# Insights into Phenylalanine Derivatives Recognition of VLA-4 Integrin: From a Pharmacophoric Study to 3D-QSAR and Molecular Docking Analyses

Antonio Macchiarulo,[§] Gabriele Costantino,[§] Mirco Meniconi,[§] Karin Pleban,[‡] Gerhard Ecker,[‡] Daniele Bellocchi,[§] and Roberto Pellicciari*[,§]

Dipartimento di Chimica e Tecnologia del Farmaco, Università di Perugia, via del Liceo 1, 06127 Perugia, Italy, and Institut für Pharmazeutische Chemie, Universität Wien, Pharmaziezentrum, Althanstrasse 14, A-1090 Wien, Austria

Received March 5, 2004

The very late antigen-4 (VLA-4), also known as integrin $\alpha 4\beta 1$, is expressed on monocytes, T- and B-lympohocytes, basophils, and eosinophils and is involved in the massive recruitment of granulocytes in different pathological conditions such as multiple sclerosis and asthma. VLA-4 interacts with its endogenous ligand VCAM-1 during chronic inflammation, and blockade of VLA-4 /VCAM-1 interaction is a potential target for immunosuppression. Two classes of VLA-4 antagonists have so far been reported: $\beta$-amino acid derivatives containing a diaryl urea moiety (BIO-1211) and phenylalanine derivatives (TR-14035). With the aim of clarifying the structural basis responsible for VLA-4 recognition by phenylalanine derivatives, we developed a combined computational study on a set of 128 antagonists available through the literature. Our computational approach is composed of three parts. (i) A VCAM-1 based pharmacophore was constructed with a restricted number of phenylalanine derivatives to identify the region of the protein that resembles synthetic antagonists. The pharmacophore was instrumental in constructing an alignment of a set of 128 compounds. This alignment was exploited to build a pseudoreceptor model with the RECEPTOR program. (ii) 3D-QSAR analysis was carried out on the computed electrostatic and steric interaction energies with the pseudoreceptor surface. The 3D-QSAR analysis yielded a predictive model able to explain much of the variance of the 128 antagonists. (iii) A homology modeling study of the headpiece of VLA-4 based on the crystal structure of $\alpha v\beta 3$ was performed. Docking experiments of TR-14035 into the binding site of VLA-4 aided the interpretation of the 3D-QSAR model. The obtained results will be fruitful for the design of new potent and selective antagonists of VLA-4.

## INTRODUCTION

Integrins are a family of cell surface receptors belonging to the cell-adhesion proteins superfamily which comprises Cadherins, Selectins, and the ImunoGlobulin (Ig) family.[1]

Integrins are constituted by the hetero-assembling of two subunits, namely alpha and beta subunits.

Each subunit contains a large extracellular region, a single transmembrane domain, and a short cytoplasmic domain.

Eighteen $\alpha$ subunits and eight $\beta$ subunits have been identified, and they associate to form at least 20 integrins. These are subclassified into eight groups on the basis of structurally distinct $\beta$-chains, and each group is endowed by diverse ligand specificity.[2]

Crystallization studies[3] and nuclear magnetic resonance analysis[4] of the extracellular region of $\alpha v\beta 3$ integrin contributed in the understanding of the structure and molecular functioning of this class of proteins.

Furthermore, the availability of the X-ray structures of apo-$\alpha v\beta 3$[3] and $\alpha v\beta 3$ in complex with an Arg-Gly-Asp containing ligand[5] have prompted structure-based studies to gain insight into recognition and rational design of integrin ligands.[6−8]

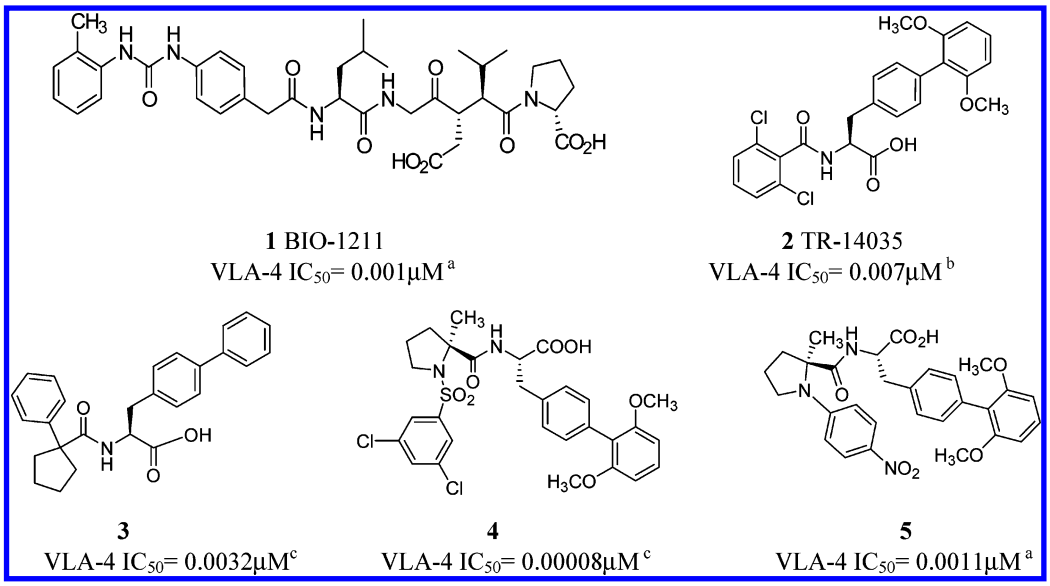
Integrins exist in an equilibrium between three conformations: a bent conformation with closed headpiece, an extended form with closed headpiece, and a second extended form with open headpiece. Each of these conformations of the extracellular region corresponds to a different conformation of the cytoplasmic tails. In particular, the cytoplasmic tails of the $\alpha$ and $\beta$ subunit are associated in the bent conformation of the extracellular region where the transmembrane domains are close. Conversely, they are dissociated in both the closed and the open extended conformations of the integrin headpiece where the transmembrane domains are separated. This means that there is a bidirectional conformational influence between the headpiece and the cytoplasmic tail of integrins. The binding of endogenous ligands favors the extended open headpiece conformation determining the activation of integrins. Similarly, cytosolic signals inducing the dissociation of the $\alpha$ and $\beta$ subunit tails promote the separation of the transmembrane domains and shift the equilibrium of the headpiece toward the open extended active conformation. This mechanism is at the basis of the bidirectional signal transmission across the membrane by integrins.[9−13]

$\beta 1$ integrins (very late antigen, VLA, family) are known to function in a dual mode, both as adhesion and signaling molecules.[14] The VLA family is comprised of six members, widely distributed among tissues and interacts principally

* Corresponding author phone: ++39 075 585 5120; fax: ++39 075 585 5124; e-mail: rp@unipg.it.
‡ Universität Wien.
§ Università di Perugia.

10.1021/ci049914l CCC: $27.50 
Published on Web 07/22/2004

**Chart 1**




[a] Inhibition of VCAM-Ig binding. [b] Jurkat cell adhesion assay. [c] Inhibition of radiolabeled VCAM-1-Ig binding.

with extracellular matrix proteins. Of particular interest is VLA-4 ($\alpha 4\beta 1$), which is expressed on monocytes, T- and B-lympohocytes, basophils, and eosinophils and binds the alternatively spliced type-III connecting-segment (CS-1) of fibronectin and the vascular cell adhesion molecule-1 (VCAM-1).[15] It provides a mechanism for selective recruitment of granulocytes during an inflammatory response thus representing an attractive target for immunosuppression in different pathological conditions. Indeed, the involvement of VLA-4 has been proven in diseases such as asthma,[16,17] multiple sclerosis,[18] rheumatoid arthritis,[19] and diabetes.[20,21] By peptide mapping studies, Komoriya et al. identified a tripeptide, Leu-Asp-Val, as the minimum sequence motif of fibronectin needed to bind VLA-4.[22]

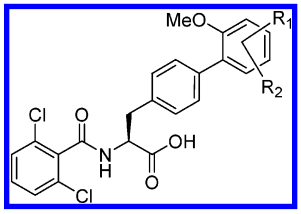
This pattern of three residues is homologous to the Leu-Asp-Ser motif found in VCAM-1 which has been reported as the VCAM-1 binding site of VLA-4,[23] thus providing that VCAM-1 and fibronectin could share an identical or overlapping VLA-4 binding site.[24]

The crystal structure of VCAM-1 has revealed that the Leu-Asp-Ser motif forms a surface loop exposed to the solvent and suitable for the interaction with VLA-4.[25,26]

Both VLA-4 binding motifs of VCAM-1 and fibronectin were used to design the majority of VLA-4 antagonists.[27] Nevertheless, it is worth to note that cyclic peptide antagonists were also derived from the Arg-Gly-Asp tripeptide, another region of fibronectin that binds other integrins such as $\alpha 5\beta 1$ and $\alpha IIb\beta 3$,[28] and that peptides derived from the complementary-determining region-3 of antibody 21/6 (an anti-$\alpha 4$ mAb)[29] were reported as competitive inhibitors of VLA-4 interaction at low micromolar concentration. Among noncyclic antagonists based on the Leu-Asp-Val motif so far reported, two distinct groups can be identified: (a) $\beta$-amino acid derivatives containing a diaryl urea moiety (BIO-1211, **1**, Chart 1)[30] and (b) phenylalanine derivatives (TR-14035, **2**).[31−34]

No experimental evidence has yet been reported that these two classes of compounds could share the same binding site on VLA-4.

