LIGAND RECOGNITION SPECIFICITY OF LEUKOCYTE INTEGRIN $\alpha_M\beta_2$ (Mac-1, CD11b/CD18) AND ITS FUNCTIONAL CONSEQUENCES

Nataly P. Podolnikova[‡], Andriy V. Podolnikov[‡], Thomas A. Haas[¶], Valeryi K. Lishko[‡], and Tatiana P. Ugarova[‡ §]

From the School of Life Sciences, the Center for Metabolic and Vascular Biology, Arizona State University, Tempe, AZ 85287, and [¶]University of Saskatchewan, Saskatoon, Canada

Running title: Integrin $\alpha_M\beta_2$ recognition specificity

[§] To whom correspondence should be addressed: Tatiana P. Ugarova, the Center for Metabolic and Vascular Biology, Arizona State University, PO Box 874501, Tempe, AZ 85287; Tel: (480)-301-4235; E-mail: Tatiana.Ugarova@asu.edu

Keywords: Integrins; integrin $\alpha_M\beta_2$ (Mac-1, CD11b/CD18); cell adhesion; protein motifs; antimicrobial peptides; LL-37; alarmins; host defense; innate immunity

ABSTRACT

The broad recognition specificity exhibited by integrin $\alpha_M\beta_2$ (Mac-1, CD11b/CD18) has allowed this adhesion receptor to play innumerable roles in leukocyte biology, yet we know little how and why $\alpha_M\beta_2$ binds its multiple ligands. Within $\alpha_M\beta_2$, the α_MI-domain is responsible for integrin's multiligand binding properties. To determine its recognition motif, we screened peptide libraries spanning sequences of many known protein ligands for the α_MI-domain binding and also selected the α_MI-domain recognition sequences by phage display. Analyses of >1400 binding and non-binding peptides derived from peptide libraries showed that a key feature of the α_MI-domain recognition motif is a small core consisting of basic amino acids flanked by hydrophobic residues. Furthermore, the peptides selected by phage display conformed to a similar pattern. Identification of the recognition motif allowed the construction of an algorithm which reliably predicts the α_MI-domain binding sites in the $\alpha_M\beta_2$ ligands. The recognition specificity of the α_MI-domain resembles that of some chaperones which enables it to bind segments exposed in unfolded proteins. The disclosure of the $\alpha_M\beta_2$ binding preferences allowed the prediction that cationic host defense peptides, which are strikingly enriched in the α_MI-domain recognition motifs, represent a new class of $\alpha_M\beta_2$ ligands. This prediction has been tested by examining the interaction of $\alpha_M\beta_2$ with the human cathelicidin peptide LL-37. LL-37 induced a potent $\alpha_M\beta_2$-dependent cell migratory response and caused activation of $\alpha_M\beta_2$ on neutrophils. The newly revealed recognition specificity of $\alpha_M\beta_2$ towards unfolded protein segments and cationic proteins/peptides suggests that $\alpha_M\beta_2$ may serve as a previously proposed "alarmin" receptor with important roles in innate host defense.

Integrins are non-covalently associated α-β heterodimer receptors that mediate adhesive interactions of cells with the extracellular matrix and other cells. Integrins regulate a diverse range of processes including cell migration, differentiation, the immune response, and maintenance of tissue architecture. Many integrins exhibit a very broad ligand binding specificity and can bind various proteins that share no obvious sequence similarity. Furthermore, even integrins that selectively recognize the RGD adhesion motif are capable of binding numerous ligands that lack this sequence and belong to diverse protein families. To date, the mechanisms underlying the broad ligand specificity exhibited by integrins remain unknown.

$\alpha_M\beta_2$ (Mac-1, CD11b/CD18), which belongs to the β_2 subfamily of leukocyte integrins, is the most promiscuous integrin with more than 40 reported protein ligands. $\alpha_M\beta_2$ is expressed predominantly on myeloid cells and mediates adhesive reactions of leukocytes during the inflammatory response. In particular, it contributes to firm adhesion of neutrophils to endothelial cells, promotes their diapedesis, and participates in neutrophil migration to sites of inflammation (1-3). Many other neutrophil and monocyte/macrophage responses, including phagocytosis, homotypic aggregation, degranulation, and adherence to microorganisms also depend on $\alpha_M\beta_2$. The complexity of $\alpha_M\beta_2$-mediated functions is believed to arise from its ability to recognize a multitude of structurally and functionally dissimilar ligands. The reported $\alpha_M\beta_2$ ligands include numerous proteins that constitute the extracellular matrix and many that become associated with the ECM during the inflammatory response (a partial list is provided in (4,5). It also binds several cellular receptors such as ICAM-1, GPIbα, and JAM-3 (6-8). Further adding to the diversity, several proteases, such as elastase, myeloperoxidase, and plasminogen (9-11), and even the non-mammalian proteins ovalbumin and keyhole limpet hemocyanin, are $\alpha_M\beta_2$ ligands. Each year, new and structurally unrelated proteins are being added to this already impressive list. Particularly notable additions are CD40L, which may contribute to the pathogenesis of atherosclerosis (12),

and HMGB1 (high-mobility group box 1, amphoterin) (13), a nuclear protein that is released from necrotic cells and activated macrophages and recently emerged as a potent inflammatory mediator (14-16). The mechanism by which this single integrin can recognize such a vast repertoire of structurally unrelated proteins, the biological significance of $\alpha_M\beta_2$'s broad specificity, and the physiological relevance of many identified ligands remain poorly understood.

