THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 28, pp. 19757–19768, July 11, 2008

Structural Basis for the Interaction of *iso*DGR with the RGD-binding Site of $\alpha v\beta 3$ Integrin*[S]

Received for publication, December 18, 2007, and in revised form, May 7, 2008 Published, JBC Papers in Press, May 13, 2008, DOI 10.1074/jbc.M710273200

Andrea Spitaleri[‡1,2], Silvia Mari[‡1], Flavio Curnis[§], Catia Traversari[¶], Renato Longhi[‖], Claudio Bordignon[¶], Angelo Corti[§3], Gian-Paolo Rizzardi[¶4], and Giovanna Musco[‡5]

From the [‡]Dulbecco Telethon Institute Biomolecular NMR Laboratory %S. Raffaele Scientific Institute, 20132 Milan, the [§]Department of Oncology, Cancer Immunotherapy and Gene Therapy Program and Italian Institute of Technology Network Research Unit of Molecular Neuroscience, S. Raffaele Scientific Institute, 20132 Milan, [¶]MolMed SpA, 20132 Milan, and [‖]Istituto di Chimica del Riconoscimento Molecolare, Consiglio Nazionale delle Ricerche, 20131 Milan, Italy

Asparagine deamidation at the NGR sequence in the 5th type I repeat of fibronectin (FN-I_5) generates *iso*DGR, an $\alpha v\beta 3$ integrin-binding motif regulating endothelial cell adhesion and proliferation. By NMR and molecular dynamics studies, we analyzed the structure of C*iso*DGRC (*iso*DGR-2C), a cyclic β-peptide mimicking the FN-I_5 site, and compared it with NGR, RGD, or DGR-containing cyclopeptides. Docking experiments show that *iso*DGR, exploiting an inverted orientation as compared with RGD, favorably interacts with the RGD-binding site of $\alpha v\beta 3$, both recapitulating canonical RGD-$\alpha v\beta 3$ contacts and establishing additional polar interactions. Conversely, NGR and DGR motifs lack the fundamental pharmacophoric requirements for high receptor affinity. Therefore, unlike NGR and DGR, *iso*DGR is a new natural recognition motif of the RGD-binding pocket of $\alpha v\beta 3$. These findings contribute to explain the different functional properties of FN-I_5 before and after deamidation, and provide support for the hypothesis that NGR → *iso*DGR transition can work as a molecular timer for activating latent integrin-binding sites in proteins, thus regulating protein function.

A number of cellular interactions with the extracellular matrix (ECM)[6] are mediated by fibronectins, which are large adhesive glycoproteins (~450 kDa) involved in several key processes, including embryogenesis, angiogenesis, inflammation, hemostasis, thrombosis, and tissue repair (1, 2). Fibronectins are soluble elements of plasma and other body fluids, as well as constituents of the insoluble ECM (2, 3). Human fibronectin (FN), which is typically composed of two almost identical subunits connected covalently by disulfide bonds at their C termini, is thus an abundant and ubiquitous ECM protein present in about 20 isoforms and consisting primarily of three types of repeating modules (FN-I, FN-II, and FN-III) (2–5). Distinct FN modules contain varying binding sites for several different molecules, including sulfated glycosaminoglycans, syndecans, DNA, gelatin, heparin, and fibrin (2, 3, 6). In addition, fibronectins contain binding sites for about half of the known cell surface integrin receptors (7, 8). Integrin receptors are present in many animal species, ranging from sponges to mammals (9), and play essential roles in cellular physiology (attachment, migration, proliferation, differentiation, and survival) and in disease (cancer, tumor metastasis, immune dysfunction, ischemia-reperfusion injury, viral infections, osteoporosis, and coagulopathies) (4, 10).

In this context, it is worth noting that FN-I_5 (*i.e.* the 5th FN-I repeat) and FN-I_7 modules contain a GNGRG loop that is conserved in human, bovine, murine, rat, bird, amphibian, and fish fibronectin, suggesting that this loop is functionally important (11). In addition, we have recently shown that the deamidation of Asn^{263} at the Asn-Gly-Arg (NGR) sequence of FN-I_5 and of peptides containing the NGR motif generates *iso*DGR (*iso*Asp-Gly-Arg), a novel cell adhesion motif binding to $\alpha v\beta 3$ integrin (12). Consistently, cells from homozygous knock-in mice carrying the RGD → RGE mutation in fibronectin exhibited normal fibronectin fibril assembly, both *in vitro* and *in vivo*, exploiting the presence in the FN-I_5 repeat of a NGR → *iso*DGR motif binding to $\alpha v\beta 3$ (13). Integrin $\alpha v\beta 3$ is a relevant receptor in tumor angiogenesis and metastasis, viral infections, inflammation, and bone resorption (8, 14), and its ligands contain the Arg-Gly-Asp (RGD) sequence (15, 16). Importantly, a cyclic β-peptide containing the C*iso*DGRC motif (*iso*DGR-2C) is a competitive antagonist of RGD-containing ligands of $\alpha v\beta 3$ and inhibits endothelial cell adhesion, proliferation, and tumor growth (12). Furthermore, analysis of competitive binding plots of RGD-2C and *iso*DGR-2C shows that both ligands have comparable binding affinity for $\alpha v\beta 3$, whereas DGR-2C and NGR-2C cyclopeptides show a >2 order of magnitude

* This work has been supported in part by MolMed SpA. C. T., C. B., and G. P. R. are MolMed employees, and A. C. and G. M. are consultants of MolMed. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3 and Tables S1–S3.

[1] Both authors contributed equally to this work.

[2] Supported by FIRB Grant RBIP06FYF7.

[3] To whom correspondence may be addressed: MolMed, via Olgettina 58, 20132 Milan, Italy. Fax: 39.02.2643.4786; E-mail: corti.angelo@hsr.it.

[4] To whom correspondence may be addressed: MolMed, via Olgettina 58, 20132 Milan, Italy, Fax: +39.02.2127.7325; E-mail: paolo.rizzardi@molmed.com.

[5] Supported by Telethon Foundation. To whom correspondence may be addressed: via Olgettina 58, 20132 Milan, Italy. Fax: 39.02.2643.4153; E-mail: musco.giovanna@hsr.it.

[6] The abbreviations used are: APBS, adaptive Poisson-Boltzmann solver program; a.u., arbitrary units; ECM, extracellular matrix; FN, fibronectin; (hTNF)-α, human tumor necrosis factor; MD, molecular dynamics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; OPLS, optimized parameters for liquid simulation; REMD, replica exchange molecular dynamics simulations; r.m.s.d., root mean square deviation; MIDAS, metal ion-dependent adhesion site.

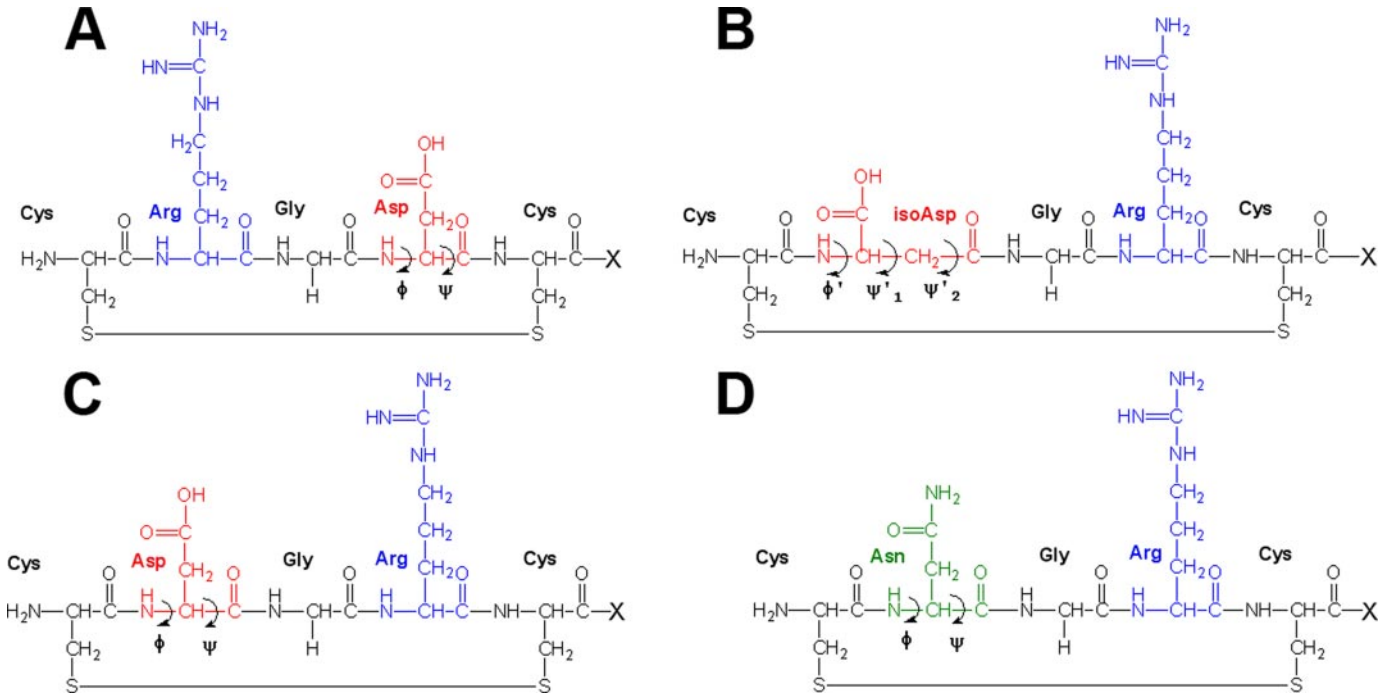

FIGURE 1. **Schematic representation of RGD-2C (*A*), *iso*DGR-2C (*B*), NGR-2C (*C*), and DGR-2C (*D*).** The torsion angles along the backbone main chain are explicitly labeled and are defined as follows: $\varphi = C'_{i-1}$–N_i–C-α_i–C'_i; $\psi = N_i$–C-α_i–C'_i–N_{i+1}. In the isoaspartic residue, torsion angles are defined as follows: $\varphi' = C_{i-1}$–N_i–Cα_i–Cβ_i; $\psi'_1 = N_i$–Cα_i–Cβ_i–C_i; $\psi'_2 = C\alpha_i$–Cβ_i–C_i–N_{i+1}. The insertion of an additional methylene group in the isoaspartic residue introduces an additional torsion angle (φ'_2). *X* denotes the C-terminal tail composed of residues GVRY.

lower affinity (12), pointing to stereospecific *iso*DGR-$\alpha v\beta 3$ interactions.