**Table 1.**



| compound | R1 | R2 | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|
| **2**(TR-14035) | 6-OMe | H | 0.007 |
| **6** | 3-OMe | H | 0.13 |
| **7** | 4-OMe | H | 0.1 |
| **8** | 5-OMe | H | 0.017 |
| **9** | 6-OMe | 4-OMe | 0.008 |
| **10** | 6-OMe | 4-OH | 0.017 |
| **11** | 6-OMe | 3-OH | 0.042 |
| **12** | 6-OMe | 3-OMe | 0.010 |
| **13** | 6-OMe | 3-Br | 0.018 |
| **14** | 6-OMe | 3-F | 0.022 |
| **15** | 6-OMe | 3,5-diF | 0.071 |
| **16** | 6-OMe | 3-$CONH_2$ | 0.010 |
| **17** | 6-OMe | 3-NHCOMe | 0.020 |
| **18** | 6-OMe | 3-NHCONHMe | 0.020 |

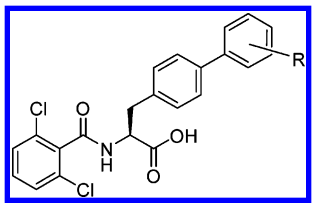
While pharmacophoric models[35,36] and CoMFA studies[37] have been reported for selected series of derivatives belonging to group-a, to our knowledge there has been no attempt to quantitatively study the structure−activity relationship of group-b derivatives.

Using the crystal structure of VCAM-1[25] coupled to the available mutagenesis experiments,[38] we built a pharmacophore model for a small set of compounds of group-b (**2**−**5**). The pharmacophore model was used as a starting point to align 128 phenylalanine derivatives[31] (Tables 1−9) and construct a pseudoreceptor model of VLA-4. Finally, a 3D-QSAR model was derived and discussed in light of recent achievements obtained with the crystallization of the $\alpha v\beta 3$ integrin in complex with an Arg-Gly-Asp containing ligand.[5]

## METHODS

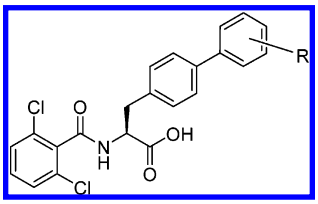
All compounds of this study were constructed within Cerius2 software package,[39] and their geometries optimized

**Table 2.**




| compound | R | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|
| **19** | H | 2.3 |
| **20** | 3-OMe | 6.2 |
| **21** | 4-OMe | 1.1 |
| **22** | 2-$OCF_3$ | 0.42 |
| **23** | 2-$O^{iso}Pr$ | 1.17 |
| **24** | 2-OBu | 1.6 |
| **25** | 2-$O^{iso}Bu$ | 2.1 |
| **26** | 2-Me | 0.75 |
| **27** | 2-$CF_3$ | 0.13 |
| **28** | 2-OH | 0.33 |
| **29** | 2-CN | 0.13 |
| **30** | 2-CHO | 0.16 |
| **31** | 2-$CH_2OH$ | 0.14 |
| **32** | 2-$CH_2OMe$ | 0.34 |
| **33** | 2-$CH_2NH_2$ | 0.30 |
| **34** | 2-$CH_2NHPh$ | 0.28 |
| **35** | 2-$CH_2NHCH_2Ph$ | 1.66 |
| **36** | 2-COOH | 6.14 |
| **37** | 2-$CONH_2$ | 0.51 |
| **38** | 2-CONHMe | 0.79 |
| **39** | 2-$CONMe_2$ | 0.79 |

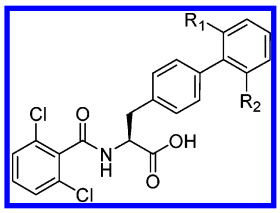
**Table 3.**




| compound | R | $\alpha_4\beta_1$-(VLA4) $C_{50}$ $\mu$M |
|---|---|---|
| **40** | 2-CONHBu | 1.44 |
| **41** | 2-$CONHCH_2Ph$ | 1.93 |
| **42** | 2-$SO_2NH^tBu$ | 0.15 |
| **43** | 2-$SO_2NH_2$ | 0.10 |
| **44** | 2-$SO_2NHCOPh$ | 0.35 |
| **45** | 2-$SO_2NHCOMe$ | 1.6 |
| **46** | 2-$NHCOO^tBu$ | 2.7 |
| **47** | 2-$NH_2$ | 0.45 |
| **48** | 2-$NHCONH_2$ | 0.40 |
| **49** | 2-$NHSO_2Me$ | 0.27 |
| **50** | 2-NHCOMe | 0.78 |
| **51** | 2-NHCOOMe | 0.57 |
| **52** | 2-$NMe_2$ | 0.20 |

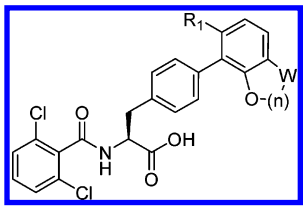
using the UNIVERSAL force field v.1.2[40] with the Smart Minimizer protocol of Open Force Field (OFF). Partial atomic charges were computed using the semiempirical MOPAC/AM1 method. Conformational analyses were carried out using the Grid Search routine with an increment value for torsional angles of 30 degree in a range from −180 to 150 degrees. The crystal structure of VCAM-1 was taken from the protein databank (PDB code: 1VCA). The pharmacophore points were manually defined by the atomic coordinates of the carboxylic oxygen atoms of Asp40 (Leu-Asp-Ser motif) and the centroid of the Leu70 side chain.

**Table 4.**



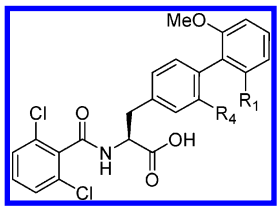

| compound | R1 | R2 | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|
| **53** | OMe | OEt | 0.027 |
| **54** | OMe | OH | 0.031 |
| **55** | OMe | $O(CH_2)_2F$ | 0.039 |
| **56** | OMe | $O^{iso}Pr$ | 0.245 |
| **57** | OEt | OEt | 0.16 |
| **58** | OH | OH | 0.5 |
| **59** | $OCH_2CH_2OH$ | $OCH_2CH_2OH$ | 0.282 |
| **60** | $OCH_2CF_3$ | $OCH_2CF_3$ | 0.676 |

**Table 5.**




| compound | R1 | W | *n* | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|---|
| **61** | H | O | 1 | 0.928 |
| **62** | OMe | O | 1 | 0.013 |
| **63** | $OCH_2OMe$ | O | 1 | 0.068 |
| **64** | $OCH_2CN$ | O | 1 | 0.069 |
| **65** | OH | O | 1 | 0.137 |
| **66** | $OCH_2CH_3$ | O | 1 | 0.109 |
| **67** | $OCH_2CH_2OH$ | O | 1 | 0.065 |
| **68** | $OCH_2CH_2OMe$ | O | 1 | 0.45 |
| **69** | $O(CH_2)_2NMe_2$ | O | 1 | 0.225 |
| **70** | OMe | O | 2 | 0.006 |
| **71** | H | $CH_2$ | 1 | 0.233 |
| **72** | H | $CH_2$ | 2 | 0.07 |