The approximately 200-residue α_MI domain within $\alpha_M\beta_2$ mediates ligand binding (17,18) and is therefore responsible for the receptor's broad substrate specificity. Accordingly, binding sites for several ligands, including C3bi, fibrinogen, neutrophil inhibitory factor (NIF), and CCN1 (Cyr61) have been localized to the α_MI-domain (19-22). Earlier studies have identified peptides derived from several proteins that bind the α_MI-domain, including the fibrinogen peptides ^{190}GWTVFQKRLDGS^{202} (P1), ^{377}YSMKKT TMKIIPFNRLTIG^{395} (P2) and CCN1-derived ^{305}SSVKKYRPKYCGS^{317} (23-25), as putative binding sites for $\alpha_M\beta_2$. The α_MI-domain–binding peptides directly support cell adhesion, inhibit $\alpha_M\beta_2$-mediated cell adhesion and are able to promote cell migration (24-26). In addition to these sequences, the α_MI-domain can bind other sequences in fibrinogen and CCN1, consistent with the existence of multiple binding sites for $\alpha_M\beta_2$ in these molecules (27). The α_MI-domain recognition sequences derived from fibrinogen and other $\alpha_M\beta_2$ ligands do not contain a particular consensus motif similar to RGD and have no apparent sequence homology. However, the fact that all these peptides bind the α_MI-domain and support $\alpha_M\beta_2$-mediated adhesion responses implies they contain a similar recognition signal. To identify this recognition signal, we previously used cellulose-bound peptide libraries to analyze a set of α_MI-domain binding sequences derived from the γC and βC domains of fibrinogen (27). This approach was widely utilized to define the mechanisms of recognition in biological systems in which promiscuity in ligand binding plays an important role including molecular chaperones (28-30). We demonstrated that the α_MI domain binds short sequences enriched in basic and

hydrophobic residues (27). However, although these analyses provided useful insights into the $\alpha_M\beta_2$ ligand binding preferences, the limited dataset was not sufficient to solve the α_MI-domain recognition motif.

To gain understanding of the principles that govern the multiligand binding properties of $\alpha_M\beta_2$, we have screened peptide libraries representing complete sequences of many known and predicted $\alpha_M\beta_2$ ligands for the α_MI-domain binding and selected recognition sequences by phage display. Analyses of a large data set allowed the identification of the α_MI-domain recognition motif which satisfactorily explained previous findings and led to important insights into functional consequences of the $\alpha_M\beta_2$ recognition specificity.

EXPERIMENTAL PROCEDURES

Proteins and Peptides

The active conformer of α_MI-domain (residues α_M Glu^{123}-Lys^{315}) was prepared with or without the GST fusion as described previously (21). The α_MI-domain without fusion was labeled with ^{125}I using IODO-GEN (Pierce, Rockford, IL). The D fragment (100 kDa) was prepared from human fibrinogen (Enzyme Research Laboratories, South Bend, IN) by digestion with plasmin as described (31). The peptides for binding experiments with the α_MI-domain (RKLRSLWRR, LQLRFPRFV, LLHNYGVYT, GDDPSDKFF, QVLRIRKRA, ARLPIWF, GRLPMPW, and NRLLLTG) were synthesized according to standard Fmoc machine protocols using an Omega 396 synthesizer (Advanced ChemTech, Louisville, KY) and analyzed by reverse-phase HPLC and mass spectrometry. The human cathelicidin peptide LL-37 (LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES) was from AnaSpec, Inc (San Jose, CA). The fibrinogen peptides TMKIIPFNRLIG (P2-C), GWTVFQKRLDGSV (P1), KYRLTYAYFAG, and SVNKYRGTAGNA were described previously (27). The mAbs 44a and IB4 directed against human α_M and β_2 integrin subunits respectively were purified from conditioned media of hybridoma cells obtained from American Tissue Culture Collection (Manassas, VA) using protein A agarose. The mAb CBRM1/5 conjugated to Alexa 488 was from Santa Cruz Biotechnology (Dallas, TX).

Synthesis and Screening of Peptide Libraries for α_MI-domain Binding

Peptide libraries representing the compete sequences of 15 proteins and 8 antimicrobial peptides (Table 1) were prepared by parallel spot synthesis using cellulose membranes as described previously (32). Protein amino acid sequences were obtained from the NCBI database using the corresponding accession numbers (Table 1). The peptide libraries of the γC and βC domains of fibrinogen were described previously (27). The libraries were synthesized as 9-mer overlapping

peptides with 3-amino acid offset. Peptides were C-terminally attached to the cellulose via a $(\beta\text{-Ala})_2$ spacer and were acetylated N-terminally. The membrane-bound peptides were tested for the ability to bind the α_MI-domain essentially as described (27). In brief, membranes were blocked with 1% BSA and incubated with 5 μg/ml ^{125}I-labeled α_MI-domain in TBS containing 1 mM $MgCl_2$, 0.1% BSA, and 2 mM dithiothreitol. Membranes were washed with TBS containing 0.05% Tween 20, dried, and α_MI-domain binding was visualized by autoradiography and analyzed by densitometry.

Determination of Energy Contributions of Amino Acids for α_MI-domain Binding and Establishment of the α_MI-binding Algorithm

The algorithm predicting the α_MI-domain binding sequences was constructed based on the statistical energy contributions of individual amino acids within the 9-mer α_MI-domain binding peptides using the strategy described previously for members of the Hsp70 family of molecular chaperones (28). The energy values were calculated according to the equation $\Delta G_K = -RT \ln Pb/Pn$, where R is the gas constant (8.31 J mol^{-1} K^{-1}), T is the absolute temperature in kelvins, and Pb and Pn are the relative frequencies with which each amino acid occurs in binding and non-binding peptides, respectively. Because no specific distance pattern of residues within the α_MI-domain binding peptides was apparent, the relative importance of each residue within the 9-mer was assumed to be equal. The combined energy value for binding and non-binding peptides was calculated using the computer program IRMA (available upon request), which detects potential α_MI-domain binding sites contained within protein primary sequences by searching for segments with the lowest ΔG_K values.

Phage Display

The phage epitope library which displays random seven residue insertions near the N-terminus of the pIII surface protein was purchased from New England Biolabs (Ipswich, MA). The library had

a complexity in excess of two billion independent clones. Affinity panning was performed according to the manufacturer's protocol. Briefly, the wells of 24-well plates were coated with the α_MI-domain (20 μg/ml in PBS) for 3 h at 37 °C and post-coated with 1% PVA in PBS for 1 h at 22 °C. The wells were incubated with the phage library overnight at 4 °C. Following extensive washes with PBS containing 0.1% Tween-20, the α_MI-domain bound phages were eluted with 0.2 M glycine buffer, pH 2.0. The released phages were expanded and after the second panning were eluted with 20 μg/ml P2-C peptide. After third panning, the phages were eluted with P2-C and the sequences of the insert were determined. Control isolations were performed using PVA. 26 selected and 10 unselected phages from the PVA control were sequenced.