The structural determinants dictating the interaction between *iso*DGR and $\alpha v\beta 3$ are still unknown. To gain a structural insight into the different $\alpha v\beta 3$ recognition mechanisms, we analyzed the conformation in solution of four ligands containing either the *iso*DGR, RGD, NGR, or DGR motifs (Fig. 1 and supplemental Fig. S1), and we created putative binding models of the four ligands with $\alpha v\beta 3$ based on the crystallographic structure of $\alpha v\beta 3$ in its "RGD-bound" conformation (15). Consistently with functional data, our docking studies show that the *iso*DGR motif can perfectly mimic the canonical RGD interactions with $\alpha v\beta 3$, whereas NGR and DGR lack the stereochemical and electrostatic requirements for a correct recognition of the $\alpha v\beta 3$-binding pocket.

EXPERIMENTAL PROCEDURES

Cell Line and Reagents—EA.hy926 cells (human endothelial cell fused with human lung carcinoma A549) were cultured as described previously (17). Human $\alpha v\beta 3$ was from Immunological Science (Rome, Italy), and streptavidin peroxidase was from Società Prodotti Antibiotici (Milan, Italy). NGR-TNF and FN-I_5 were prepared as described previously (12).

Preparation and Characterization of Synthetic Peptides—CRGDCGVRY (RGD-2C), CDGRCGVRY (DGR-2C), CNGRCGVRY (NGR-2C), and C*iso*DGRCGVRY (*iso*DGR-2C) were prepared as described (12). The biotinylated peptides CNGRCGVRSSSRTPSDKYGK-bio and CARACGVRSSSRTPSDKYGK-bio (called bio-CNGRC-hTNF-(1–11) and bio-CARAC-hTNF-(1–11)) consist of CNGRCG or CARACG fused to the N-terminal sequence of human tumor necrosis factor (hTNF)-α followed by a Tyr, to enable detection, and a biotinylated Lys. These peptides were prepared and purified as described (11). Deamidation of bio-CNGRC-hTNF-(1–11) was obtained by diluting bio-CNGRC-hTNF-(1–11) in 0.1 M ammonium bicarbonate buffer, pH 8.5, and incubating for 16 h at 37 °C. The resulting β-peptide was called "heat-treated" bio-CNGRC-hTNF-(1–11). All peptides were dissolved in sterile water and stored in aliquots at −20 °C. The molecular mass of each peptide was checked by matrix-assisted laser desorption ionization-time-of-flight analysis.

Binding of Peptides to $\alpha v\beta 3$ Integrin—The binding of biotinylated peptides (bio-CNGRC-hTNF-(1–11), heat-treated bio-CNGRC-hTNF-(1–11) at 37 °C, and bio-CARAC-hTNF-(1–11)) to $\alpha v\beta 3$ integrin immobilized on microtiter plates was analyzed as described previously, using streptavidin peroxidase-peptide complexes (12). These complexes were prepared by mixing various amounts of biotinylated peptides (4 to 0.062 μg) in phosphate-buffered saline with Ca^{2+} and Mg^{2+} (DPBS, Cambrex) containing 3% bovine serum albumin with 0.03 units of streptavidin peroxidase (binding capacity 1 μg of biotin/unit of streptavidin peroxidase, final volume of 15 μl). Complexes were diluted in 3% bovine serum albumin/DPBS (1:300), added to microtiter plates coated with purified human $\alpha v\beta 3$ integrin (0.5 μg/ml), and incubated for 2 h at room temperature. After washing with DPBS, bound peroxidase was detected by chromogenic reaction with *o*-phenylenediamine.

Cell Adhesion Assay—The effect of peptides on EA.hy926 cell adhesion was investigated by seeding EA.hy926 cells (7.5×10^3) in 96-well flat bottom plates in the presence of increasing concentration of ligands (NGR-2C, RGD-2C, *iso*DGR-2C, and DGR-2C). After 4 h of incubation at 37 °C, nonadherent cells were washed out. The amount of adherent cells was measured

by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide colorimetric assay (18).

NMR Experiments and Structure Calculations—For each ligand NMR spectra of an ~5 mM sample (90% H_2O, 10% D_2O) at pH 3 were recorded at 280 K on a Bruker Avance-600 spectrometer (Bruker BioSpin) equipped with a triple-resonance TCI cryoprobe with an *x*, *y*, *z* shielded pulsed-field gradient coil. The experiments were performed at acidic pH to avoid both signal loss because of high exchange rates at neutral pH and deamidation of the asparagine in the NGR-2C peptide (12). Proton resonances were assigned by conventional two-dimensional experiments as follows: total correlation spectroscopy (t_{mix} = 60 ms), nuclear Overhauser effect spectroscopy (NOESY), and rotational nuclear Overhauser effect spectroscopy (t_{mix} = 100–400 ms) (19). Cross-peaks intensities were measured from NOESY spectra at 200 ms. No differences were observed in the experiments at higher mixing times. Water proton signals were suppressed with excitation sculpting sequence (20).

All resonances of the four ligands have been assigned with the only exception of the amide proton of C1, which has a higher exchange rate with the solvent even at low pH (Biological Magnetic Resonance Data Bank accession codes: RGD, 1928339, *iso*DGR, 38926841; NGR, 32415285; and DGR, 52915018). The $^3J_{HN\text{-}HA}$ coupling constants were obtained directly from the resolved amide proton resonances of well digitized mono-dimensional spectrum. The temperature coefficients of the amide protons were obtained from linear fits of the chemical shift data from mono-dimensional spectra acquired in a temperature range from 280 to 300 K (5 K-increasing steps). Data were processed with NMRPipe (21) and analyzed using the NMRView software (22).

Structure Calculations—Structures were calculated using ARIA 1.2, ambiguous restraints for iterative assignment (23), in combination with CNS 1.2, crystallography, and NMR systems (24) using only manually assigned NOEs as experimental restraints. Coupling constants that were in the range between 6 and 8.5 Hz were not used in calculations because their values were suggestive of rapidly interconverting conformers coexisting in solution. To avoid bias in the calculations, hydrogen bonds inferred from temperature coefficients data were not included. These data were only used for structure validation. Calculations were carried out in the simplified all-hydrogen PARALLHDG5.3 force field with nonbonded interactions modeled by PROLSQ force field (22). Parameters for isoaspartic residue were derived from the aspartic residue applying appropriate dihedrals and improper angles. A total of eight iterations (200 structures per iteration) were performed. The ARIA default water refinement was performed on the 30 best structures of the final iteration. The stereochemical quality of the structures was assessed with PROCHECK-NMR program (25).

In the case of *iso*DGR-2C, residue 2 was excluded from torsion angle check. However, its torsion angle compared well with φ and ψ angles of isoaspartic residues measured in deposited x-ray structures (Protein Data Bank codes 1AT6, 1DY5, 1RTU, 2FI5, 2FTM, 1C9P, 1LSQ, 2FI4, and 2FTL).

Molecular Dynamics Simulations—Simulations were performed on the lowest energy NMR structures of each ligand using the GROMACS 3.3.1 package (26) with the optimized parameters for liquid simulation (OPLS) force field (27). All trajectories were calculated in periodic cubic boxes (5 × 5 × 5 nm) of explicit SPC water molecules (28).

The system was neutralized by 1–2 chloride ions according to the charge of the ligand. Bond lengths were constrained using the LINCS algorithm (29), and Lennard-Jones interactions were calculated with a 0.9-nm twin-range cutoff. Full electrostatic potentials were computed using the PME method (30) with a cutoff of 0.9 nm. The system was first minimized using steepest descent algorithm and then equilibrated at 300 K for 100 ps under NPT (number of molecules, volume, temperature) and periodic boundary conditions. REMD simulations were performed with 16 replicas run in parallel at the following temperatures from 293 to 353 K with 4 K-increasing steps. The temperatures were chosen so as to maintain an exchange rate of 5–10%. Transitions between adjacent temperatures were attempted every 500 MD steps (1 ps) using a Metropolis transition probability (31), which gives a probability of exchange between two replicas *i* and *j*: $P(i,j) = \exp(-(\beta_i - \beta_j)(E_j - E_i))$, where $\beta = 1/k_BT$; *E* is the potential energy of the system; *T* is the absolute temperature, and k_B is Boltzmann's constant. Each replica was simulated for 2 ns, with an integration time step of 0.002 fs under NVT (number of molecules, volume, temperature) conditions, yielding a total sampling time of 32 ns. Configurations were saved prior to every attempted transition, leading to an ensemble at each temperature containing 2000 structures. The full trajectories were clustered over the backbone atoms of the macrocycle and the disulfide bridge using the GROMOS algorithm g_cluster (32) as implemented in the GROMACS 3.3.1 package with a cutoff of 0.07 nm.

Electrostatic Surface Potential Calculations—The electrostatic potential of the four ligands has been calculated using the adaptive Poisson-Boltzmann solver program (APBS 0.5) (33). The charges (*Q*) and radii (*R*) of atoms in the PQR file required by APBS were taken from the OPLS force field. Electrostatic potential was visualized using the PyMOL program (Delano Scientific LLC) with positive potential in blue and negative potential in red in a range between −5 and +5 *kT*/e.

Molecular Docking Calculations—Docking calculations of the four ligands on the globular head of the extracellular part of $\alpha v\beta 3$ in its ligand-bound conformation (Protein Data Bank code 1L5G), have been performed using the docking program HADDOCK2.0 (34, 35).

For each ligand an ensemble of the best 30 NMR structures in terms of energy was docked onto $\alpha v\beta 3$. The protocol follows a three-stage docking procedure, which includes the following: (*a*) randomization of orientations and rigid body minimization, (*b*) simulated annealing in torsion angle space, and (*c*) refinement in Cartesian space with explicit water. Ambiguous interaction restraints ($\alpha v\beta 3$: Asp^{150}, Asp^{218}, Tyr^{122}, Arg^{214}, Asn^{215}, and Arg^{216}; ligand, residues 2–4) (35) were derived from the known interactions of the RGD motif of the cyclic pentapeptide in the Protein Data Bank structure 1L5G. OPLS force field was used. During the rigid body docking step 1000 structures were calculated, allowing the ligand to explore solutions rotated by 180°, thus increasing the sampling of the solutions. The best 200 solutions in terms of intermolecular energies were selected for

a semiflexible simulated annealing in which the side-chains of $\alpha v\beta 3$ and of the cyclic peptides located at the binding interface ($\alpha v\beta 3$, residues 148–152, 216–220 on the α_v domain and 118–120, 124–126 on $\beta 3$ domain; ligand, residues 1–9) were allowed to move in a semi-rigid body docking protocol to search for conformational rearrangements. The models were then subjected to a water refinement step (TIP3P model). Backbone and side-chain of amino acids 6–9 were fully flexible in iteration 1 and in the water refinement step. Analogous calculations have been performed using for each ligand an ensemble of conformers composed by the centroids representing the highest populated clusters (80%) identified in the clusters of REMD trajectories (4, 23, 10, and 5 structures for RGD, *iso*DGR, NGR, and DGR, respectively).