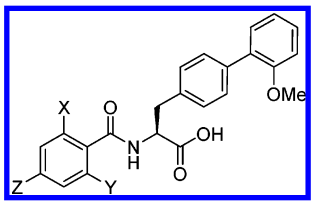
**Table 6.**




| compound | R4 | R1 | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|
| **73** | OMe | H | 0.040 |
| **74** | CH(OH)Me | H | 0.027 |
| **75** | CH(OH)Et | H | 0.047 |
| **76** | CH(OH)Me | OMe | 0.008 |
| **77** | $CH_2Me$ | H | 0.21 |
| **78** | NHCOMe | OMe | 0.030 |
| **79** | NHCOOEt | OMe | 0.123 |
| **80** | $NHSO_2Me$ | OMe | 0.010 |

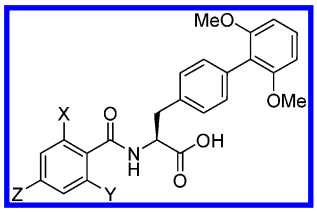
The conformations of compounds **2**−**5** were superimposed on VCAM-1 by fitting their carboxylic group and the centroid of the aromatic ring of the phenylalanine unit onto the pharmacophoric points. The conformation with the best fit and lowest energy resulting from the superimposition

**Table 7.**




| compound | X | Y | Z | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|---|
| **81** | H | H | H | 5.08 |
| **82** | Cl | H | H | 0.56 |
| **83** | Br | H | H | 0.77 |
| **84** | $CF_3$ | H | H | 1.1 |
| **85** | Cl | H | OMe | 0.3 |
| **86** | Cl | H | Cl | 0.31 |
| **87** | Cl | H | Br | 1.4 |
| **88** | Cl | H | $NO_2$ | 0.27 |
| **89** | Cl | H | $NH_2$ | 0.31 |
| **90** | Cl | H | $SO_2Me$ | 0.15 |
| **91** | Cl | H | OH | 0.14 |
| **92** | Cl | H | $CO_2{}^tBu$ | 0.46 |
| **93** | Cl | H | $CO_2H$ | 0.015 |
| **94** | Cl | H | $CONH_2$ | 0.12 |
| **95** | Cl | H | $CONHSO_2Me$ | 0.060 |
| **96** | Cl | Cl | H | 0.077 |
| **97** | Cl | F | H | 0.17 |
| **98** | F | F | H | 0.33 |
| **99** | Me | Me | H | 0.34 |

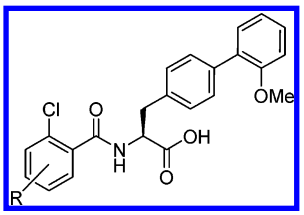
**Table 8.**




| compound | X | Y | Z | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|---|---|
| **100** | Br | Br | H | 0.046 |
| **101** | F | F | H | 0.059 |
| **102** | Cl | Me | H | 0.018 |
| **103** | Cl | Cl | Br | 0.029 |
| **104** | Cl | Cl | OH | 0.018 |
| **105** | Cl | Cl | CN | 0.013 |
| **106** | Cl | Cl | $CONH_2$ | 0.0025 |
| **107** | Cl | Cl | $CO_2H$ | 0.00033 |
| **108** | Cl | Cl | $NH_2$ | 0.007 |
| **109** | Cl | Cl | NHCOMe | 0.014 |

protocol was stored. The alignment of all remaining compounds was performed over their respective structurally related parent compound with the Align module and using the common subgraph match routine. The pseudoreceptor model was generated using the Receptor approach.[41,42] Briefly, the pseudoreceptor model represents the essential features of the binding site by assuming complementarity between the shape and properties of the binding pocket of the putative receptor and the set of VLA-4 antagonists. 3D surfaces of the binding site enclose the most active members (after appropriate alignment) of the starting set of compounds and represent the pseudoreceptor. The surface is generated from a "Shape Field". The atomic coordinates of the contributing molecules are used to compute field values on

**Table 9.**




| compound | R | $\alpha_4\beta_1$-(VLA4) $IC_{50}$ $\mu$M |
|---|---|---|
| **110** | 4-NHMe | 0.49 |
| **111** | 4-NHCOOEt | 0.20 |
| **112** | 4-$NHCONH_2$ | 0.14 |
| **113** | 4-$NHCSNH_2$ | 0.14 |
| **114** | 4-NHCSNHMe | 0.030 |
| **115** | 4-NHCSNHEt | 0.12 |
| **116** | 4-$NHCSNH^{iso}Pr$ | 0.32 |
| **117** | 4-NHCSNHPh | 0.31 |
| **118** | 4-NHCSNHBn | 1.5 |
| **119** | 4-$NHSO_2Me$ | 0.041 |
| **120** | 4-$NMeSO_2Me$ | 0.20 |
| **121** | 4-$NHSO_2NMe_2$ | 0.052 |
| **122** | 4-$NHSO_2CF_3$ | 0.013 |
| **123** | 4-$NHSO_2Ph$ | 0.021 |
| **124** | 4-$NHSO_2$(4-FPh) | 0.049 |
| **125** | 4-$NHSO_2$(4-MePh) | 0.073 |
| **126** | 4-$NHSO_2$(4-$CF_3$Ph) | 0.15 |
| **127** | 4-$NHSO_2$(4-OMePh) | 0.084 |
| **128** | 4-$NHSO_2$(2-MePh) | 0.12 |
| **129** | 4-$NHSO_2$(2-thienyl) | 0.011 |
| **130** | 3-$NHSO_2CF_3$ | 0.361 |
| **131** | 5-$NHSO_2CF_3$ | 0.060 |
| **132** | 6-$NHSO_2CF_3$ | 20 |

each point of a 3D grid using a van der Waals function (eq 1)

$$V(r) = r - \mathrm{VdWr}$$

where $r$ is the distance from the atomic coordinate to the grid point (called *surface fit* and set here to 0.1) and VDWr is the van der Waals radius of the atom.

The contribution of compounds for surface generation was equally weighted. Thus the generated pseudoreceptor was characterized by a volume of 2665 $Å^3$ and a surface of 1531 $Å^2$.

The interaction energies of aligned compounds with the pseudoreceptor were calculated for all points constituting the surface. The interaction at each point represents an independent variable, and a total of 738 variables were calculated. The 3D-QSAR analysis was carried out by using an in-house genetic algorithm based partial least-squares program (MiPLS). MiPLS includes the following capabilities: pretreatment of data (autoscaling, deletion of variable with SD close to zero); correlation analysis between variables (PCA analysis and plot of results); PLS analysis and plot of results; external predictions; scrambling test in PLS analysis; and variable selection via Genetic Algorithm in two different ways: (a) one-pot GAs variable selection and (b) two-phases GAs variable selection.[43]

One-pot GAs variable selection consists of stopping the genetic evolution of models when the chosen number of chromosomes is reached and performing a statistical analysis on the variables constituting the final chromosomes. Variables are then selected and added one by one into following generated models. The model with the highest predictivity is then chosen.

The first part of the two-phases GAs variable selection consists of performing a given number of runs on the original data set and on a variable scrambled data sets. The fitness function (e.g: $Q^2$ or $r_{cv}^2$) is evaluated for both the data sets, and the stop criterion is defined when the highest response is obtained for the "true" data set. A high number of cycles (number of cycles = 100) is then carried out with the stop criterion as above-defined. A statistical analysis of the occurrence of individual variables in the fittest chromosome for each of the 100 cycles is performed. In the second part of the two-phases GAs variable selection, the variables to be included into the analysis are chosen as the ones with the highest occurrence and a partial least-squares analysis is carried out to generate the final model. For our purpose we used the two-phases GAs variable selection. In particular, the following parameters for the first phase were defined: initial population size (ga_popol_ini) = 30; chromosomes generated at each cycle = 370; number of variable selection in the scrambled data set = 25; number of variable selection in the "true" data set = 25. The stop criteria of genetic evolution is given after the generation of 45 chromosomes. This value was chosen to improve the signal/noise ratio in a reasonable cost of CPU time. The cross-validation protocol for all analysis was performed with both the leave-one-out and 5 groups definition procedures and a number of 30 randomization with 5 maximum number of principal components.