Solid-phase Binding Assays

Selected peptides derived from different protein sequences were tested for their ability to compete with binding of the α_MI-domain to immobilized D fragment. 96-well Immulon 4HBX microtiter plates were coated with 1 μg/ml D fragment for 3 h at 37 °C and post-coated with 1% BSA for 1 h. The α_MI-domain as a fusion with GST (10 μg/ml) in TBS containing 1 mM $MgCl_2$ and 0.05% Tween 20 was preincubated with different concentrations of peptides for 1 h at 22 °C and 0.1-ml aliquots were added to the wells. After incubation for 1.5 h at 37 °C, the wells were washed and anti-GST mAb (dilution 1:5000) was added. The bound α_MI-domain was quantified after reaction with goat anti-mouse IgG conjugated with alkaline phosphatase.

Leukocyte Migration Assays

Chemotaxis assays of U937 monocytic cells were performed on 22×22 mm coverslips. Agarose (1%; Life Technologies, Carlsbad, CA) was dissolved in Hank's balanced salt solution and mixed with 2 mg/ml LL-37 to a final concentration of 15 μg/ml. A 10-μl drop of warm agarose solution containing LL-37 was placed at one corner of the cover glass ~1.5 mm from the edge. A control

agarose drop was placed in the diagonally opposite corner of the coverslip and the agarose was allowed to polymerize for 5 min. The coverslips were placed into wells of a 6-well plate containing 5 ml RPMI 1640+10% FBS. A 10-μl aliquot containing 5×10^4 U937 cells was loaded in the center of the cover glass and the plate was incubated for 2 h at 37 °C in a humidified atmosphere containing 5% CO_2. In this experimental format, the cells sediment approximately 5 min after loading to form an ~4-6 mm circle and begin to migrate towards the agarose drop containing LL-37. Photographs of cell migration were taken at 2-mm intervals.

In the second format, migration assays were performed with $\alpha_M\beta_2$-expressing HEK293 cells using Transwell inserts (5-μm pore size) as previously described (26,33). Briefly, the upper chamber of the Transwell system contained 3×10^5 cells and the lower chamber contained LL-37 (0.1-2 μg/ml). For inhibition experiments, cells were pretreated for 15 min with 20 μg/ml of anti-α_M mAb 44a, anti-β_2 mAb IB4, or non-inhibitory mAb OKM1. After incubation for 16 h at 37 °C, cells adherent to the underside of the filter were fixed with paraformaldehyde and stained with Hematoxylin. Selected Transwell assays were performed with thioglycollate-elicited monocyte/macrophages isolated from the peritoneum of wild-type and $\alpha_M\beta_2$-deficient mice (The Jackson Laboratory, Bar Harbor, Maine). Macrophages were purified using the EasySep Mouse selection kit (StemCell Technologies) with mAb against F4/80 conjugated to PE and allowed to migrate for 90 min.

Cell Adhesion Assays

Adhesion assays with $\alpha_M\beta_2$-expressing HEK293 and U937 monocytoid cells were performed essentially as described (24,34). Briefly, the wells of polystyrene microtiter plates (Immulon 4HBX, Dynex Technologies, Chantilly, VA) were coated with 2.5 μg/ml fibrinogen for 3 h at 37 °C and post-coated with 0.5% PVP for 1 h at 37 °C. Cells were labeled with 10 μM Calcein AM (Invitrogen), preincubated for 15 min at 22 °C with selected concentrations of peptides in

DMEM+0.1% BSA, and 100 μl aliquots (4.5×10^4 cells) were added to the wells. Immediately before cell addition, 10 μl 1% BSA was added to the wells. After 30 min incubation at 37 °C, the nonadherent cells were removed and fluorescence was measured in a fluorescence plate reader (CytoFluorII, Applied Biosystems, Foster City, CA).

Flow Cytometry

FACS analyses were performed to assess $\alpha_M\beta_2$ activation and expression on the cell surface of neutrophils induced by LL-37. Neutrophils isolated from human blood under sterile conditions were suspended in HBSS+0.1% BSA at 10^6/0.1 ml and incubated with different concentrations of LL-37 (0.5-10 μg/ml) or fMLP (200 nM) for 5 min at 22 °C. MAb CBRM1/5 (20 μl) conjugated to Alexa 488 was added to cells and incubated for 30 min on ice. The cells were analyzed in a FACS Scan™ (BD Biosciences) as described (5).

RESULTS

Screening Peptide Libraries for α_MI-domain Binding

To determine the recognition specificity of $\alpha_M\beta_2$ we screened cellulose-bound peptide libraries representing the complete sequences of 15 proteins for α_MI-domain binding (Table 1). These proteins were selected because they are known $\alpha_M\beta_2$ ligands (γC and βC domains of fibrinogen, CCN1, N-terminal Pg peptide, ICAM-1, myeloperoxidase, elastase, azurocidin, and soybean trypsin inhibitor) or because they are predicted to be candidate ligands based on knowledge of their amino acid sequences and physico-chemical properties (cathepsin G, proteinase 3, and protein C). Indeed, cathepsin G and protein C support strong adhesion of $\alpha_M\beta_2$-expressing HEK293 cells and monocytoid U937 cells (unpublished data). In addition, bone sialoprotein (BSP) and ovomucoid were selected as predicted non-binding proteins. Figure 1*A* shows representative peptide scans derived from the $\alpha_M\beta_2$-binding and predicted non-binding proteins. The peptide libraries were composed of 9-mer peptides that overlap by six residues. Since 6-mer peptides present minimal α_MI-domain binding signals (34), the scans contain all potential linear binding sites for the α_MI-domain. The use of this screening approach is validated by previous findings that P1 (γ190-202) and P2 (γ377-395), recognition peptides derived from fibrinogen (27) and CCN1-H2 (305-317), the recognition peptide from CCN1 (25), bound the α_MI-domain when covalently coupled to the cellulose membrane (Fig. 1*A*). Furthermore, the specificity of interactions is confirmed by inhibition of the α_MI-domain binding to the γC library in the presence of soluble P2 peptide (27), and the lack of binding of radiolabeled α_LI-domain to the γC and other peptide libraries.