The analysis of the simulations was performed applying in-house Python and Tcl scripts. Root mean square deviation (r.m.s.d.) values were calculated using the ProFit program (available on line). The fitting of the protein was performed on the flexible residues (Table 3) using the McLachlan algorithm (36). The r.m.s.d. values of the cyclopeptides were calculated on the backbone of residues 1–5 and on the sulfur and Cβ atoms of residues 1 and 5. The final r.m.s.d. matrix was then clustered using the algorithm described in Daura *et al.* (32), where a cluster is defined as an ensemble of at least two conformations displaying a r.m.s.d. smaller than 1 Å. In the case of NGR-2C, the cut off was increased to 1.3 Å.

The final structures after water refinement were clustered and scored using a combination of energy terms defined as follows: $1.0 \times E_{vdW} + 1.0 \times E_{elec} + 0.1 \times E_{AIR} + 1.0 \times E_{desolv}$, the subscripts are as follows: vdW, van der Waals energy; elec, the electrostatic energy; AIR, the ambiguous interaction restraint energy; and desolv, the desolvation energy calculated using the atomic desolvation parameters of Fernandez-Recio and colleagues (37).

RESULTS

Unlike NGR-2C and DGR-2C, isoDGR-2C and RGD-2C Can Efficiently Bind $\alpha v\beta 3$ Integrin—We have shown previously that peptides containing the NGR motif, such as CNGRC and FN-I_5, can rapidly deamidate (half-life, 2–3 h) when incubated for 16 h at pH 8.5 and 37 °C. This reaction leads to the generation of species containing the DGR and *iso*DGR motifs (12). To assess the effect of NGR deamidation on $\alpha v\beta 3$ binding, we have analyzed the interaction of a biotinylated NGR peptide (bio-CNGRC-hTNF-(1–11)) to $\alpha v\beta 3$ integrin before and after incubation at 37 °C. To this aim, we prepared complexes of biotinylated peptides and streptavidin-peroxidase, and we analyzed their binding to $\alpha v\beta 3$ integrin adsorbed on microtiter plates. Although little or no binding was observed with bio-CNGRC-hTNF-(1–11) and bio-CARAC-hTNF-(1–11), a control peptide, a marked increase in binding occurred after heat treatment of bio-CNGRC-hTNF-(1–11) (Fig. 2*A*). Similar results were observed also with a FN-I_5 peptide (data not shown). This suggests that peptide deamidation is very critical for $\alpha v\beta 3$ binding.

Given that NGR deamidation can lead to the formation of *iso*DGR and DGR, we performed additional assays to assess which of these molecular species was responsible for the increase in binding after heat treatment. To this aim we compared the capability of synthetic DGR-2C, *iso*DGR-2C, and NGR-2C to compete the binding of heat-treated bio-CNGRC-hTNF-(1–11) to $\alpha v\beta 3$ integrin adsorbed on microtiter plates. The RGD-2C peptide was analyzed in parallel as a positive control. The results showed that *iso*DGR-2C and RGD-2C could inhibit the binding of heat-treated bio-CNGRC-hTNF-(1–11) to $\alpha v\beta 3$ with similar potency, whereas >100-fold higher concentration of DGR and NGR were necessary to induce partial competition (Fig. 2*B*). These results confirm the hypothesis that *iso*DGR, unlike NGR and DGR, can efficiently mimic RGD in binding $\alpha v\beta 3$.

To assess whether these ligand-integrin interactions can also occur in living cells, we investigated the effect of RGD-2C,

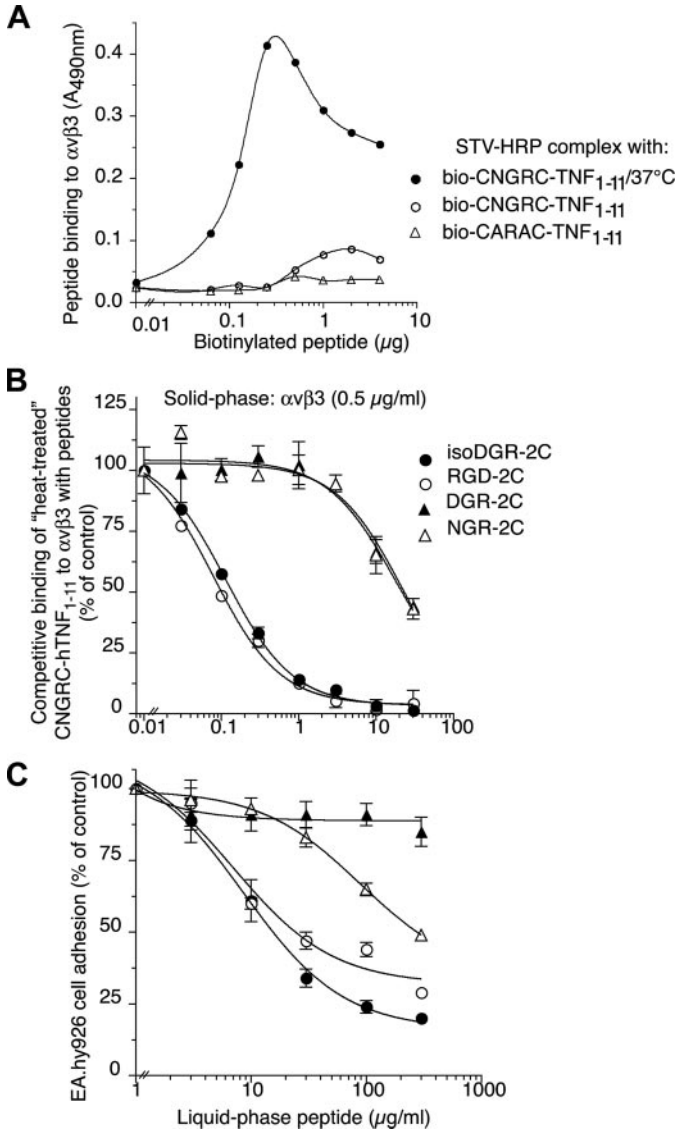

FIGURE 2. ***iso*DGR-2C and RGD-2C, unlike NGR-2C and DGR-2C, can efficiently bind $\alpha v\beta 3$ integrin.** *A,* binding of bio-CNGRC-hTNF-(1–11), bio-CARAC-hTNF-(1–11), and heat-treated bio-CNGRC-hTNF-(1–11) (complexed with streptavidin peroxidase) to solid-phase coated with purified $\alpha v\beta 3$ integrin. Binding was detected as described under "Experimental Procedures." *B,* competitive binding of heat-treated bio-CNGRC-hTNF-(1–11)-streptavidin peroxidase to $\alpha v\beta 3$ with various amounts of *iso*DGR-2C, RGD-2C, DRG-2C, and NGR-2C. *C,* adhesion of EA.hy926 cells to culture plates in the presence of various amounts of *iso*DGR-2C, RGD-2C, DGR-2C, and NGR-2C.

TABLE 1
Summary of conformational constraints and statistics for the 30 best structures of RGD-2C, *iso*DGR-2C, NGR-2C and DGR-2C

	RGD-2C	*iso*DGR-2C	NGR-2C	DGR-2C
Experimental NMR constraints				
Intra-residue ($i - j = 0$)	62	40	56	58
Sequential ($\|i - j\| = 1$)	29	20	30	30
Medium range ($2 \leq \|i - j\| \leq 4$)	4	0	4	3
Total	95	60	90	91
Deviation from idealized covalent geometry				
Bonds (Å)	0.004 ± 0.0002	0.004 ± 0.0003	0.004 ± 0.0003	0.004 ± 0.0002
Angles (°)	0.45 ± 0.05	0.47 ± 0.05	0.44 ± 0.04	0.45 ± 0.04
Coordinate precision (Å)[a]				
Backbone (residues 1–5)	1.1 ± 0.3	1.3 ± 0.4	1.1 ± 0.2	0.9 ± 0.3
All heavy atoms (residues 1–5)	1.6 ± 0.4	2.4 ± 0.5	1.8 ± 0.4	1.7 ± 0.4
Structural quality[b]				
Residues in most favored region of Ramachandran plot (% value)	53.3	56.7	82.2	64.4
Residues in additionally allowed region (% value)	46.7	40.0	15.6	35.6
Residues in generously allowed region (% value)	0.0	3.3	2.2	0.0

[a] r.m.s.d. between the ensemble of structures was calculated on the backbone atoms (N, Cα, and C′) of the cycle, including the Cβ and Sγ atoms of Cys[1] and Cys[5].
[b] Structural quality was measured only on residues 1–5, and the *iso*-Asp residue was excluded from the analysis.

DGR-2C, *iso*DGR-2C, and NGR-2C on the adhesion of EA.hy926 cells. As expected both RGD-2C and *iso*DGR-2C, but not DGR-2C, efficiently inhibited cell adhesion (Fig. 2*C*). Furthermore, a 10-fold higher concentration of NGR-2C was necessary to compete to comparable levels. Considering that assay incubation was 3.5 h and that the half-life of NGR deamidation in cell culture medium is 2–3 h (12), it is very likely that NGR competition was actually related to *iso*DGR formation during assay incubation.

Taken together, these results suggest that *iso*DGR, unlike DGR and NGR, can functionally mimic RGD in the interaction with adhesion receptors, such as $\alpha v\beta 3$ integrin. The modest activity observed with NGR could be due to deamidation occurring during assay incubation, considering the rapid kinetics of peptide deamidation (12).

Conformational Analysis of RGD-2C, isoDGR-2C, DGR-2C, and NGR-2C Macrocycles Free in Solution—To investigate the structural determinants at the basis of the different biological activities of RGD-2C, *iso*DGR-2C, NGR-2C, and DGR-2C, their conformational properties in the unbound state have been characterized by standard two-dimensional solution NMR methods and molecular dynamics simulations.