The homology model of the headpiece of VLA-4 was performed using Modeler[44] and Charmm[45] modules of Insight-II software package.[39] As a template we used the headpiece of integrin $\alpha v\beta 3$ in complex with an RGD ligand (PDB code: 1L5G) to ensure the active conformation of the binding site for further docking experiments. The alignment was performed employing the "Align123" module of Insight II on the sequences of the N-terminal domain of $\alpha 4$ and $\alpha v$ subunits and the I-like domain of $\beta 1$ and $\beta 3$ subunits of human integrins (retrieved from the swiss-prot database). The Blosum matrix was used with a gap penalty of 11 and gap extension penalty of 1. Furthermore the alignment was checked in order to avoid gaps in conserved secondary structures and loops were manually refined. The thus derived homology model was submitted to minimization protocols using Charmm force-field. Prior to the minimization the atomic coordinates of manganese ($Mn^{2+}$) of the MIDAS motif were merged into the binding site as present in the crystal structure. The energy minimization was performed with the peptide ligand in the binding site to avoid its closure until reaching a gradient of 0.05 kcal/mol.

Docking experiments were performed using Autodock version 3.0.5.[46] Briefly, 120 runs were carried out using the genetic algorithm with a population size of 50 individuals and 2 500 000 energy evaluations. If not specified, other parameters were left to their respective default values. The search was conducted in a grid of 80 points *per* dimension and a step size of 0.375 Å centered on the coordinates of the $Mn^{2+}$ atom. Results were clustered and ranked in terms of binding energy and occurrence for a given docked conformation. All calculations were carried out on a SGI O2 R12000 machine.

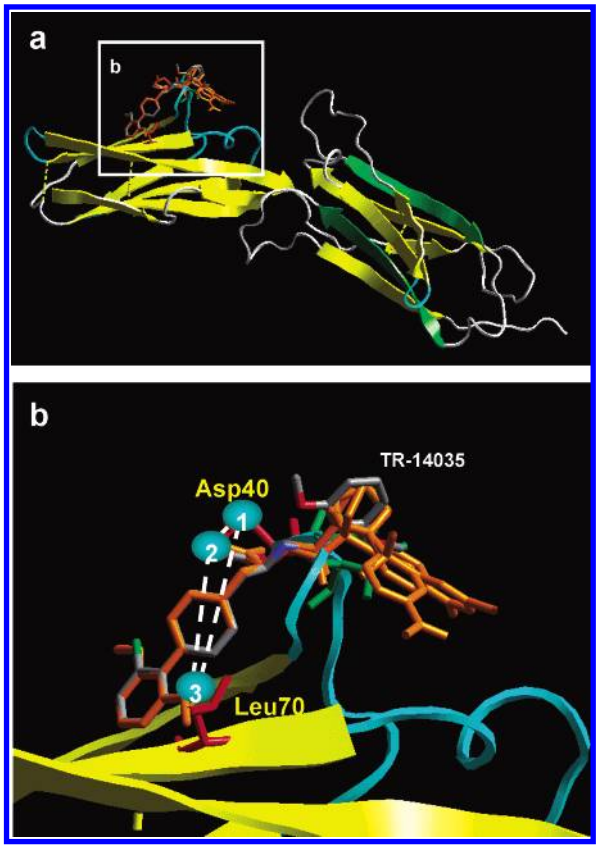



**Figure 1.** (a) Superposition of VCAM-1 and phenylalanine derivatives (**2**−**5**) on the basis of the pharmacophoric points. (b) Close-up view of VCAM-1 based pharmacophore.

## RESULTS

**Pharmacophore Model of Phenylalanine Derivatives Based on VCAM-1.** VCAM-1 is the endogenous ligand of VLA-4. It binds the extracellular domain of the integrin by means of three residues, Leu39-Asp40-Ser41, which forms a surface loop exposed to the solvent. Furthermore, site-directed mutagenesis pinpointed a key role for a residue of leucine (Leu70), adjacently located on domain 1 of VCAM-1, in the interaction with VLA-4.

Thus, we defined a protein-based pharmacophore model by taking into account the physicochemical and spatial relationship of these four residues.

In particular, the pharmacophore points were constituted by the atomic coordinates of the carboxylic oxygen atoms of Asp40 (Leu-Asp-Ser motif) and the centroid of the Leu70 side chain.

Four phenylalanine derivatives (**2**−**5**) were selected through the literature, and their conformations were fitted on VCAM-1 superimposing their carboxylic group and centroid of the aromatic ring of the phenylalanine unit over the pharmacophoric points (Figure 1).

The conformation with the best fit and lowest energy was stored and used to align a set of 128 VLA-4 antagonists tested for VLA-4 activity in RPMI and Jurkat cell adhesion assays.[31]

**Pseudoreceptor Construction and 3D-QSAR Analysis.** The alignment of the 128 compounds (Tables 1−9) based on VCAM-1 pharmacophore was instrumental to generate a pseudoreceptor model of the VLA-4 binding site (Figure 2).

The pseudoreceptor model is intended to provide a continuous description of the chemical environment of the

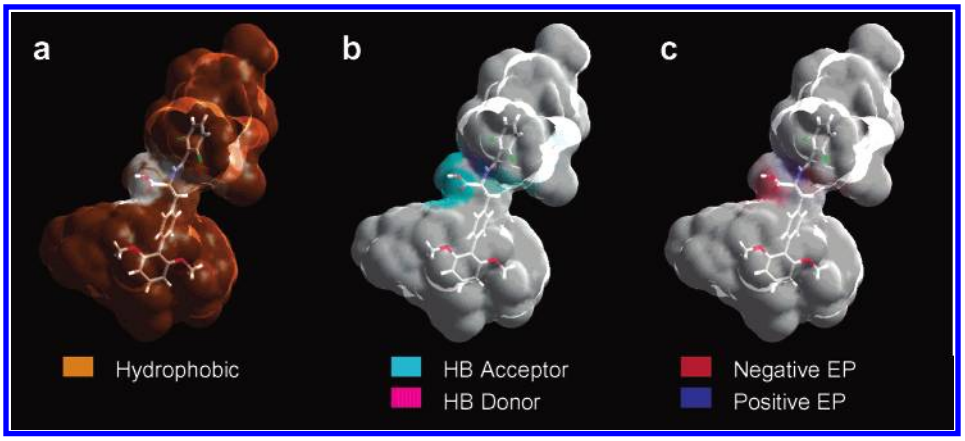



**Figure 2.** Pseudoreceptor model of VLA-4. (a) Mapping of hydrophobic interactions (brown); (b) mapping of hydrogen bond acceptor (cyan) and donor (purple) interactions; (c) mapping of negative (red) and positive (blue) electrostatic potential interactions.

binding pocket of VLA-4. Pseudoreceptor models differ from pharmacophore models in that they do not have any spatial arrangement of atoms or functional groups common to the set of active molecules. Conversely, they represent the essential features of an active site by assuming complementarity between the shape and properties of the receptor site and the bioactive conformations of the set of active compounds. It should be mentioned that errors in alignment can lead to incorrect, poorly predictive receptor surface models. Another possible caveat could be ought to the method of calculating atomic charges. Indeed, partial atomic charges directly affect the charge and electrostatic properties of the pseudoreceptor model and, in turn, the interaction energies. Herein we used the semiempirical MOPAC/AM1 method since its level of accuracy was good enough for our purpose. However, the influence of different calculation methods of atomic charges on the values of interaction energies was not investigated because out of the scope of this work.

Thus, electrostatic, hydrophobic, and hydrogen bond interaction features of the pseudoreceptor model were used to calculate the interaction energies of all aligned compounds with the binding site. In particular, the interaction energies were calculated for all points constituting the surface of the pseudoreceptor model. The values of the interaction at each point were considered as energetical descriptors and thus a total of 738 descriptors were calculated.

A 3D-QSAR study was carried out on 128 objects and 739 variables (738 descriptors plus the inhibitory activity of compounds as resulted from the RPMI and Jurkat cell adhesion assays). The whole set of 128 compounds and 738 energetical descriptors was submitted to a principal component analysis (PCA) to highlight possible outliers or the presence of subgroups.