The α_MI-domain bound to only a subset of peptides in each library thus providing an internal control for specificity (Fig. 1*A*). α_MI-domain binding peptides were present in all libraries tested, with no obvious pattern in their distribution within the scans. However, the binding peptides

frequently occurred as clusters, indicating that neighboring peptides with overlapping sequences share α_MI-domain binding sites. Furthermore, the frequency with which the α_MI-domain binding peptides occurred in different libraries varied. Peptide libraries derived from predicted non-ligands contain the smallest number of binders (e.g., BSP in Fig. 1*A*). Based on their ability to bind the α_MI-domain, as assessed by densitometry and visual inspection, the peptides were grouped into three populations: strong binders (196), good binders (462), and non-binding (748).

Distribution of Amino Acids within the α_MI-domain Binding Peptides

We analyzed the relative occurrence of the 20 amino acids in 1,406 peptides representing the entire library. The distribution of amino acids within the library was similar to that found in natural proteins (Fig. 2*A*), except that Ala and Lys were less frequent (~1.4-fold) and Arg was more frequent (1.4-fold). The higher occurrence of Arg probably reflects the fact that established $\alpha_M\beta_2$ ligands such as elastase, myeloperoxidase, and azurocidin and the predicted ligand cathepsin G are cationic proteins that are unusually enriched in this amino acid. The distribution of amino acids in the α_MI-domain binding and non-binding peptides deviated substantially from that in the total library. As shown in Figure 2*B*, peptides that bind the α_MI-domain strongly were enriched in basic residues Arg and Lys (~2- to 2.5-fold) and somewhat in the large hydrophobic residues Leu, Ile, Phe, Val, and Met (~1.1- to 1.5-fold). While the increase in each hydrophobic residue was small, the combined enrichment was ~2-fold (Fig. S1*A*). Negatively charged residues were strongly disfavored and polar residues His, Asn, Gln and Cys were depleted (~2.7-fold) (Fig. 2*B* and Fig. S*1A*). A similar trend existed in the population of good binders; basic residues were over-represented up to ~1.4-fold whereas hydrophobic residues were enriched to the same degree as in strong binders (not shown). In contrast, non-binding peptides were enriched in negatively charged and polar amino acids (Fig. 2*C* and S1*B*) whereas basic residues were strongly disfavored and hydrophobic residues were slightly disfavored. Furthermore, Tyr was enriched in strong (Fig. 2*B*),

but not in non-binding peptides (Fig. 2*C*). In agreement with these findings and our previous data (27), the basicity of peptides was consistent with their ability to bind the α_MI-domain; the positively charged peptides exhibited the highest affinity for the α_MI-domain, whereas neutral peptides had lower affinity and negatively charged peptides were not active. Furthermore, the role of hydrophobic residues is illustrated by the finding that only simultaneous mutation of basic and hydrophobic residues in peptides constituting the γC-library ablated the α_MI-domain binding (27).

Identification of the α_MI-domain recognition motif (IRM)

Apart from being enriched in basic and hydrophobic residues, the α_MI-domain binders displayed significant variability in amino acid sequences with no obvious consensus motif. However, we noted that hydrophobic residues often existed in the immediate proximity of basic residues forming small cores composed of 3-5 residues. The nature of these clusters and the amino acids involved in α_MI-domain binding were determined by computational analyses of 196 strong binders. A list of strong binders is shown in Table S1. As shown in Table 2, the relative occurrence of various motifs in which hydrophobic residues surround basic residues was ~3.5- to 7.0-fold higher in the population of α_MI-domain binding peptides than in non-binding peptides. It is noteworthy that the α_MI-domain binders contain uncommon combinations of amino acids in which basic residues are surrounded by one or two hydrophobic residues on both sides. For example, the occurrence of a rare motif HyHy**B**HyHy in which four hydrophobic residues flank basic residues on both sides (motif 10) was 6.2-fold higher in binders than in non-binding peptides. Likewise, motif 9 (Hy**B**HyHy) occurred 7 times more often in the population of strong binders than in non-binding peptides. Even simple combinations such as BHy or HyB were ~5 times more frequent in the population of α_MI-domain binders. Furthermore, strong binders often contained several short motifs. Analyses indicated that ~80% of all hydrophobic residues in the population were assembled into individual cores that were no further than two residues away from

basic residues. Hydrophobic residues were found more frequently at positions -1 and +1 than at -2 and +2 (Fig. 3). For example, Ile and Met occurred 3.2- and 5.8-fold more often, respectively, in the immediate neighborhood of basic residues than at the secondary positions (Fig. 3). Specific positioning of hydrophobic residues within the cores was not detected with the exception of Met, which was abundant at the -1 position (Fig. 3). The positioning of the small basic/hydrophobic cores within 9-mer peptides does not appear to be important. Although the population of strong binders was enriched in Tyr (1.2-fold), which occurred slightly more frequently at the -3 position, no other preferences for this residue were noted. These analyses suggest that the principal feature of the α_MI-domain recognition motif (IRM) is a short sequence composed of a central basic residue surrounded by hydrophobic residues.