NMR Studies Indicate Flexibility of RGD-2C, isoDGR-2C, DGR-2C, and NGR-2C Macrocycles—Analysis of two-dimensional NOESY and rotational nuclear Overhauser effect spectroscopy spectra of the free ligands revealed the presence of only intra-residue and sequential nuclear Overhauser effect (NOE) and rotational nuclear Overhauser effect patterns with very few medium range contacts (Table 1), suggesting conformational variability for the ligands. The four ligands showed similar number of NOEs (~90) with the exception of *iso*DGR-2C, which had the lowest number of NOEs (~60). The latter suggests that the β-bond of the isoaspartic acid induces a greater conformational flexibility than that observed in the other three macrocycles.

Furthermore, the $^3J_{\text{HN-HA}}$ couplings inside the macrocycles were all comprised in values ranging between 6 and 8.5 Hz (supplemental Table S1), indicating averaging between rapidly interconverting conformers. In addition, some residues outside the ring presented higher coupling constants (*e.g.* Tyr[9] of RGD-2C $^3J_{\text{HN-HA}}$ = 8.7 Hz), suggestive of a more extended conformation for residues inside the tail.

Finally, the temperature coefficients ($|\delta\Delta/\Delta T| > 3.5$ ppb/K) of the amide protons of the ligands inside the macrocycle (from residue 1 to 4) indicated partially or fully solvent exposure, consistent with macrocycle flexibility (supplemental Table S1) (20). In a different way, residue Cys[5] showed lower temperature coefficients, suggesting a certain degree of solvent protection probably because of intra-molecular hydrogen bonds (supplemental Table S1).

Solution Structures of the Four Macrocycles Indicate That isoDGR-2C Displays the Highest Conformational Heterogeneity—Using the available short and medium range NOEs, solution structures of the four macrocycles have been calculated with the program CNS 1.2 (24), interfaced with ARIA 1.2 (23). For each ligand, we selected the 30 conformers with lowest energy for detailed structural analysis. The calculated structures were consistent with the experimental restraints, with no NOE distance violations greater than 0.4 Å and with all residues in the allowed regions of the Ramachandran plot (Table 1). The final ensembles of low energy structures of the four macrocycles are illustrated in Fig. 3, *A–D*. The C-terminal tails (residues 6–9) are highly disordered, whereas the macrocycles adopt a more ordered conformation with r.m.s.d. values ranging between 0.9 and 1.1 Å, calculated over both the backbone atoms and the disulfide bridge. In contrast, *iso*DGR-2C displays the highest conformational heterogeneity (r.m.s.d. ~ 1.3 Å), consistently with the lower number of NOEs (Table 1).

The dynamic behavior of the four macrocycles was further investigated by REMD, performed on the lowest energy conformer of each bundle of structures. The methodology is based on the parallel tempering Monte Carlo method (31, 38, 39), where multiple copies (or replicas) of identical systems are simulated in parallel at different temperatures, thus enabling an enhanced conformational sampling. We ran 16 replicas, for a total simulation length of 32 ns, over a temperature range between 293 and 353 K. For each ligand, we performed cluster analysis to separate the pool of conformations produced by the REMD simulations into families of structures of similar geo-

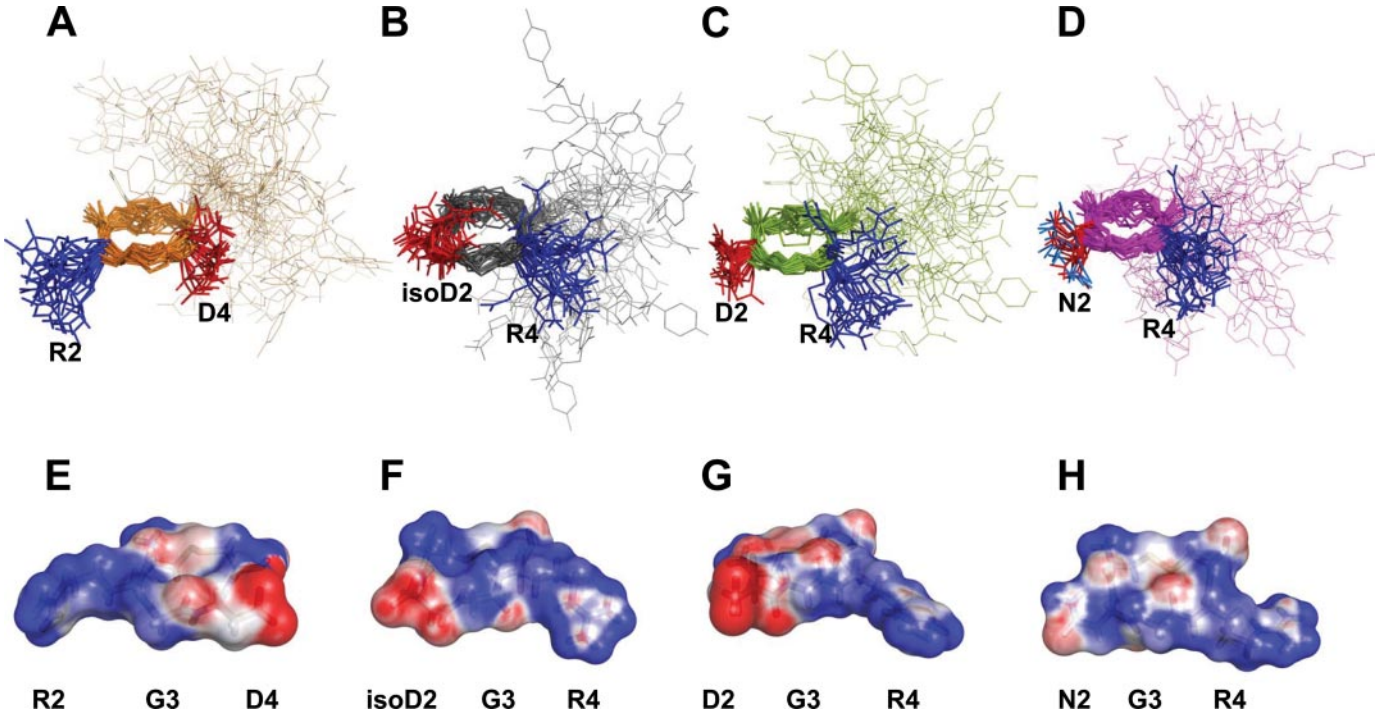

FIGURE 3. **Solution structures (*A–D*) and electrostatic surface plots of the lowest energy structure (*E–H*) of RGD-2C, *iso*DGR-2C, DGR-2C, and NGR-2C.** NMR ensembles of the 30 best structures of RGD-2C (*A*), *iso*DGR-2C (*B*), DGR-2C (*C*), and NGR-2C (*D*); the Cα atoms of the macrocycles and the disulfide bridges are represented with *sticks*; arginine and aspartate (isoaspartate) side-chains are represented in *blue* and *red*, respectively; the amino group of Asn^2 is represented in *cyan*; the side-chains of the C-terminal tails (residues 6–9) are represented with *lines*. The electrostatic surface potentials of RGD-2C (*E*), *iso*DGR-2C (*F*), DGR-2C (*G*), and NGR-2C (*H*) have been calculated using APBS; positively and negatively charged areas are represented in *blue* and *red* (*iso*-values from +5 *kT*/e to −5 *kT*/e), respectively.

metric properties. Then we used these families of structures to identify the main structural features of the systems, focusing on the few structures (centroids) that are representative of the clusters of each macrocycle. We applied a root mean square deviation criterion of 0.7 Å over the heavy atoms of the macrocycle during pairwise comparisons of individual structures. Clustering yielded 20, 32, and 45 clusters for RGD-2C, DGR-2C, and NGR-2C, respectively. Notably, applying the same cut-off for *iso*DGR clustering, we identified 89 clusters, suggesting a larger conformational variability for the *iso*DGR macrocycle, in agreement with experimental NMR results. For each macrocycle, the bundle of centroids was compared with its lowest energy NMR conformer, yielding an r.m.s.d. of ~1.75 Å over the backbone atoms and the disulfide bond, thus confirming the flexibility of the macrocycles (supplemental Table S2).

Backbone Torsion-Angle Analysis Reveals That an Inverse γ-Turn Conformation Is Highly Populated in DGR-2C and NGR-2C—The bundles of centroids representing the conformations of the four macrocycles were consistent with the corresponding NOE distance restraints, with all residues in the allowed regions of the Ramachandran plot (supplemental Table S2).

We then inspected these bundles for the presence of elements of secondary structure. Analysis of both φ and ψ angles and of hydrogen bonds provides support for the presence of an inverse γ-turn centered on residue 4 of each ligand. The inverse γ-turn was stabilized by the formation of an hydrogen bond between the amide proton of residue 5 and the carbonyl of residue 3. In addition, this inverse γ-turn recurred more frequently in both NGR-2C and DGR-2C than in RGD-2C and *iso*DGR-2C (~30–40 *versus* 10%), in agreement with the lower temperature factors of amide proton Cys^5 measured in NGR-2C and DGR-2C as compared with those observed in RGD-2C and *iso*DGR-2C (supplemental Table S1). Of note, consistent results have been obtained in the NMR solution structure analyses.

NGR-2C Has a Different Electrostatic Surface as Compared with RGD-2C, isoDGR-2C, and DGR-2C—During REMD simulations, both RGD-2C and *iso*DGR-2C adopted an extended conformation with the basic and the acidic moieties pointing in opposite directions, thus creating a highly polarized structure with the positive and negative poles located at a distance of ~13–14 Å (Fig. 3, *E–H*, and supplemental Table S3). Likewise, residues 2 and 4 of DGR-2C and NGR-2C pointed in opposite directions, although their distances were slightly shorter (~12–13 Å). Most importantly, the presence of the amino group instead of the carboxylic moiety in NGR-2C markedly changes the electrostatic surface strongly reducing the negative electrostatic contribution (Fig. 3*H*).

Overall, NMR methods and molecular dynamics simulations indicate that there are important differences among the four macrocycles. First, *iso*DGR-2C has the greatest conformational heterogeneity induced by the β-bond of the isoaspartic acid. Second, NGR-2C, presenting a prevalent positively charged surface electrostatic potential, lacks the characteristic dipolar anchor points that represent a crucial pharmacophoric requirement for $\alpha v\beta 3$ recognition (40–45).