Results indicate that the variance of the independent variable is explained mainly by the first component (90%). From a close inspection of the PCA plot (Figure 3a), compound **132** results a possible candidate to be an outlier. Indeed, this compound is the only one in the set that is characterized by a trifluoromethylsulfonyl substituent on the *ortho* position of one benzene ring. Moreover, compounds **22**, **81**, **84**, **85**, and **91** seem to constitute an independent subgroup on the first component. At this stage of the analysis only compound **132** was eliminated from the initial set. Thus 127 objects and 738 variables were used to perform a partial least-squares (PLS) analysis and genetic variable selection

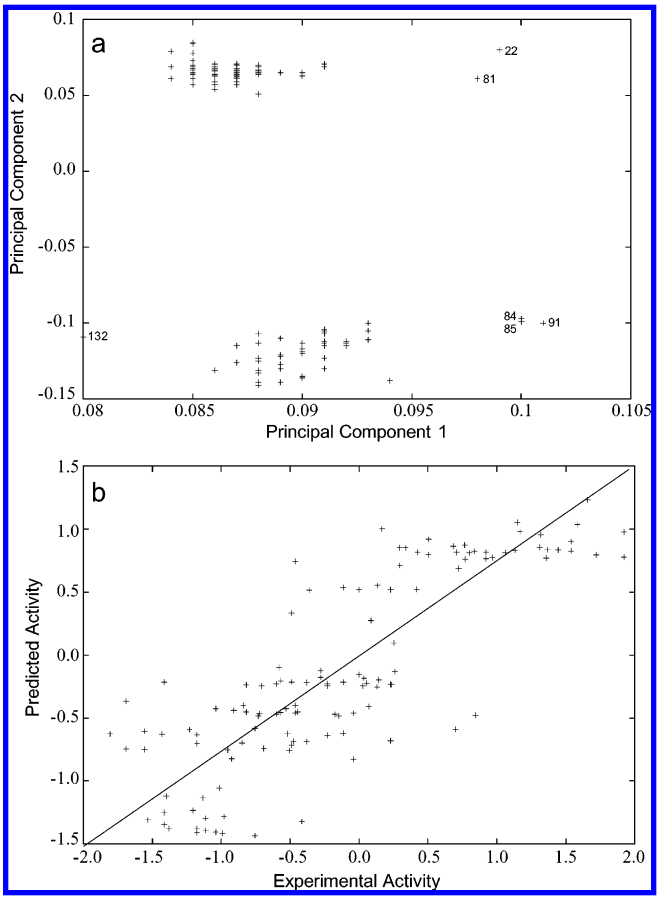



**Figure 3.** (a) PCA score plot of 128 objects on the basis of the first two principal components (respectively *x* and *y* axis) constituted with 738 energetical descriptors. Possible outliers on the first component (*x* axis) are labeled. (b) PLS plot of predicted (*y* axis) versus experimental activity (*x* axis).

**Table 10.** PLS Models

| model | objects | variables | $q^2$ (5-GR)[a] | $r^2$ | SDEC[b] | SDEP[c] | no. of components |
|---|---|---|---|---|---|---|---|
| 1 | 127 | 738 | 0.51 | 0.63 | 0.57 | 0.65 | 4 |
| 2 | 123 | 738 | 0.59 | 0.68 | 0.51 | 0.58 | 4 |
| 3 | 123 | 17 | 0.68 | 0.69 | 0.50 | 0.51 | 2 |

[a] 5-GR = five groups cross-validation. [b] SDEC = standard error of calculation. [c] SDEP = standard error of prediction.

(Table 10). The first QSAR analysis yielded a model defined by 4 principal components and a predictive index of 0.51.


**Figure 4.** Alignment of the headpiece of αv (residues 31−468) and α4 (residues 40−470) subunit. Residues involved in TR-14035 (**2**) binding are gray shaded.

The analysis of the model reveals the presence of four outliers, namely compound **59**, **81**, **85**, and **107**. Thus, these compounds were removed, and a new model was calculated (Table 10, model 2).

Before performing a variable selection, the system of variables was tested with a set of 50 randomizations to check if the genetic selection protocol could lead to erroneously model the noise. This procedure yielded a predictive index of −0.029 with 0.07 as the maximum $Q^2$ value and 18% as the number of cycles with the $Q^2 > 0$.

Thus the model was submitted to a variable selection based on the genetic algorithm described in the method part. In particular, 17 variables were selected. This optimized model yielded a better predictive index of 0.68 (model 3, Table 1 and Figure 3b) than the model based on the starting pool of 738 variables.

**Homology Modeling of the Headpiece of VLA-4 and Docking Experiments of TR-14035.** The headpiece of VLA-4 is constituted with the assembling of the von Willebrand factor (vWF)-type A domain, known as the I-like domain, of the $\beta$ subunit and the N-terminal region of the α subunit. The I-like domain adopts an $\alpha/\beta$ Rossmann fold and presents a metal-ion-adhesion site (MIDAS). Indeed, integrins require divalent cations to bind ligands. In particular, we used $Mn^{2+}$ since it is present in the crystal structure of $\alpha v\beta 3$.

Furthermore, the I-like domain is endowed by a large loop that is important for ligand selectivity and is known as the specificity-determining loop (SDL). The N-terminal region of the α subunit folds into a $\beta$-propeller domain which contains seven repeats. Since the binding site of VLA-4 ligands is located at the interface between the I-like domain of the $\beta$ subunit and the $\beta$-propeller domain of the α subunit, we modeled only this region on the basis of the crystal structure of $\alpha v\beta 3$ as described in the method section (Figures 4 and 5).

Docking experiments were carried out on TR-14035 (**2**). The pharmacophoric bioactive conformation of the biphenylalanine moiety of TR-14035 (**2**) is found among the best energetically favored solutions (Figure 6). Nevertheless, the N-benzoyl group adopts a different position compared to that used in the alignment of the test set. This is expected since we have no pharmacophoric reference in the VCAM-1 structure to address the bioactive conformation of the N-benzoyl fragment.

Thus, the N-benzoyl fragment and the carboxylic group of TR-14035 (**2**) interact with the $\beta$ subunit, while the biphenylalanine moiety points toward the α subunit. In particular, the aromatic ring of the N-benzoyl fragment makes hydrophobic interactions with the side chain of $\beta$-Tyr153, while the amido group forms hydrogen bonds with $\beta$-Asn244. The carboxylic group of TR-14035 (**2**) interacts with the $Mn^{2+}$ ion of the MIDAS motif of the I-like domain and the side chains of $\beta$-Ser152, $\beta$-Ser154. The biphenylalanine moiety makes hydrophobic interactions with the side chains of α-Tyr226 and α-Phe253.

## DISCUSSION

Starting from a VCAM-1 based pharmacophore model, we obtained an alignment of 128 inhibitors of VLA-4 adhesion. The alignment of compounds was used to construct a pseudoreceptor model of the binding site of VLA-4. The