Identification of the α_MI-domain binding motif by phage display

We also determined the α_MI-domain binding sequences by an independent approach. Affinity panning of libraries of bacteriophages that display random peptide sequences at the N-terminus of protein pIII was used to characterize peptides that bind the α_MI-domain. Because peptides containing at least six residues are sufficient for efficient binding to the α_MI-domain, we chose to pan the phage library that displays 7-mer peptides. The library was incubated with the immobilized α_MI-domain and the bound phage were isolated using elution with the P2-C peptide. This approach is based on the finding that P2-C inhibits α_MI-domain binding not only to fibrinogen (from which this peptide is derived) but to many other ligands. After the three rounds of panning, the sequences of the insert of 26 clones were determined. Sequencing data demonstrated that 6 clones contained the ARLPIWF sequence, 2 clones contained ARLPLLW, 2 contained SMKPLWT, and 1 contained GRLPMPW, all of which resemble the most abundant ARLPIWF motif. Although there were no other apparent consensus sequences, 7 clones with affinity for the α_MI-domain contained the sequences in which Arg and Lys were flanked by

hydrophobic residues (LTMPMIR, MAPHVRS, AATKLAF, LEPFFHR, SFRXVSP, AKQPFFW, APHQXRS). The remaining clones contained no basic residues but were enriched in hydrophobic amino acids, especially Leu and Phe (data not shown). Only one Glu and one Asp were detected among 26 sequences. The sequences of inserts derived from phages eluted from PVA were enriched in Ser, Thr, and Gln and had no similarity. PVA is typically used as a post-coat in solid-phase binding assays with the α_MI-domain and adhesion assays with $\alpha_M\beta_2$-expressing cells. These analyses indicate that the sequences revealed by phage display conform to the IRM pattern determined in the analyses of strong binders from peptide libraries.

Development of Algorithm Predicting α_MI-domain Binding Sites

To further analyze the substrate binding preferences of $\alpha_M\beta_2$ we developed the α_MI-domain recognition motif algorithm (IRMA), which allows the prediction of α_MI-domain binding sites in naturally occurring protein sequences and synthetic peptides. The IRMA is based on scoring the statistical energy contributions of each amino acid in the nine residues constituting the proposed α_MI-domain binding motif. The segment composed of nine residues was chosen arbitrarily but its size is justified by findings that continuous stretches of that length are enriched in small cores composed of basic and hydrophobic residues and that their composition deviates significantly from the total library (Table 2). Although 6-mer peptides have the potential to bind the α_MI-domain, this ability is realized only if they do not contain acidic residues. In contrast, 9-mer peptides can bind the α_MI-domain even if they contain acidic residues although as a rule the presence of each acidic residue needs to be compensated by an additional basic residue. The energy values were derived from the fold difference in the overall occurrence of each residue within strong binders and non-binding peptides (Table S2). We then wrote a computer program to calculate the combined energy value for any sequence present in the nanopeptides. This value serves as a measure of probability that the α_MI-domain binds this sequence: the lower the energy

value, the higher the likelihood that the sequence binds the α_MI-domain. As shown in Figure 4, the energy values obtained for most of the population of strong binders ranged from –18 to +2, and those for the general population of non-binding peptides ranged from –2 to +20. A small population of peptides (~ 1%) displayed energy values that were either below or above these limits. The energy distribution within a pool of good binders (range –8 to +8) fell between these two groups (Fig. S2). Based on these data, correct prediction is difficult for peptides with energy values between –2 and +8 but outside of this "gray area" the ability to predict that a peptide binds (or does not bind) the α_MI-domain is high.

To test the quality of the IRMA, we determined the energy distribution within peptide libraries covering the sequences of $\alpha_M\beta_2$ ligands and compared them with experimental data. Within protein sequences, the algorithm scans for the energy minima that represent the regions with the highest probability of α_MI-domain binding. As shown in Figure 5 for the γC-domain of fibrinogen, there was an excellent correlation between the energy minima and the experimentally found α_MI-domain binding sites (27). It is noteworthy that several overlapping peptides encompassing the P2 sequence (spots 76-82) had the lowest energy values. Likewise, the strongest cluster in the scan of CCN1 (spots 91-96; Fig. 1*A*) contains the peptide ^{305}SSVKKYRPKYCGS^{317}, a previously identified binding site for $\alpha_M\beta_2$ (25), and had the lowest energy value. Similar relationships were obtained for all ligands tested, suggesting that a length of nine residues is optimal to predict the majority of sites.

Localization of the α_MI-domain Binding Sequences within Native Protein Structures

To determine the localization of identified α_MI-domain binding sequences within native folded protein structures, we mapped the sequences onto the corresponding three-dimensional structures of neutrophil cationic proteins (see Table 1 for the PDB codes). Only continuous stretches composed of several strong binders were analyzed. α_MI-domain binding sequences were found in

segments representing different secondary structure elements, with slightly higher occurrence in the exposed loops (52% of segments analyzed) than in α-helices (20%) and β-strands (28%). These analyses suggest that the α_MI-domain does not have strong preferences for specific structural motifs. Although some of the α_MI-domain recognition segments were partially or completely buried within the protein cores, many side chains of critical basic and hydrophobic residues constituting the binding sites were well exposed (shown for cathepsin G in Fig. S4). It should be noted that the high percentage of α_MI-domain binding sequences found in exposed loops might reflect the fact that the proteins included in the analyses are cationic proteins.

Peptides Derived from Peptide Libraries and Phage Display Inhibit α_MI-domain Binding

We next examined whether the α_MI-domain binding peptides identified in the scans of various proteins and by phage display block α_MI domain binding. We also examined the relationship between the energy values of peptides predicted by the IRMA and their inhibitory activities. Selected 9-mer peptides corresponding to sequences derived from several proteins and covering a range of energy values were synthesized by traditional Fmoc chemistry and tested in solid phase binding assays (Table 3). In addition, two 7-mer peptides revealed by phage display ARLPIWF and GRLPMPW were synthesized. Many of the selected peptides inhibited α_MI domain binding in a dose-dependent manner. As shown in Table 3, their inhibitory activities varied widely with the highest potency (i.e., lowest IC_{50} values) observed for peptides with the lowest calculated energy values. The peptide RKLRSLWRR (MP-9) derived from myeloperoxidase (-21.3 kJ/mole) was the most potent and, on a molar basis, ~12-fold more active than P2-C (24,27). A strong relationship between the peptides' activities and their calculated energy scores was observed for the group of peptides with energy values in the range of ~ -20 to ~ -7 kJ/mole (Fig. S3). As expected, the peptides with lower energy scores were weak inhibitors (Table 3, Fig. S3). No significant correlation between the peptides' activities and their energy scores was noted (Fig. S3).