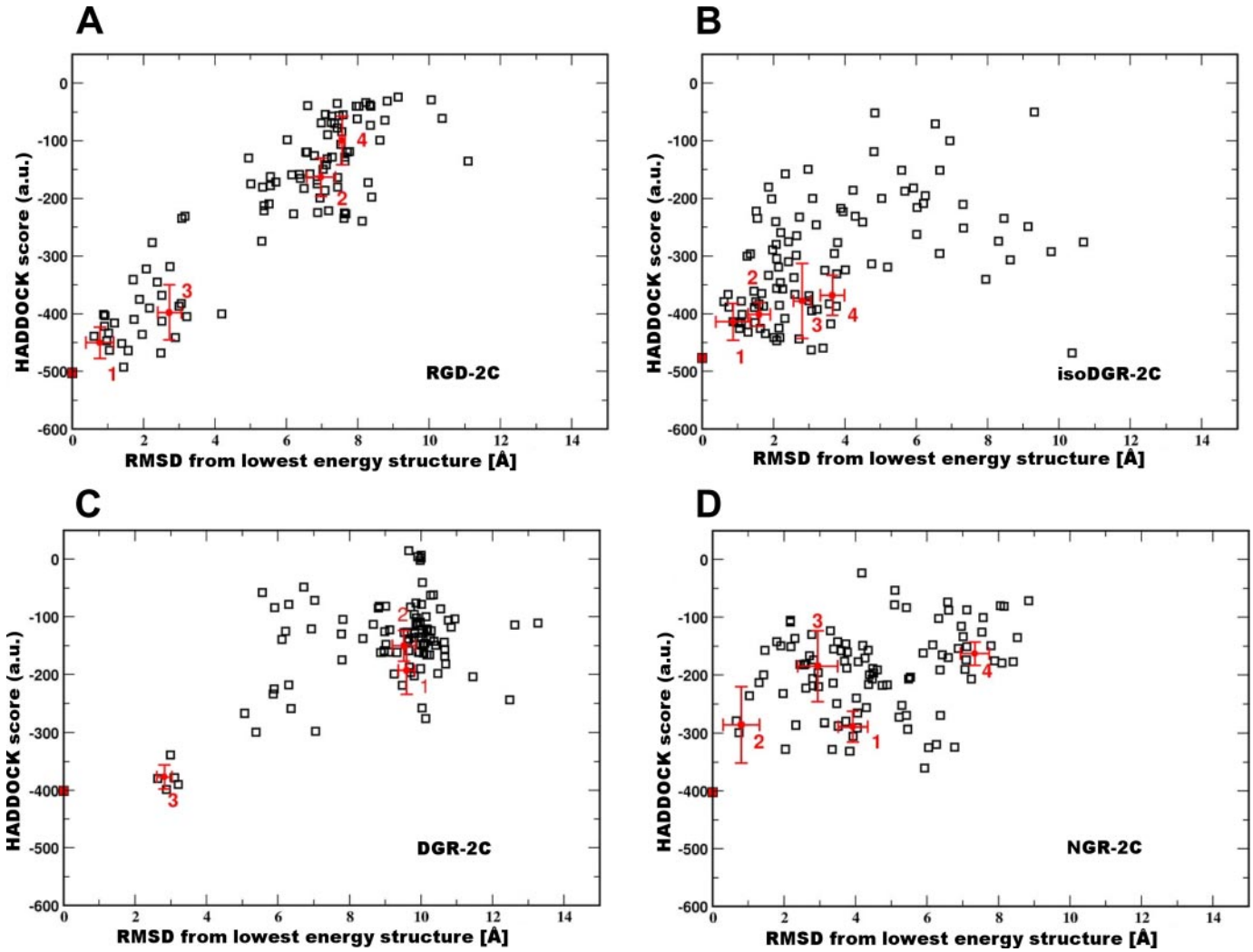

FIGURE 4. **HADDOCK score *versus* r.m.s.d. from the lowest energy complex structure for docking runs performed on RGD-2C (*A*), *iso*DGR-2C (*B*), DGR-2C (*C*), and NGR-2C (*D*).** The *open squares* correspond to the individual structures; *the red circles* correspond to the cluster averages with the standard deviations indicated by *bars*; the *red numbers* identify the clusters. Clusters were obtained first fitting the models on the backbone of the receptor-binding site and then calculating the r.m.s.d. values from the lowest energy solutions (indicated by the *red square* in each panel) over the heavy atoms of the ligand cycle. Structures belonged to the same cluster if they differed by less than 1 Å in the pairwise r.m.s.d. matrix. For NGR-2C, the cluster cutoff was increased up to 1.3 Å. The HADDOCK score corresponds to the weighted sum of different energy terms (van der Waals, electrostatic, and restraint energies) (see Table 2).

Docking Models of RGD-2C, isoDGR-2C, DGR-2C, and NGR-2C onto αvβ3—To gain a structural insight into the different binding affinities of the four macrocycles for αvβ3, we created putative binding models of the four ligands using the docking program HADDOCK2.0 (34, 35), based on the crystallographic structure of αvβ3 in complex with RGDf(NMV), a cyclic RGD pentapeptide (15, 16).

HADDOCK2.0 relies on a data-driven docking protocol consisting of a rigid-docking step followed by a semi-flexible step, which includes side-chain flexibility for both the receptor and the ligand. To account for the degree of conformational heterogeneity of a ligand, the protocol also allows for the docking of an ensemble of structures during the same run. Thus, for each ligand, the bundle of the 30 lowest energy NMR structures was docked onto the x-ray structure of αvβ3 following removal of the ligand. The program was first successfully tested for its ability to model the experimental αvβ3-RGDf(NMV) interactions, being able to reproduce the crystallographic binding mode (15) with an r.m.s.d. of ~0.5 Å, calculated on the heavy atoms of the RGD sequence (supplemental Fig. S2*A*).

Docking calculations are in agreement with data describing the different biological activity of the four macrocycles. RGD-2C and *iso*DGR-2C converged to unique low energy docking poses, contained in the mostly populated cluster (Fig. 4, *A* and *B*, and Table 2). Conversely, DGR-2C and NGR-2C failed to find highly populated binding solutions with a favorable HADDOCK score (Fig. 4, *C* and *D*, and Table 3). Noteworthy, the program generated essentially the same results when we used a bundle of structures containing conformers representative of the highest populated clusters produced by REMD simulations, thus confirming the reliability of the obtained models (supplemental Fig. S3). For RGD-2C and *iso*DGR-2C, HADDOCK calculations identified a highly populated cluster (~15 structures) with a consistent set of interactions with the receptor, showing a favorable HADDOCK score (−450 a.u.) along with an r.m.s.d. of

≤1 Å (Table 2) with respect to the lowest energy structure (Fig. 4, *A* and *B*).

RGD-2C Reproduces the Experimental αvβ3-RGDf(NMV) Interactions—Analysis of the best scored models of RGD-2C in complex with αvβ3 shows that the binding mode of the RGD motif reproduces the experimental αvβ3-RGDf(NMV) interactions (Table 3), with an r.m.s.d. of 0.25 Å calculated over the backbone atoms of the RGD sequence (supplemental Fig. S2*B*). The Arg^2 and Asp^4 side-chains point in opposite directions ($d_{CZ\text{-}CE} \sim 13.7$ Å) (supplemental Table S3), contacting specific residues of the α- and β-subunits, respectively, whereas the C-terminal tail does not contribute to the complex formation pointing outside the receptor. As shown in Fig. 5*A* and *F*, the guanidinium of Arg^2 inserts into the groove of the α-subunit, forming a salt-bridge with Asp^{150} and hydrogen bonds with both the carbonyl of Gln^{180} side chain and the carboxylate of Asp^{218}. On the opposite site, one of the carboxylate oxygen atoms contacts the ion in the metal ion-dependent adhesion site (MIDAS) region, whereas the other oxygen forms a hydrogen bond with the backbone amide of Asn^{215} and Tyr^{122}. The complex is further stabilized by backbone interactions between the amide proton of Asp^4 and the carbonyl of Arg^{216}, and between the carbonyl of Asp^4 and the guanidinium of Arg^{214}. In addition, although conformational studies on the free peptide revealed the presence of a minority of structures with an inverse γ-turn conformation, receptor-bound RGD-2C, similarly to RGDf(NMV), adopts this inverse γ-turn conformation to optimize the interaction with αvβ3.

isoDGR Docks onto αvβ3 Both Recapitulating the Classical RGD-αvβ3 Interactions and Establishing Additional Interactions—The model of *iso*DGR-2C in complex with αvβ3 shows that the ligand, although having a reversed sequence with respect to RGD-2C, preserves the characteristic polar interactions with the receptor described in RGD-2C, by means of a 180° rotation of the macrocycle inside the binding pocket (Fig. 5, *B* and *F*, and Table 3). Also in this case, the C-terminal tail is not involved in the binding. The carboxylate of the isoaspartic residue points toward the MIDAS region of the β domain, coordinating the metal ion and forming two hydrogen bonds with the backbone amide of Asn^{215} and Tyr^{122}, whereas its amide and carbonyl groups form hydrogen bonds with the carbonyl of Tyr^{122} and the guanidinium of Arg^{214}, respectively. On the opposite side, Arg^4 forms a bidentate salt bridge with Asp^{218} and an hydrogen bond with Asp^{150} and Gln^{180}, thus reproducing the interactions observed in the crystallographic complex structure (supplemental Fig. S2*C*). Unlike RGD-2C,

TABLE 2
Statistics of the clusters obtained during HADDOCK calculations

	HADDOCK energy[a]	r.m.s.d.[b]	BSA[c]	No. of structures
	a.u.	*Å*	*Å²*	
RGD-2C				
Cluster 1	−450 ± 27	0.8 ± 0.3	949 ± 96	15
Cluster 2	−163 ± 34	7.0 ± 0.4	973 ± 105	9
Cluster 3	−398 ± 48	2.7 ± 0.3	1045 ± 91	6
Cluster 4	−100 ± 42	7.6 ± 0.1	839 ± 49	6
***iso*DGR-2C**				
Cluster 1	−414 ± 32	0.9 ± 0.4	970 ± 118	13
Cluster 2	−378 ± 65	2.8 ± 0.2	998 ± 130	6
Cluster 3	−401 ± 21	1.6 ± 0.3	902 ± 89	6
Cluster 4	−368 ± 35	3.7 ± 0.3	740 ± 119	6
NGR-2C				
Cluster 1	−289 ± 26	3.9 ± 0.4	973 ± 123	7
Cluster 2	−286 ± 66	0.8 ± 0.5	890 ± 219	6
Cluster 3	−185 ± 62	2.9 ± 0.6	803 ± 164	5
Cluster 4	−163 ± 20	7.3 ± 0.4	613 ± 11	5
DGR-2C				
Cluster 1	−193 ± 42	9.6 ± 0.2	754 ± 145	14
Cluster 2	−150 ± 27	9.5 ± 0.3	728 ± 106	7
Cluster 3	−377 ± 21	2.8 ± 0.2	755 ± 87	5

[a] HADDOCK energy includes the following terms: intermolecular van der Waals and electrostatic energies, empirical desolvation energy, and the native contact analysis (34, 35). Energies are reported as mean ± S.D. over the solutions of each cluster.
[b] r.m.s.d. from the lowest energy structure was calculated over all heavy atoms of the ligand after fitting on the backbone of the protein only. Values of r.m.s.d. are reported as mean ± S.D. over the solutions of each cluster.
[c] BSA means buried surface area. Values are reported as mean ± S.D. over the solutions of each cluster.