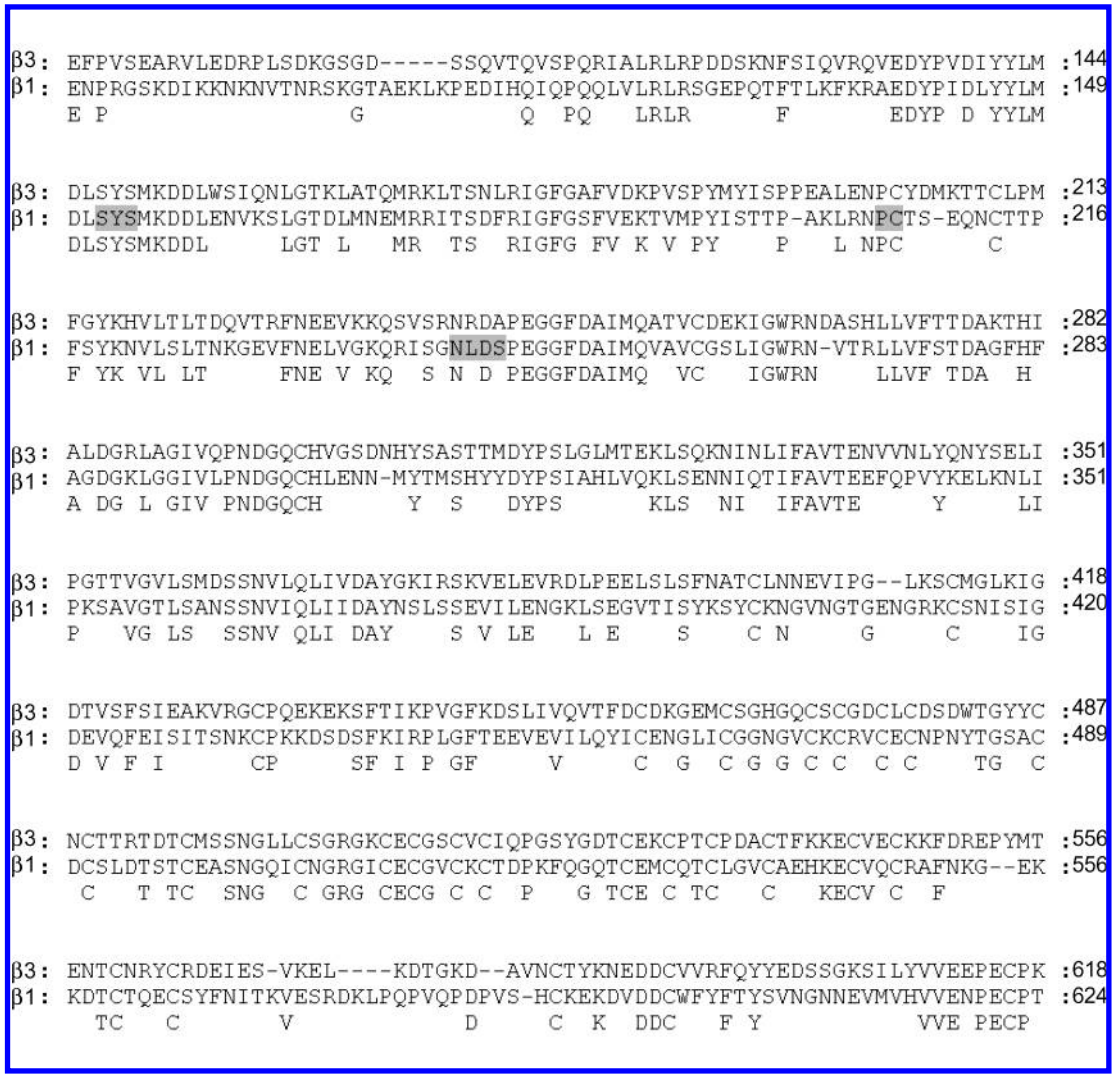



**Figure 5.** Alignment of the headpiece of $\beta$3 (residues 81−618) and $\beta$1 (residues 81−624) subunit. Residues involved in TR-14035 (**2**) binding are gray shaded.

surface of the pseudoreceptor was instrumental to calculate 738 energetical descriptors of the interaction between each compound and each point of the pseudoreceptor surface. The system, composed with 739 variables (738 descriptors plus the inhibitory activity of compounds) and 128 objects, was submitted to a 3D-QSAR analysis using a genetic algorithm/PLS based program. The starting model, after the removal of outliers, yielded a $Q^2$ of 0.52 with 4 principal components. Genetic variable selection improved the predictive index to 0.68 with only 2 principal components. In particular, the optimized model was composed with 17 variables endowed with the highest occurrence in the fittest chromosome (Table 11).

Thirteen out of 17 were hydrophobic interaction points (var. 1−4, 6−14) of the pseudoreceptor, while the remaining 4 were hydrogen bond interaction points (hydrogen bond acceptor: var. 5; hydrogen bond donor: var. 15−17). The first component of the model explains the 55% of the variance and is constituted mainly with variables 17 (loading = 1.469), 5 (loading = 0.603), 15 (loading = 0.447), 9 (loading = 0.284), and 16 (loading = 0.221).

The spatial plot (Figure 7) of these variables around TR-14035 (**2**) shows that the hydrogen bond descriptors are distributed around the carboxylic group and amide moiety of the N-benzoyl fragment. Conversely, the hydrophobic point is located near the biphenylalanine moiety. It is worth noting how variable 17, presenting the highest loading on the first component, lays along the lone pair vector of the carboxylic group.

These results are in agreement with the experimental evidences that the carboxylic group is a required moiety for the activity of these compounds. Indeed, the principal interaction of VCAM-1 with VLA-4 is a coordinating bond between the carboxylic group of the aspartate belonging to the IDS peptide and the $Mn^{2+}$ ion located in the metal-ion dependent adhesion site (MIDAS) of the $\beta$ subunit. In the first component, variable 9 presents the highest loading score among the hydrophobic points. It is located near the biphenylalanine group. On the basis of the VCAM-1 based pharmacophore, this group should mimick the hydrophobic side chain of Leu70. Since site-directed mutagenesis pinpointed a key role for Leu70 in the interaction with VLA-4, the selection of variable 9 confirms the likelihood of our studies.

The second component explains the 40% of the variance. Many of the variables with a positive value on the first component assume negative loadings in the second component with the only exception of variable 5. This is explained

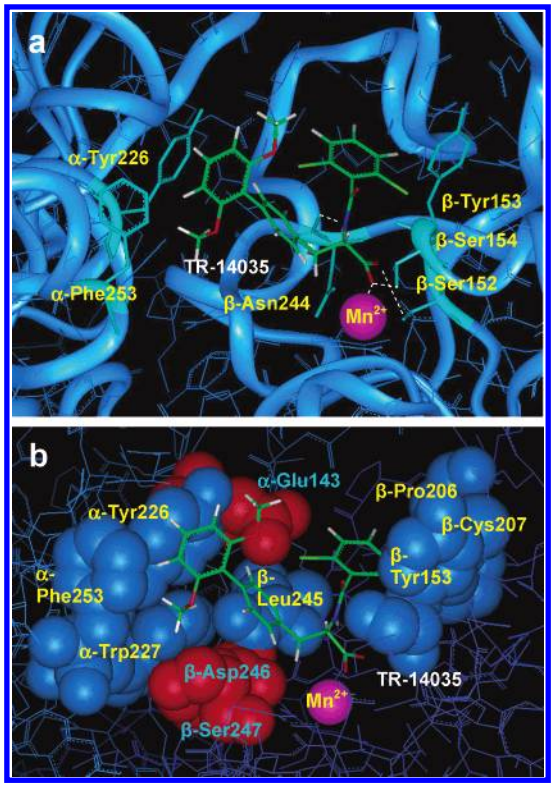



**Figure 6.** (a) Docking result of TR-14035 (**2**) into the binding site of VLA-4. Residues interacting with compound **2** are labeled. $Mn^{2+}$ is shown in CPK. (b) Binding pocket of TR-14035. Residues constituting the binding pocket of TR-14035 are in CPK and color coded on the basis of their physicochemical property. Hydrophobic residues are shown in blue. Polar and charged residues are colored in red.

**Table 11.** Physicochemical Properties and PLS Loadings of Selected Variables

| surface point no. of variable | properties | | | | PLS-loading | |
|---|---|---|---|---|---|---|
| | charge[a] | ESP[b] | H-bond[c] | HY[d] | compd 1 | compd 2 |
| 1 | 0.04811 | −0.00297 | −0.08230 | 0.9518 | 0.079 | −0.182 |
| 2 | −0.00277 | −0.00576 | −0.16346 | 0.9972 | 0.043 | −0.120 |
| 3 | 0.04666 | 0.00581 | 0.03125 | 0.9533 | 0.049 | −0.139 |
| 4 | −0.01783 | −0.00709 | 0.00000 | 0.9821 | 0.090 | −0.173 |
| 5 | 0.05656 | 0.03799 | 0.29688 | 0.0000 | 0.603 | 0.036 |
| 6 | 0.02436 | 0.00474 | 0.02344 | 0.9756 | 0.040 | −0.113 |
| 7 | −0.00287 | −0.00003 | −0.00434 | 0.9971 | 0.042 | −0.119 |
| 8 | 0.04238 | 0.00639 | 0.00019 | 0.9576 | 0.081 | −0.173 |
| 9 | −0.04569 | −0.00582 | −0.33087 | 0.9543 | 0.284 | −0.264 |
| 10 | 0.01311 | −0.00337 | −0.10187 | 0.9868 | 0.041 | −0.114 |
| 11 | 0.05643 | 0.00345 | 0.00000 | 0.9435 | 0.069 | −0.175 |
| 12 | 0.04963 | 0.00695 | 0.00000 | 0.9503 | 0.072 | −0.125 |
| 13 | 0.01853 | 0.00581 | 0.00781 | 0.9814 | 0.042 | −0.116 |
| 14 | 0.03887 | 0.00498 | 0.02835 | 0.9611 | 0.146 | −0.200 |
| 15 | −0.33110 | −0.03874 | −0.99099 | 0.0000 | 0.447 | −0.009 |
| 16 | −0.32576 | −0.03363 | −0.98249 | 0.0000 | 0.221 | −0.220 |
| 17 | −0.32380 | −0.07159 | −1.00000 | 0.0000 | 1.469 | −0.785 |

[a] Charge = point charge of the pseudoreceptor surface point. [b] ESP = electrostatic potential. [c] H-bond = hydrogen bond interaction (positive/negative values are referred to acceptor/donor property). [d] HY = hydrophobic property (range of values 0−1).

by the fact that the binding event described is not linear. Thus, the second component should represent the correction of the first component for the nonlinearity of the binding event. Beside, since many hydrophobic interaction points (var. 1−4, 6−8, and 10−14) assume the highest loading in the second component rather than in the first one, the second component should describe an overall negative effect on the biological activity of the hydrophobic interaction in these points.