The difference between the two groups of peptides showing a biphasic character of the relationship between inhibitory potencies and energy scores remains to be determined. The peptides that displayed positive energy values were inactive.

To corroborate studies with isolated α_MI-domains using whole receptors, the myeloperoxidase-derived peptide MP-9 was tested for its ability to block adhesion of $\alpha_M\beta_2$-expressing cells. The peptide inhibited adhesion in a dose-dependent manner with an IC_{50} of ~10 μM and 3.5 μM for $\alpha_M\beta_2$-expressing HEK293 and U937 monocytic cells, respectively. On a molar basis, MP-9 was a ~4-fold more potent inhibitor of U937 cell adhesion than P2-C. MP-9 also directly supported strong cell adhesion (data not shown).

Binding of the α_MI-domain to Cationic Antimicrobial Peptides

The cationic nature of host defense peptides and the abundance of hydrophobic residues in their highly variable sequences suggest that they may contain α_MI-domain binding sites. To investigate this possibility, we applied the developed algorithm to search The Antimicrobial Peptide Database (http://aps.unmc.edu/AP/main.html). The search revealed that numerous antimicrobial peptides, especially those with a net positive charge >5, contain multiple IRMs. To confirm that these molecules interact with the α_MI-domain, we synthesized peptide libraries covering the sequences of selected mammalian and non-mammalian host defense peptides, including human cathelicidin peptide LL-37, HNP-1, hBD-1, bovine bactenecin 5, fruit fly drosocin, pig tripticin, horseshoe crab polyphemusin, and the synthetic derivative IDR-1 (innate defense regulator) (Table 1). Fig. 1*B* provides examples of selected peptides and shows that they all interacted with the α_MI-domain at multiple sites. Notably, many peptides, such as LL-37 and bactenecin 5, contained long uninterrupted stretches of α_MI-domain binding sequences.

The Human Cathelicidin LL-37 Peptide Induces $\alpha_M\beta_2$-mediated Responses in Monocytes

We next examined the ability of cationic host defense peptides to bind the receptor using the human cathelicidin peptide LL-37 (Fig. 6*A*), an important host defense peptide that is produced by phagocytic and epithelial cells (35). The protective effect of LL-37 seen *in vivo* (36) has been ascribed to its immunomodulatory properties, including the ability to induce a chemotactic response, release of cytokines, gene expression, and activation of intracellular signaling in human monocytes (Reviewed in (37,38)). We analyzed whether any of these responses are mediated by $\alpha_M\beta_2$. As shown in Figure 6*B*, LL-37 induced a chemotactic response in U937 monocytic cells that was inhibited by the function-blocking anti-α_M mAb 44a. LL-37 also induced migration of $\alpha_M\beta_2$-expressing HEK293 cells in a Transwell system (Fig. 6*C*). This response was dependent on LL-37 concentration, occurred at low (0.1-2 μg/ml) LL-37 concentrations (Fig. 6*D*), and was inhibited by function-blocking anti-α_M mAb 44a and anti-β_2 mAb IB4 but not by the non-blocking anti-α_M mAb OKM1 (Fig. 6*C*). LL-37 did not induce migration of macrophages isolated from $\alpha_M\beta_2$-deficient mice suggesting that the loss of $\alpha_M\beta_2$ is specific for the function of LL-37 (Fig. 6*E*). As a control, $\alpha_M\beta_2$-deficient macrophages migrated to fMLP. The ability of LL-37 to induce $\alpha_M\beta_2$-mediated cell migration is reminiscent of that of the P2-C peptide, which duplicates the binding site for $\alpha_M\beta_2$ in fibrinogen (26). It is noteworthy though that LL-37 is the first naturally occurring bioactive peptide shown to bind $\alpha_M\beta_2$.

To investigate the possibility that LL-37 is capable of inducing an activated state of $\alpha_M\beta_2$, we assessed the reactivity of LL-37–treated neutrophils with mAb CBRM1/5, which specifically recognizes an activation-dependent epitope in the α_MI-domain (39). As shown in Fig. 6*F*, LL-37 induced expression of CBRM1/5 epitope in a concentration-dependent manner. The number of cells expressing the activation epitope induced by 10 μg/ml LL-37 was slightly higher than that stimulated by fMLP. When neutrophils were first incubated with the neutrophil inhibitory factor (NIF), a specific inhibitor that blocks ligand binding to the α_MI-domain and then treated with LL-

37, no binding of CBRM1/5 was detected (Fig. 6*F*). These data indicate that binding of LL-37 to $\alpha_M\beta_2$ induces receptor activation and initiates leukocyte migration.

DISCUSSION

By analyzing a large number of sequences present in peptide and phage libraries, we identified the binding motif(s) for the α_MI-domain responsible for the broad ligand recognition exhibited by integrin $\alpha_M\beta_2$. This motif is defined by the following general features: First, the α_MI-domain recognizes short (6-9-mer) sequences enriched in basic and hydrophobic residues that contain small cores consisting of a central basic residue flanked by hydrophobic residues. The motifs recognized by the α_MI-domain are best described as Hy**B**Hy, HyHy**B**Hy, Hy**B**HyHy, and HyHy**B**HyHy patterns in which Hy is any hydrophobic residue and B is Arg or Lys. These motifs are enriched 5- to 7-fold in the population of the α_MI-domain binding compared to non-binding peptides. Motifs in which one of the hydrophobic residues is substituted for any other residue except Asp and Glu are also good α_MI-domain binders. Motifs such as **B**XHy, HyX**B**, HyHy**B**X, and X**B**HyHy (where X is any residue) are enriched 4- to -7-fold in the population of the α_MI-domain binders. Second, acidic residues are strongly disfavored. Third, hydrophobic residues occur more frequently in the immediate neighborhood of basic residues, i.e., at -1 and +1 positions, than at -2 and +2 positions. Fourth, the presence of several basic residues in the 9-mer strongly increases the likelihood of α_MI-domain binding. Fifth, the absence (or strong depletion) of acidic and hydrophilic residues in the regions flanking the 9-mer recognition sequence increases its probability of serving as an α_MI-domain binding site. Thus, the α_MI-domain recognition motif is a relatively small segment and has a high degree of redundancy in the hydrophobic residues that surround critical basic residues. These characteristics are consistent with the capacity of $\alpha_M\beta_2$ to recognize a wide variety of unrelated sequences and thus form the basis for $\alpha_M\beta_2$ ligand binding promiscuity.