TABLE 3
Summary of the distances (Å) between the ligands and αvβ3 observed in the highest populated clusters (cluster 1)
Distances are reported as mean ± S.D. calculated over the solutions belonging to cluster 1.

Ligand	αvβ3	RGD x-ray	RGD x-ray dock	RGD-2C	*iso*DGR-2C	NGR-2C	DGR-2C
C_{HNTerm}[a]	$Tyr^{122}{}_{O}$[b]				1.8 ± 0.1		
D_{OX}[c]/IAS_{OX}	$Tyr^{122}{}_{HN}$[d]	1.9	2.1 ± 0.1	2 ± 0.1	2.2 ± 0.1		
D_{OX}/IAS_{OX}	$Asn^{215}{}_{HN}$	1.8	2.2 ± 0.1	2.2 ± 0.1	2.3 ± 0.2		
D_{OX}/IAS_{OX}	Ca^{2+}	2.7	1.7 ± 0.1	1.7 ± 0.1	1.8 ± 0.1		
D_O	$Arg^{214}{}_{HX}$[e]	3.0	2.0 ± 0.1	1.9 ± 0.1	2.3 ± 0.1		1.9 ± 0.2
D_{HN}	$Asp^{150}{}_{OX}$						
D_{HN}	$Arg^{216}{}_{O}$	2.7	2.4 ± 0.1	2.2 ± 0.1			
$isoD_{HN}$	$Tyr^{122}{}_{O}$				2.0 ± 0.2		
N_{HX}	$Asn^{215}{}_{O}$					2.1 ± 0.2	
N_O	$Arg^{214}{}_{HX}$					1.9 ± 0.1	
N_{OX}	$Asn^{216}{}_{HN}$					2.2 ± 0.1	
G_{HN}	$Arg^{216}{}_{O}$				2.4 ± 0.1		
R_{HX}	$Asp^{150}{}_{OX}$	2.9	2.2 ± 0.2	1.9 ± 0.2	1.9 ± 0.1	1.9 ± 0.2	
R_{HX}	$Asp^{218}{}_{OX}$	1.9	1.7 ± 0.1	1.8 ± 0.2	2.0 ± 0.4	2.0 ± 0.3	
R_{HX}	$Glu^{180}{}_{OX}$	2.7	1.9 ± 0.2	1.7 ± 0.1	2.0 ± 0.3		
R_{HX}	$Asn^{215}{}_{O}$						1.8 ± 0.1
R_{HX}	$Arg^{216}{}_{O}$						2.3 ± 0.1
R_{HX}	$Tyr^{122}{}_{O}$						2.0 ± 0.2
R_{HN}	$Asp^{218}{}_{OX}$						2.4 ± 0.1

[a] HNTerm denotes N-terminal protons.
[b] O denotes carbonyl oxygen.
[c] OX denotes carboxylate oxygen.
[d] HN denotes amide proton.
[e] HX denotes guanidinium proton.

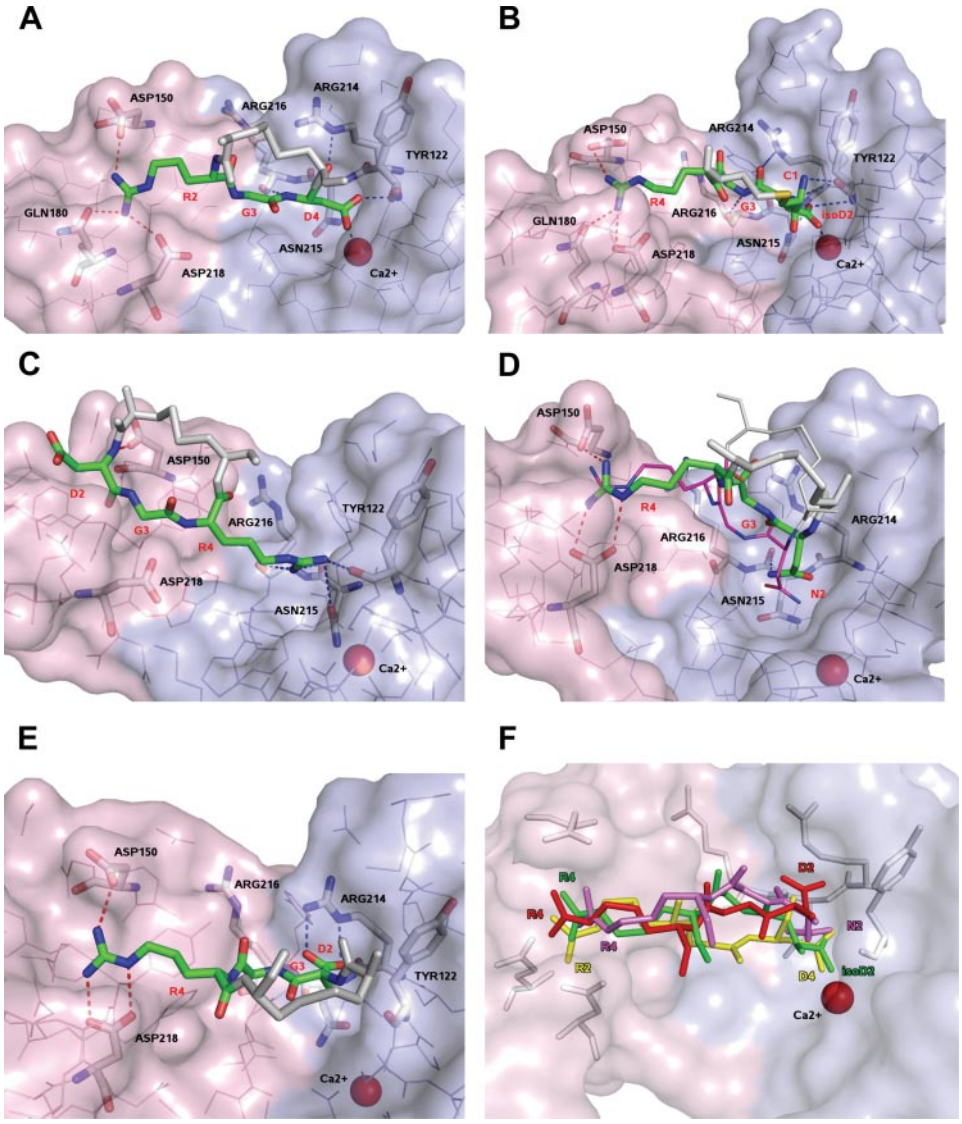

FIGURE 5. **HADDOCK models of ligand-integrin-binding site.** Surface representation of αvβ3 binding pocket in complex with RGD-2C (*A*), *iso*DGR-2C (*B*), DGR-2C (*C*), and NGR-2C (*D*), with the α- and β-subunits represented in *pink* and *pale cyan*, respectively. In each panel, the side-chains of the ligand interacting with the receptor are shown in *green*, with nitrogen, oxygen, and sulfur atoms in *blue*, *red*, and *yellow*, respectively; the disulfide bond is represented with *gray sticks*; the MIDAS cation is represented with a *red sphere*. The side chains of both αvβ3 and the ligand directly involved in the binding are labeled with the three- and one-letter code, respectively. *Red* and *blue dotted lines* denote the hydrogen bonds of the ligand with α- and β-subunits, respectively. The models (*A–C*) of the binding site correspond to the best (lowest energy) HADDOCK score structures of the highest populated cluster (see corresponding cluster 1 in each panel of Fig. 4). *D,* because HADDOCK calculations on NGR-2C produced two clusters with both similar score and populations (cluster 1 and 2 in Fig. 4*D*), the best structure of cluster 2 (*thin sticks* with NGR residues in *magenta*) is superimposed on the best structure of cluster 1 to allow comparison. *E,* model of DGR-2C in complex with αvβ3 as defined in cluster 3 of Fig. 4. The pose is rotated by 180° with respect to cluster 1, and it is poorly populated and lacks the fundamental stabilizing interactions with the metal ion. *F,* superposition of HADDOCK structures of RGD-2C (*yellow*, cluster 1), *iso*DGR-2C (*green*, cluster 1), NGR-2C (*purple*, cluster 1), and DGR-2C (*red*, cluster 3) in complex with αvβ3. The carboxyl group of both RGD-2C and *iso*DGR-2C coordinate the metal ion, whereas the carboxyl- and amino-group of DGR-2C and NGR-2C, respectively, move away from the MIDAS site.

receptor-bound *iso*DGR-2C does not adopt an inverse γ-turn conformation to interact with αvβ3.

It is worth noting that *iso*DGR motif displays some additional interactions with αvβ3 as compared with RGD sequence. First, the *iso*DGR complex is also stabilized by an hydrogen bond between the backbone amide of Gly^{3} and the carbonyl of Arg^{216} (Fig. 5*B* and Table 3). As a matter of fact, the *iso*DGR central glycine contributes to the recognition of the receptor via polar interactions, while both in the crystallographic structure and in the αvβ3-RGD-2C model the RGD central glycine interacts with the protein only through weak hydrophobic interactions (Fig. 5*A* and Table 3). Second, the reversed orientation inside the binding pocket of *iso*DGR allows the formation of an additional interaction between the carbonyl of Tyr^{122} and the N terminus of Cys^{1} (Fig. 5*B* and Table 3). Yet this interaction is missing both in the crystallographic structure and in the αvβ3-RGD-2C model, suggesting that the N-terminal residue flanking the *iso*DGR motif might confer binding specificity.