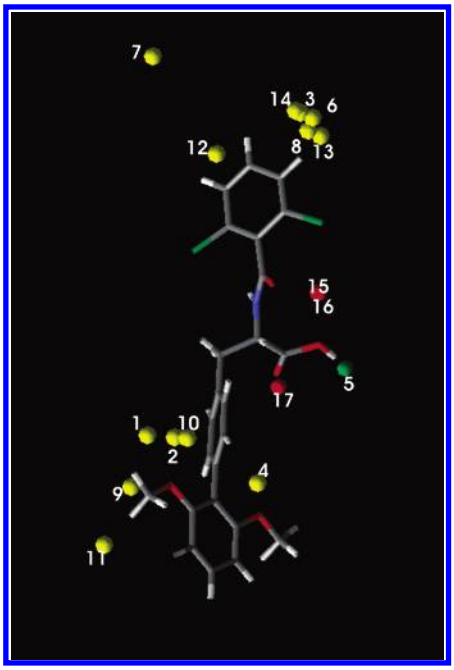



**Figure 7.** Plot of the selected variables around TR-14035 (**2**). Variables describing the hydrophobic interactions are colored in yellow; variables of the hydrogen bond donor interactions are colored in red; the variable indicating the hydrogen bond acceptor interaction is colored in green.

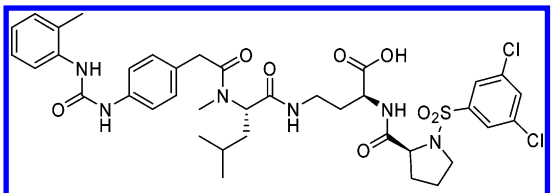
To interpret the meaning of the selected variables in term of favorable/unfavorable interactions, we constructed a homology model of the VLA-4 binding site on the basis of the crystal structure of $\alpha v\beta 3$. Docking experiments of TR-14035 (**2**) pinpointed a bioactive conformation similar to that one found using the pharmacophoric approach. In particular, the only difference is the orientation of the N-benzoyl fragment that was not constrained by the definition of a specific pharmacophoric point on the VCAM-1.

In agreement with the spatial location of hydrogen bonds descriptors, the carboxylic group and amide moiety of TR-14035 (**2**) are surrounded by polar residues. Thus, hydrogen bond interactions are observed between the amide moiety and $\beta$-Asn244. The carboxylic group of TR-14035 (**1**) not only coordinates the $Mn^{2+}$ ion but also forms a hydrogen bond network with $\beta$-Ser152 and $\beta$-Ser154.

The position of hydrophobic descriptors around TR-14035 (**2**) is in agreement with the docking experiments as well. Indeed, the benzoyl fragment interacts with $\beta$-Tyr153. However the presence of variables 3, 6−8, and 12−14 around this moiety pinpoints that hydrophobic interactions are unfavorable for the biological activity in this region of the binding site. The reason is that the negatively charged side chain of α-Glu143, facing the binding site of the benzoyl group, disrupts the hydrophobic pocket constituted with $\beta$-Tyr153, $\beta$-Leu245, $\beta$-Pro206, and $\beta$-Cys207 (Figure 6).

Docking experiments report that the biphenylalanine moiety lays in a cleft between the side chains of α-Tyr226, α-Trp227, and α-Phe253. The location of variable 9 near this moiety indicates the favorable nature of these hydrophobic interactions for the biological activity. However, since variables 1, 2, 4, 10, and 11 are present as well around the

**Chart 2**



**133** BIO-5192

contour of the aromatic rings, the effect of the above hydrophobic interactions is smoothed by the presence of the polar side chains of $\beta$-Asp246 and $\beta$-Ser247.

Recently a new selective and potent inhibitor of VLA-4 has been reported.[47] This molecule (BIO-5192, **133** Chart 2) represents a hybrid between compounds **1** and **4**. Indeed, it contains a diaryl urea moiety and a sulfonamide cap. It is worth to note how BIO-5192 could represent the key to join both classes of $\beta$-amino acid derivatives and phenylalanine derivatives into one interaction model to describe qualitatively VLA-4 inhibition. Indeed, if the sulfonamide cap would occupy the N-benzoyl binding pocket of TR-14035 (**2**), then the diaryl urea moiety would point toward the $\alpha$ subunit. Interestingly, following this interaction model, the leucine side chain of BIO-5192 would be aligned with the location of the hydrophobic variable 9 and the Leu70 of VCAM-1.

## CONCLUSIONS

The development of VLA-4 antagonists is a challenging area for the treatment of many diseases involving a massive recruitment of granulocytes during an inflammatory response.

So far two classes of VLA-4 antagonists have been reported: $\beta$-amino acid derivatives containing a diaryl urea moiety (BIO-1211, **1**) and phenylalanine derivatives (TR-14035, **2**). In this work we have reported for the first time a combined approach to study the structure−activity relationship of compounds belonging to the class of phenylalanine derivatives.

In particular, using a VCAM-1 based pharmacophore model, 3D-QSAR analysis, and docking studies, we get insights into the structural basis of VLA-4 recognition for a set of 128 phenylalanine derivatives. The agreement found among experimental data, docking studies, and 3D-QSAR analysis indicates the goodness of our predictive model.

The results herein reported may be used to design novel selective VLA-4 antagonists and to screen virtual libraries of compounds. Furthermore, the good predictivity shown by the 3D-QSAR model will be fruitful to prioritize a list of potential antagonists to be synthesized.

## REFERENCES AND NOTES


(1) Hynes, R. O. Cell adhesion: old and new questions. *Trends Cell Biol.* **1999**, *9*, M33−37.

(2) van der Flier, A.; Sonnenberg, A. Function and interactions of integrins. *Cell Tissue Res* **2001**, *305*, 285−298.

(3) Xiong, J. P.; Stehle, T.; Diefenbach, B.; Zhang, R.; Dunker, R. et al. Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **2001**, *294*, 339−345.

(4) Beglova, N.; Blacklow, S. C.; Takagi, J.; Springer, T. A. Cysteine-rich module structure reveals a fulcrum for integrin rearrangement upon activation. *Nat. Struct. Biol.* **2002**, *9*, 282−287.

(5) Xiong, J. P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M. et al. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **2002**, *296*, 151−155.

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

(7) Marinelli, L.; Lavecchia, A.; Gottschalk, K. E.; Novellino, E.; Kessler, H. Docking studies on alphavbeta3 integrin ligands: pharmacophore refinement and implications for drug design. *J. Med. Chem.* **2003**, *46*, 4393−4404.

(8) 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**, *3*, 470−473.

(9) Takagi, J.; Erickson, H. P.; Springer, T. A. C-terminal opening mimics 'inside-out' activation of integrin alpha5beta1. *Nat. Struct. Biol.* **2001**, *8*, 412−416.

(10) Takagi, J.; Petre, B. M.; Walz, T.; Springer, T. A. Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **2002**, *110*, 599−511.

(11) Lu, C.; Takagi, J.; Springer, T. A. Association of the membrane proximal regions of the alpha and beta subunit cytoplasmic domains constrains an integrin in the inactive state. *J. Biol. Chem.* **2001**, *276*, 14642−14648.