The finding that the presence of basic and hydrophobic amino acids is an important feature of the α_MI-domain binding peptides supports and extends our earlier studies with fibrinogen (24,27)

and CCN1 (25) as well as the localization of $\alpha_M\beta_2$ recognition sequences reported in other studies. For example, the identification of the α_MI-domain recognition motif explains the ability of $\alpha_M\beta_2$ to bind LYQAKRFKV or SSVKKYRPKYCGS, the peptides derived from Factor X (40). Although the $\alpha_M\beta_2$-binding nanopeptide CPCFLLGCC (LLG-C4) derived from phage display (41) does not contain basic residues, the presence of Phe-Leu-Leu hydrophobic core may account for its capacity to interact with $\alpha_M\beta_2$. It is interesting that several negatively charged peptides tested in our scans (1.8%) bound the α_MI-domain and those were anomalously enriched in Leu (2.6-fold) and Trp (3.7-fold). However, LLG-C4 was also shown to interact with the monospecific integrin $\alpha_L\beta_2$ whereas typical $\alpha_M\beta_2$ ligands, such as P2-C (γTMKIIPFFNRLTIG), do not (5). Therefore, the presence of basic residues appears to endow the peptide sequences with recognition activity toward $\alpha_M\beta_2$. Although Ile, Leu, Phe, Val and Met are enriched in the α_MI-domain binding peptides, no apparent dominant hydrophobic residues were identified and no specific positioning of hydrophobic residues (except for methionine at -1 position) has been noted (Fig. 2B and Fig. 3). However, we cannot exclude the possibility that some hydrophobic residues are more important than others and further studies of their contribution, as well as that of aromatic residues, are needed to dissect further structural features of the α_MI-domain recognition motif.

The nature of the α_MI-domain recognition motif (IRM) explains why the capacity of $\alpha_M\beta_2$ to bind proteins increases dramatically after their chemical or thermal denaturation (4) or after their unfolding as a consequence of adsorption to plastic surfaces (4,42): the IRMs are rich in hydrophobic residues whose side chains are normally buried in the interior of folded proteins and protein unfolding results in the exposure of such sequences. This behavior is exemplified by P2-C, which is part of the β-sheet in the γC domain of fibrinogen. This sequence is hidden in the soluble intact protein, but becomes unmasked as a result of its unfolding upon adsorption onto various surfaces or during deposition in the extracellular matrix (42,43). The masking of potential α_MI-

domain binding sites in the interior of folded proteins could, at least in part, explain the observation that many intact soluble ligands bind poorly to $\alpha_M\beta_2$-expressing cells in suspension but support efficient adhesion when presented to the receptor in an immobilized form.

Recognition specificity of the α_MI-domain is remarkably reminiscent of that of molecular chaperones, especially those belonging to the Hsp70 family (28,44). This finding supports the original proposal by Davies (4) that the $\alpha_M\beta_2$'s ability to bind unfolded proteins shares similarities with chaperones. Molecular chaperones represent one of several biological systems in which promiscuity in ligand binding plays important roles. Chaperones assist in protein folding in the cell by transient association with a large variety of short hydrophobic sequences that are generally accessible only in non-native conformers. Although differences exist with respect to the positioning of basic residues within the hydrophobic patch, the general requirements regarding size, the presence of basic and hydrophobic residues, and the exclusion of negatively charged residues appear to be comparable for the α_MI-domain and chaperone sequences. It is noteworthy that N<u>RLLL</u>TG, a classic chaperone peptide (45,46) binds the α_MI-domain (Table 3) and inhibits $\alpha_M\beta_2$-mediated cell adhesion. The basis for this activity is likely to be the presence of Arg and adjacent leucines at +1 and +2 positions. Notably, NRLLLTG closely resembles γ^{390}NRLTIG^{395}, the α_MI-domain recognition motif within P2-C (24,34) which has been confirmed as the $\alpha_M\beta_2$-binding site in fibrinogen by genetic manipulation in mice (47,48). The overlap in recognition specificity between $\alpha_M\beta_2$ and chaperons is not entirely unexpected since both proteins prefer unfolded conformers as their substrates. However, they are involved in opposite processes: while chaperones recognize peptide sequences to ensure their correct folding, the α_MI domain only binds to sequences that are exposed in denatured unfolded conformers.

The characteristics of IRM also explain why many neutrophil cationic proteins, including elastase, myeloperoxidase, and azurocidin, are $\alpha_M\beta_2$ ligands. As illustrated here for cathepsin G

(Fig. S4), the side chains of many critical basic and adjacent hydrophobic residues within the α_MI-domain binding clusters are exposed on the surface of these proteins, which probably explains why these proteins bind $\alpha_M\beta_2$ even when presented to the receptor in soluble form. Indeed, free myeloperoxidase released from stimulated neutrophils during inflammation binds neutrophils through an $\alpha_M\beta_2$-dependent mechanism (49). We propose that other cationic proteins that are normally sequestered in granules of resting neutrophils can also bind $\alpha_M\beta_2$ after their release upon cell activation.