Our model clearly shows that *iso*DGR-2C fulfills the pharmacophoric requirements for an αvβ3 ligand, adopting both the proper distance (~13.7 Å) (supplemental Table S3) and the correct orientation between the two charged groups, which can simultaneously bind the α- and β-subunits. In this context, it is of note that the β-bond of the isoaspartic residue, with the consequent insertion of an additional torsion angle (Fig. 1*B* and supplemental Fig. S1) inside the main-chain, confers to the backbone and the side-chains of the ligand the necessary flexibility to adapt well inside the crevice, thus optimizing the backbone and side-chain interactions with the receptor.

DGR-2C Fails to Identify a Favorable Position Inside the Binding Pocket—HADDOCK calculations performed on DGR-2C failed to generate highly populated low energy solutions (Fig. 4*C* and Table 2), clearly indicating the inability of DGR-2C to favorably interact with αvβ3 (Fig. 5*C* and Fig. 5*F*). The highest populated cluster (cluster 1) displays higher HADDOCK energy than clusters found in both RGD-2C and *iso*DGR-2C (approximately −200 a.u.), and differs by more than 9 Å from the lowest energy structure. Remarkably, in cluster 1, DGR-2C was docked onto the receptor with an opposite orientation as compared with *iso*DGR-2C, thus losing all the important stabilizing interactions of the complex, including the metal ion coordination (Table 3). In agreement with its unfavorable HADDOCK score, DGR-2C creates only superficial interactions as indicated by its

lower buried surface area (~780 Å^2) in comparison with the *iso*DGR and RGD-2C ones (~950 Å^2). In attempting the binding with $\alpha v\beta 3$, DGR-2C maintains the inverse γ-turn conformation observed in NMR and molecular dynamics (MD) studies along with the proper distance (~14 Å) between the charged groups. However, the reduced number of interactions with the receptor suggests that the macrocycle backbone conformation and the side-chain orientations are not appropriate to accommodate in the binding crevice of $\alpha v\beta 3$.

Furthermore, we identified few poses (rotated by 180° with respect to cluster 1) in cluster 3 where the ligand penetrates inside the binding crevice orienting Asp^2 and Arg^4 toward the β-subunit and the α-subunit, respectively (Fig. 5, *E–F*). However, this cluster differed by 2.5 Å from the lowest energy complex structure and was poorly populated (five structures), suggesting that this binding mode has low probability to occur. In fact, in this cluster, the macrocycle has to change its backbone conformation to fit inside the binding pocket, breaking the γ-turn and strongly reducing the distance between Asp^2 and Arg^4 (~11 Å). Therefore, these data provide further support for the hypothesis that DGR-2C does not have the correct stereochemistry and the necessary flexibility to accommodate inside $\alpha v\beta 3$.

Overall, our results from both cell-adhesion experiments and docking performed on DGR-2C are consistent with previous binding and conformational studies performed on retro-peptides. The latter clearly indicated a dramatic decrease in activity for the DGR sequence as compared with the parent compound structure containing the RGD motif, ascribing this decrease in activity to the inappropriate amide bond directions and side-chain topology adopted by the retro-molecules (46).

NGR-2C Fails to Converge to Unique Binding Solutions—HADDOCK calculations on NGR-2C failed to identify unique docking solutions inside $\alpha v\beta 3$, clearly showing poor convergence of the docking poses, with significantly higher energies (−200 a.u.) than those calculated for RGD-2C and *iso*DGR-2C (Fig. 4*D* and Table 2). Because we could not identify any cluster with a 1-Å cut-off, we increased it up to 1.3 Å, obtaining only poorly populated clusters (less than seven poses per cluster, see Table 2) with two opposite orientations inside the binding pocket (Fig. 4*D*, clusters 1–3 *versus* cluster 4). Moreover, in the first two clusters, displaying both similar HADDOCK energy and populations, the macrocycle adopted different backbone conformations inside the binding pocket (r.m.s.d. ~ 3.5 Å), providing further support for the inability of the NGR motif to specifically recognize $\alpha v\beta 3$ (Fig. 5, *D* and *F*). Consistently, the poses contained in clusters 1 and 2 (Fig. 4*D*) displayed a reduced pattern of interactions, although adopting the same orientation as *iso*DGR-2C inside the binding pocket. On one end, the Arg^4 guanidinium group maintains its interactions with Asp^{150} and Asp^{218}. On the other end, the positive electrostatic surface of the N-2 amino group causes repulsive interactions with the metal ion and with the side-chains of both Asn^{215} and Arg^{214}, thus forcing the Asn^2 amino group to move away from the MIDAS site (Fig. 5*D*), strongly destabilizing the NGR-2C-receptor complex.

Similarly to what observed for DGR-2C, in cluster 4 (Fig. 4*D*) the macrocycle does not maintain the γ-turn when bound to $\alpha v\beta 3$ and the distance between Asn^2 and Arg^4 is significantly reduced (~11 Å) as compared with the distances measured in RGD-2C and *iso*DGR-2C (supplemental Table S3). This lends support to the hypothesis that the macrocycle backbone conformation and the side-chain orientations are not appropriate to optimize the ligand-receptor interactions. Therefore, our data provide a structural rationale to explain the inability of NGR-2C to dock onto $\alpha v\beta 3$.

DISCUSSION

The main finding of this work is that *iso*DGR-2C has the ability to properly interact with $\alpha v\beta 3$, both recapitulating the canonical RGD-$\alpha v\beta 3$ contacts and establishing additional polar interactions, as demonstrated by both functional and structural studies, thus defining a new natural $\alpha v\beta 3$ recognition motif that does not comply with the strict RGD sequence requirement rule.

To this extent, it is well established that $\alpha v\beta 3$ integrin binds RGD motif-containing proteins, such as fibronectin, vitronectin, fibrillin-1, von Willebrand factor, osteopontin, and echistatin (47–52). Generally, the integrin recognition mechanism follows strict sequence and stereospecific requirements, as characterized in detail throughout both functional and structural studies (15, 52–54). In contrast, here we show that this strict RGD sequence requirement rule can be bent. In fact, our biological data indicate that *iso*DGR-2C and RGD-2C bind to $\alpha v\beta 3$ with similar affinity, and that these interactions can occur also in living cells, because *iso*DGR-2C and RGD-2C are also able to inhibit cell adhesion of endothelial cells. Consistently, we have recently shown that fibronectin can interact with $\alpha v\beta 3$ not only via the classical RGD mechanism but also through an *iso*DGR site in the FN-I$_5$ module, generated by the deamidation of an NGR sequence (12). Likewise, cells from homozygous knock-in mice carrying the RGD → RGE mutation in fibronectin can be assembled into fibrils *in vivo* and *in vitro* via $\alpha v\beta 3$ binding to the *iso*DGR motif in FN-I$_5$, generated by NGR deamidation (13).

Our three-dimensional model of *iso*DGR-$\alpha v\beta 3$ interactions indicates that this binding is dictated by a novel and stereospecific recognition pattern, offering a structural rationale for $\alpha v\beta 3$ binding specificity of deamidated-Asn^{263} FN-I$_5$. Here, a key finding of our structural studies is that *iso*DGR docks onto $\alpha v\beta 3$ in an inverted orientation as compared with the RGD ligand. This orientation allows *iso*DGR to anchor the α and β domains like an electrostatic clamp through the isoaspartic and arginine side-chains, reproducing the canonical interactions characterizing the $\alpha v\beta 3$ recognition via RGD (15). Furthermore, the acidic and basic residues of isoaspartate and arginine are at the correct distance (13.72 Å, supplemental Table S3) and orientation to engage stabilizing interactions with the polar region of the receptor, which is a crucial pharmacophoric requirement of an $\alpha v\beta 3$ ligand. Most importantly, the reversed orientation of *iso*DGR-2C inside the binding pocket allows for additional interactions as compared with RGD-2C and RGDf(NMV), involving the N-terminal cysteine flanking the *iso*DGR sequence and the central glycine (Table 3). In particular, *iso*DGR-2C cysteine 1 points toward the binding site contributing to receptor recognition, whereas in RGD-2C it faces

out from the receptor. Furthermore, the backbone atoms of *iso*DGR-2C central glycine engage additional polar interactions, whereas in the RGD motif the backbone adopts an inverse γ turn conformation centered on the glycine, thus contributing to αvβ3 recognition only through weak hydrophobic contacts.

In this context, it is worth noting that the presence of the β-bond in *iso*DGR-2C confers to the macrocycle a high degree of backbone flexibility, as shown by both NMR analysis and MD simulations. Therefore, the ligand can easily accommodate inside the shallow crevice of the receptor orienting the isoaspartate toward the MIDAS cation, which is a central anchor point for the ligand, thus completing its coordination sphere. Under this view, the greatest conformational heterogeneity of *iso*DGR-2C induced by the β-bond of the isoaspartic acid is conceivably a favorable feature. As a matter of fact, the entropic penalty paid by the flexible *iso*DGR macrocycle upon binding to αvβ3 is largely compensated by the stabilizing backbone and side-chains interactions engaged with the receptor. Notably, our results are in agreement with rational drug design, because peptidomimetics employing β-amino acids mimicking the aspartate residue are potent antagonists of αvβ3 (45, 55), thus suggesting that the conformational plasticity of the β-bond favors the binding to the receptor exploiting the adaptation ability of ligand-receptor interaction.

Both NMR and MD studies on DGR-2C provide further support for the importance of the β-bond in contributing to ligand binding. As a matter of fact, DGR-2C has a reduced conformational freedom of the cycle and does not have the correct stereochemistry to interact with the receptor. Thus, to fit inside the receptor binding groove DGR-2C should reduce the distance between the arginine and the aspartic residues breaking the stable inverse γ-turn, and therefore paying a highly unfavorable energy penalty. These results along with data from both binding and cell adhesion experiments are consistent with previous binding and conformational studies performed on retropeptides, indicating a dramatic decrease in activity for the DGR sequence as compared with the parent compound structure containing the RGD motif (46). This decrease in activity can be ascribed to the inappropriate amide bond directions and side-chain topology adopted by the retro-molecules, consistently with the strict RGD sequence requirement rule.

