construction without manual intervention. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 369−384.

(49) Hanessian, S.; Moitessier, N.; Therrien, E. A comparative docking study and the design of potentially selective MMP inhibitors. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 873−881.

(50) Development and validation of a genetic algorithm for flexible docking. Jones, G.; Willett, P.; Glen, R.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.

(51) Joseph-McCarthy, D.; Thomas, B. E., IV; Berlmarsh, M.; Moustaka, D.; Alvares, J. C. Pharmacophore-based docking to account for ligand flexibility. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 172−188.

(52) Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129−149.

(53) *InsightII 2000*, Molecular Simulations Inc.: San Diego, CA, 1995.

(54) *Sybyl 6.9.1*, Tripos Inc.: St. Louis, MO, 2003.

(55) Gurrath, M.; Müller, G.; Kessler, H.; Aumailley, M.; Timpl, R. Conformation/activity studies of rationally designed potent anti-adhesive RGD peptides. *Eur. J. Biochem*. **1992**, *210*, 911−921.
(56) Tozer, E. C.; Liddington, R. C.; Sutcliffe, M. J.; Smeeton, A. H.; Loftus, J. C. Ligand binding to integrin $\alpha_{IIb}\beta_3$ is dependent on a MIDAS-like domain of the β3 subunit. *J. Biol. Chem*. **1996**, *271*, 21978−21984.
(57) Böhm, H.-J. The development of a simple empirical scoring function to estimate the binding constant for a protein−ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des*. **1994**, *8*, 243−256.

JM0311386