(12) Vinogradova, O.; Velyvis, A.; Velyviene, A.; Hu, B.; Haas, T. et al. A structural mechanism of integrin alpha(IIb)beta(3) "inside-out" activation as regulated by its cytoplasmic face. *Cell* **2002**, *110*, 587−597.

(13) Kim, M.; Carman, C. V.; Springer, T. A. Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science* **2003**, *301*, 1720−1725.

(14) Hynes, R. O. Integrins: bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673−687.

(15) Elices, M. J. *Cell Adhesion Molecules and Matrix Proteins: Role in Health and Diseases*; Springer-Verlag: Berlin, 1998; 133.

(16) Abraham, W. M.; Sielczak, M. W.; Ahmed, A.; Cortes, A.; Lauredo, I. T. et al. Alpha 4-integrins mediate antigen-induced late bronchial responses and prolonged airway hyperresponsiveness in sheep. *J. Clin. Invest.* **1994**, *93*, 776−787.

(17) Sagara, H.; Matsuda, H.; Wada, N.; Yagita, H.; Fukuda, T. et al. A monoclonal antibody against very late activation antigen-4 inhibits eosinophil accumulation and late asthmatic response in a guinea pig model of asthma. *Int. Arch. Allergy Immunol.* **1997**, *112*, 287−294.

(18) Podolsky, D. K.; Lobb, R.; King, N.; Benjamin, C. D.; Pepinsky, B. et al. Attenuation of colitis in the cotton-top tamarin by anti-alpha 4 integrin monoclonal antibody. *J. Clin. Invest.* **1993**, *92*, 372−380.

(19) Seiffge, D. Protective effects of monoclonal antibody to VLA-4 on leukocyte adhesion and course of disease in adjuvant arthritis in rats. *J. Rheumatol.* **1996**, *23*, 2086−2091.

(20) Yang, X. D.; Karin, N.; Tisch, R.; Steinman, L.; McDevitt, H. O. Inhibition of insulitis and prevention of diabetes in nonobese diabetic mice by blocking L-selectin and very late antigen 4 adhesion receptors. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10494−10498.

(21) Burkly, L. C.; Jakubowski, A.; Hattori, M. Protection against adoptive transfer of autoimmune diabetes mediated through very late antigen-4 integrin. *Diabetes* **1994**, *43*, 529−534.

(22) Komoriya, A.; Green, L. J.; Mervic, M.; Yamada, S. S.; Yamada, K. M. et al. The minimal essential sequence for a major cell type-specific adhesion site (CS1) within the alternatively spliced type III connecting segment domain of fibronectin is leucine-aspartic acid-valine. *J. Biol. Chem.* **1991**, *266*, 15075−15079.

(23) Humphries, M. J.; Sheridan, J.; Mould, A. P.; Newham, P. Mechanisms of VCAM-1 and fibronectin binding to integrin alpha 4 beta 1: implications for integrin function and rational drug design. *Ciba Found Symp.* **1995**, *189*, 177−191; discussion 191−179.

(24) Makarem, R.; Newham, P.; Askari, J. A.; Green, L. J.; Clements, J. et al. Competitive binding of vascular cell adhesion molecule-1 and the HepII/IIICS domain of fibronectin to the integrin alpha 4 beta 1. *J. Biol. Chem.* **1994**, *269*, 4005−4011.

(25) Jones, E. Y.; Harlos, K.; Bottomley, M. J.; Robinson, R. C.; Driscoll, P. C. et al. Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 A resolution. *Nature* **1995**, *373*, 539−544.

(26) Wang, J. H.; Pepinsky, R. B.; Stehle, T.; Liu, J. H.; Karpusas, M. et al. The crystal structure of an N-terminal two-domain fragment of vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on the domain 1 C−D loop can inhibit VCAM-1-alpha 4 integrin interaction. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5714−5718.

(27) Lin, K. C.; Castro, A. C. Very late antigen 4 (VLA4) antagonists as antiinflammatory agents. *Curr. Opin. Chem. Biol.* **1998**, *2*, 453−457.

(28) Nowlin, D. M.; Gorcsan, F.; Moscinski, M.; Chiang, S. L.; Lobl, T. J. et al. A novel cyclic pentapeptide inhibits alpha 4 beta 1 and alpha

5 beta 1 integrin-mediated cell adhesion. *J. Biol. Chem.* **1993**, *268*, 20352−20359.
(29) Thorsett, E. D.; Yenosk, T. A.; Pleiss, M. A. *Inhibitors of leukocyte and adhesion*; WO9602644, 1996.
(30) Lin, K.; Ateeq, H. S.; Hsiung, S. H.; Chong, L. T.; Zimmerman, C. N. et al. Selective, tight-binding inhibitors of integrin alpha4beta1 that inhibit allergic airway responses. *J. Med. Chem.* **1999**, *42*, 920−934.
(31) Sircar, I.; Gudmundsson, K. S.; Martin, R.; Liang, J.; Nomura, S. et al. Synthesis and SAR of N-benzoyl-L-biphenylalanine derivatives: discovery of TR-14035, a dual alpha(4)beta(7)/alpha(4)beta(1) integrin antagonist. *Bioorg. Med. Chem.* **2002**, *10*, 2051−2066.
(32) Li, B.; de Laszlo, S. E.; Kamenecka, T. M.; Kopka, I. E.; Durette, P. L. et al. N-(arylacetyl)-biphenylalanines as potent VLA-4 antagonists. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2141−2144.
(33) Hagmann, W. K.; Durette, P. L.; Lanza, T.; Kevin, N. J.; de Laszlo, S. E. et al. The discovery of sulfonylated dipeptides as potent VLA-4 antagonists. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2709−2713.
(34) Kamenecka, T. M.; Lanza, T., Jr.; de Laszlo, S. E.; Li, B.; McCauley, E. D. et al. N-aryl-prolyl-dipeptides as potent antagonists of VLA-4. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2205−2208.
(35) Singh, J.; Van Vlijmen, H.; Liao, Y.; Lee, W. C.; Cornebise, M. et al. Identification of potent and novel alpha4beta1 antagonists using in silico screening. *J. Med. Chem.* **2002**, *45*, 2988−2993.
(36) You, T. J.; Maxwell, D. S.; Kogan, T. P.; Chen, Q.; Li, J. et al. A 3D structure model of integrin alpha 4 beta 1 complex: I. Construction of a homology model of beta 1 and ligand binding analysis. *Biophys. J.* **2002**, *82*, 447−457.
(37) Singh, J.; van Vlijmen, H.; Lee, W. C.; Liao, Y.; Lin, K. C. et al. 3D QSAR (COMFA) of a series of potent and highly selective VLA-4 antagonists. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 201−211.
(38) Chiu, H. H.; Crowe, D. T.; Renz, M. E.; Presta, L. G.; Jones, S. et al. Similar but nonidentical amino acid residues on vascular cell adhesion molecule-1 are involved in the interaction with alpha 4 beta 1 and alpha 4 beta 7 under different activity states. *J. Immunol.* **1995**, *155*, 5257−5267.
(39) Accelrys. *Cerius-2*, *Insight-II*; San Diego, CA.
(40) Casewit, C. J.; Colwell, K. S.; Rappe, A. K. Application of a universal force field to main group compounds. *J. Am. Chem. Soc.* **1992**, *114*, 10046−10053.
(41) Hahn, M.; Rogers, D. Receptor surface models. 2. Application to quantitative structure−activity relationships studies. *J. Med. Chem.* **1995**, *38*, 2091−2102.
(42) Hahn, M. Receptor surface models. 1. Definition and construction. *J. Med. Chem.* **1995**, *38*, 2080−2090.
(43) Leardi, R.; González, A. L. Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemom. Intell. Lab. Syst.* **1998**, *41*, 195−207.
(44) Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **1995**, *23*, 318−326.
(45) MacKerell, J., A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(46) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E. et al. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(47) Leone, D. R.; Giza, K.; Gill, A.; Dolinski, B. M.; Yang, W. et al. An assessment of the mechanistic differences between two integrin alpha 4 beta 1 inhibitors, the monoclonal antibody TA-2 and the small molecule BIO5192, in rat experimental autoimmune encephalomyelitis. *J. Pharmacol. Exp. Ther.* **2003**, *305*, 1150−1162.

CI049914L