Another important finding of this work is that NGR-2C is not an αvβ3 recognition motif because it lacks the fundamental pharmacophoric features for high receptor affinity, as shown by both NMR and MD studies in agreement with functional results from binding and cell adhesion experiments. Similarly to DGR-2C, NGR-2C adopts an unfavorable backbone conformation to accommodate inside the binding groove, and most importantly it misses the acidic group, which completes the coordination of the divalent metal ion of the MIDAS site. In addition, NGR-2C presents a prevalent positively charged surface electrostatic potential and thus lacks the characteristic dipolar anchor points that represent a crucial pharmacophoric requirement for αvβ3 recognition (40–45). Altogether, these data provide a structural rationale to explain the inability of NGR-2C to bind to αvβ3. Consistently, in agreement with previous reports (12, 13), functional studies, based on inhibition of endothelial cell adhesion and *in vitro* binding of peptides to purified αvβ3 integrin, suggest that NGR neither affects cell adhesion nor competes the binding to αvβ3, whereas the heat-induced NGR → *iso*DGR deamidation causes a marked increase in binding (>100-fold).

In this context, we therefore argue that NGR → *iso*DGR transition may represent an important physiological mechanism able to work as a molecular timer for activating latent integrin-binding sites in fibronectin (12). Furthermore, the NGR motif by itself binds to aminopeptidase N (CD13) (56, 57), which affects major biological events, including cell proliferation, secretion, invasion, and angiogenesis, thus participating in the regulation of tumorigenesis. In particular, aminopeptidase N is specifically expressed in endothelial and subendothelial cells in angiogenesis and in various types of tumors (56–59). An interesting possibility is that the NGR motif contained in fibronectin interacts with CD13 on one hand, and on the other hand gives origin to a natural recognition motif (*iso*DGR) of αvβ3 integrin, conferring to fibronectin multifunctionality and plasticity in mediating a variety of physiological and pathological processes, including development, angiogenesis, inflammation, wound repair, and tumor growth.

In conclusion, our data provide the structural basis to define *iso*DGR as a new natural recognition motif of αvβ3, which indicates that NGR lacks the necessary pharmacophoric requirements to bind αvβ3, thus helping to explain the functional differences of fibronectin NGR and *iso*DGR sequences.

Acknowledgment—We thank Dr. Massimo Degano for critical reading of the manuscript.

REFERENCES

1. Humphries, M. J., Obara, M., Olden, K., and Yamada, K. M. (1989) *Cancer Investig.* **7,** 373–393
2. Pankov, R., and Yamada, K. M. (2002) *J. Cell Sci.* **115,** 3861–3863
3. Mohri, H. (1997) *Peptides* (*N. Y.*) **18,** 899–907
4. Hynes, R. O. (1992) *Cell* **69,** 11–25
5. Kornblihtt, A. R., Pesce, C. G., Alonso, C. R., Cramer, P., Srebrow, A., Werbajh, S., and Muro, A. F. (1996) *FASEB J.* **10,** 248–257
6. Yamada, K. M. (1989) *Curr. Opin. Cell Biol.* **1,** 956–963
7. Johansson, S., Svineng, G., Wennerberg, K., Armulik, A., and Lohikangas, L. (1997) *Front. Biosci.* **2,** 126–146
8. Plow, E. F., Haas, T. A., Zhang, L., Loftus, J., and Smith, J. W. (2000) *J. Biol. Chem.* **275,** 21785–21788
9. Humphries, M. J. (2000) *Biochem. Soc. Trans.* **28,** 311–339
10. Arnaout, M. A. (1990) *Immunol. Rev.* **114,** 145–180
11. Di Matteo, P., Curnis, F., Longhi, R., Colombo, G., Sacchi, A., Crippa, L., Protti, M. P., Ponzoni, M., Toma, S., and Corti, A. (2006) *Mol. Immunol.* **43,** 1509–1518
12. Curnis, F., Longhi, R., Crippa, L., Cattaneo, A., Dondossola, E., Bachi, A., and Corti, A. (2006) *J. Biol. Chem.* **281,** 36466–36476
13. Takahashi, S., Leiss, M., Moser, M., Ohashi, T., Kitao, T., Heckmann, D., Pfeifer, A., Kessler, H., Takagi, J., Erickson, H. P., and Fassler, R. (2007) *J. Cell Biol.* **178,** 167–178
14. Eliceiri, B. P., and Cheresh, D. A. (1999) *J. Clin. Invest.* **103,** 1227–1230
15. Xiong, J. P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S. L., and Arnaout, M. A. (2002) *Science* **296,** 151–155
16. Dechantsreiter, M. A., Planker, E., Matha, B., Lohof, E., Holzemann, G., Jonczyk, A., Goodman, S. L., and Kessler, H. (1999) *J. Med. Chem.* **42,** 3033–3040
17. Curnis, F., Gasparri, A., Sacchi, A., Cattaneo, A., Magni, F., and Corti, A. (2005) *Cancer Res.* **65,** 2906–2913

18. Traversari, C., van der Bruggen, P., van den Eynde, B., Hainaut, P., Lemoine, C., Ohta, N., Old, L., and Boon, T. (1992) *Immunogenetics* **35,** 145–152
19. van Well, R. M., Marinelli, L., Altona, C., Erkelens, K., Siegal, G., van Raaij, M., Llamas-Saiz, A. L., Kessler, H., Novellino, E., Lavecchia, A., van Boom, J. H., and Overhand, M. (2003) *J. Am. Chem. Soc.* **125,** 10822–10829
20. Hwang, T. L., and Shaka, A. J. (1995) *J. Magn. Reson. Ser. A* **112,** 275–279
21. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
22. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
23. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) *Methods Enzymol.* **339,** 71–90
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
25. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
26. Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) *Comp. Phys. Comm.* **91,** 43–56
27. Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) *J. Am. Chem. Soc.* **118,** 11225–11236
28. Hermans, J., Berendsen, H. J. C., van Gunsteren, W. F., and Postma, J. P. M. (1984) *Biopolymers* **23,** 1513–1518
29. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) *J. Comput. Chem.* **18,** 1463–1472
30. Darden, T., Perera, L., Li, L., and Pedersen, L. (1999) *Structure* (*Camb.*) **7,** 55–60
31. Sugita, Y., and Okamoto, Y. (1999) *Chem. Phys. Lett.* **1–2,** 141–151
32. Daura, X., Antes, I., van Gunsteren, W. F., Thiel, W., and Mark, A. E. (1999) *Proteins* **36,** 542–555
33. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10037–10041
34. de Vries, S. J., van Dijk, A. D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V., Wassenaar, T., and Bonvin, A. M. (2007) *Proteins* **69,** 726–733
35. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) *J. Am. Chem. Soc.* **125,** 1731–1737
36. McLachlan, A. (1982) *Acta Crystallogr. Sect. A* **38,** 871–873
37. Fernandez-Recio, J., Totrov, M., and Abagyan, R. (2004) *J. Mol. Biol.* **335,** 843–865
38. Hansmann, U. H., and Okamoto, Y. (1997) *Phys. Rev.* **56,** 2228–2233
39. Hukushima, K., and Nemoto, K. (1996) *J. Phys. Soc. Japan* **65,** 1604–1608
40. Belvisi, L., Riccioni, T., Marcellini, M., Vesci, L., Chiarucci, I., Efrati, D., Potenza, D., Scolastico, C., Manzoni, L., Lombardo, K., Stasi, M. A., Orlandi, A., Ciucci, A., Nico, B., Ribatti, D., Giannini, G., Presta, M., Carminati, P., and Pisano, C. (2005) *Mol. Cancer Ther.* **4,** 1670–1680
41. Casiraghi, G., Rassu, G., Auzzas, L., Burreddu, P., Gaetani, E., Battistini, L., Zanardi, F., Curti, C., Nicastro, G., Belvisi, L., Motto, I., Castorina, M., Giannini, G., and Pisano, C. (2005) *J. Med. Chem.* **48,** 7675–7687
42. Heckmann, D., and Kessler, H. (2007) *Methods Enzymol.* **426,** 463–503
43. Marinelli, L., Lavecchia, A., Gottschalk, K. E., Novellino, E., and Kessler, H. (2003) *J. Med. Chem.* **46,** 4393–4404
44. Moitessier, N., Henry, C., Maigret, B., and Chapleur, Y. (2004) *J. Med. Chem.* **47,** 4178–4187
45. Heckmann, D., Meyer, A., Marinelli, L., Zahn, G., Stragies, R., and Kessler, H. (2007) *Angew. Chem. Int. Ed. Engl.* **46,** 3571–3574
46. Wermuth, J., Goodman, S. L., Jonczyk, A., and Kessler, H. (1997) *J. Am. Chem. Soc.* **119,** 1328–1335
47. Chen, F. H., Thomas, A. O., Hecht, J. T., Goldring, M. B., and Lawler, J. (2005) *J. Biol. Chem.* **280,** 32655–32661
48. Ensslin, M. A., and Shur, B. D. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104,** 2715–2720
49. Kiyozumi, D., Osada, A., Sugimoto, N., Weber, C. N., Ono, Y., Imai, T., Okada, A., and Sekiguchi, K. (2005) *Exp. Cell Res.* **306,** 9–23
50. Osada, A., Kiyozumi, D., Tsutsui, K., Ono, Y., Weber, C. N., Sugimoto, N., Imai, T., Okada, A., and Sekiguchi, K. (2005) *Exp. Cell Res.* **303,** 148–159
51. Steck, E., Braun, J., Pelttari, K., Kadel, S., Kalbacher, H., and Richter, W. (2007) *Matrix Biol.* **26,** 30–41
52. Takagi, J. (2007) *Curr. Opin. Cell Biol.* **19,** 557–564
53. Gottschalk, K. E., and Kessler, H. (2002) *Angew. Chem. Int. Ed. Engl.* **41,** 3767–3774
54. Xiao, T., Takagi, J., Coller, B. S., Wang, J. H., and Springer, T. A. (2004) *Nature* **432,** 59–67
55. Scarborough, R. M. (1999) *Curr. Med. Chem.* **6,** 971–981
56. Curnis, F., Arrigoni, G., Sacchi, A., Fischetti, L., Arap, W., Pasqualini, R., and Corti, A. (2002) *Cancer Res.* **62,** 867–874
57. Pasqualini, R., Koivunen, E., Kain, R., Lahdenranta, J., Sakamoto, M., Stryhn, A., Ashmun, R. A., Shapiro, L. H., Arap, W., and Ruoslahti, E. (2000) *Cancer Res.* **60,** 722–727
58. Luan, Y., and Xu, W. (2007) *Curr. Med. Chem.* **14,** 639–647
59. Sato, Y. (2004) *Biol. Pharm. Bull.* **27,** 772–776