Our analyses of the α_MI-domain binding preferences allowed us to develop an algorithm that predicts the α_MI-domain binding sites in $\alpha_M\beta_2$ ligands with high accuracy. This information, in conjunction with the crystal structure when available, may be useful for the prediction of sequences that are displayed on the surface of proteins and potentially serve as α_MI-domain binding sites. The algorithm also appears to be reliable in predicting the potency of α_MI-domain binding peptides. Its application has already uncovered several peptides with affinities for the α_MI domain several fold higher than that previously reported for the fibrinogen peptide P2-C (Table 3). This could be important in the search for antagonists because *in vivo* modulation of $\alpha_M\beta_2$ can be effective in limiting inflammatory injury (reviewed in (50)). Finally, application of the algorithm to search the Antimicrobial Peptide Database (51) revealed that numerous mammalian and non-mammalian cationic peptides contain α_MI-domain recognition patterns and can potentially bind $\alpha_M\beta_2$ (Fig. 1*B*). The prediction that one of the host defense peptides, human cathelicidin LL-37, binds $\alpha_M\beta_2$ was confirmed experimentally.

Previous studies have demonstrated that LL-37 triggers migration of neutrophils and monocytes, induces activation of MAP kinases, production of chemokines, gene expression, and degranulation of mast cells (reviewed in (37,38). The finding that LL-37 contains multiple α_MI-domain binding sites provides new insights into the mechanisms by which LL-37 may elicit

numerous immunomodulatory responses. The mechanism by which LL-37 exerts leukocyte-modulating effects has been controversial. Although the direct chemotactic activity of LL-37 was attributed to G-protein-coupled fMLP-like receptor 1 (52), many other responses induced by this peptide in monocytes are independent of GPCRs (38). The finding that migration of U937 monocytic cells in response to LL-37 is blocked by $\alpha_M\beta_2$ reagents (Fig. 6) indicates that $\alpha_M\beta_2$ is the LL-37 receptor that triggers a migratory signal in these cells.

The $\alpha_M\beta_2$ binding specificity revealed in this study may have broad biological implications and provides a basis for new investigations into the biology of this integrin. First, because of its central role in neutrophil and macrophage biology and its significance as a validated therapeutic target for inflammatory diseases, $\alpha_M\beta_2$ is the subject of intensive research. As a result, the list of $\alpha_M\beta_2$ ligands grows every year and may include many biologically irrelevant molecules. The nature of the α_MI-domain recognition motif suggests that the extensive collection of $\alpha_M\beta_2$ ligands might simply reflect the receptor's potential to bind sequences exposed by protein denaturation. Immobilization of proteins on plastic surfaces, which represents a standard method for testing a protein's capacity to serve as a potential integrin's ligand, inevitably leads to protein unfolding and unmasking of the α_MI-domain binding segments that are normally buried inside the protein's three-dimensional structure. Our findings suggest that some of the ligands that have been identified based on their ability to support $\alpha_M\beta_2$-mediated adhesion may need to be re-evaluated in terms of their physiological relevance.

Second, the identification of the α_MI-domain recognition motif may help to identify new molecules that repel $\alpha_M\beta_2$ and thus render surfaces anti-adhesive for phagocytic leukocytes, an important biomaterial application. Third, since many integrins exhibit promiscuity in ligand binding it will be interesting to determine whether the principles governing $\alpha_M\beta_2$ ligand promiscuity are shared by other members of the integrin family. Fourth, the connection between

the α_MI-domain and chaperones is intriguing. Although the similarities in recognition specificity displayed by both molecules endow them with the ability to recognize diverse ligands, how these recognition principles evolved is unknown.

Finally, the nature of the α_MI-domain recognition motif suggests that $\alpha_M\beta_2$ ligands may serve as alarm/danger signals. It has been proposed that proteins released by damaged or dead cells alarm the immune system (53,54). The original "danger" model postulated that segments of proteins that initially are buried inside the folded molecules, especially their hydrophobic portions, would function as alarm signals when exposed (53). Consequently, if a cell is disrupted the hydrophobic sequences of nascent proteins synthesized on ribosomes, which are normally bound to chaperones, will be exposed. The characteristics of the α_MI-domain recognition sequences with their abundance of hydrophobic and positively charged residues, their resemblance to the segments recognized by chaperones, and an enormous diversity of α_MI-domain binding sequences is consistent with the idea that $\alpha_M\beta_2$ is an alarm/danger-sensing molecule, or the so-called "alarmin" receptor. The term alarmin was initially coined to include activities of cathelicidin LL-37, defensins, HMGB 1, and eosinophil-derived neurotoxin (EDN) which, despite their diverse structure, all have chemotactic and activating effects on leukocytes (55). Moreover, the three of these are $\alpha_M\beta_2$ ligands (Fig. 1 and (13)). As an extension of the "alarm/danger" model, we propose that neutrophil cationic proteins/peptides that are sequestered in granules of resting neutrophils and secreted during the immune-inflammatory response would also qualify as alarmins. Indeed, myeloperoxidase activates neutrophils via $\alpha_M\beta_2$-dependent MAPK activation (49), and azurocidin induces Ca^{2+} mobilization in monocytes via β_2 integrins, most likely $\alpha_M\beta_2$ (56). Thus, it is reasonable to assume that any intracellular molecule that carries the IRMs and is released from damaged cells during tissue injury might be an alarmin that signals through the $\alpha_M\beta_2$ receptor. The recently reported HMGN1 (high-mobility nucleosome-binding protein 1) is an

alarmin (57) and also a potential $\alpha_M\beta_2$ ligand (Fig. S5). Although several receptors have been implicated in triggering alarmin responses in leukocytes (55), $\alpha_M\beta_2$ is the first molecule for which a common recognition pattern present in a large and diverse group of alarmin molecules has been identified.

In summary, we have revealed the molecular basis for the broad ligand specificity exhibited by integrin $\alpha_M\beta_2$, solving the long-standing puzzle of why ligand binding by this receptor is driven by protein denaturation and leading to several conjectures. The prediction that the host defense peptide LL-37, which harbors several IRMs, triggers immunomodulatory responses via $\alpha_M\beta_2$ has been confirmed experimentally. Another proposal, based partly on experimental evidence (Fig. 1*B*), predicts that host defense peptides are a new class of $\alpha_M\beta_2$ ligands. The newly solved α_MI-domain recognition motif could be used to identify molecules that are induced in injured tissues and might act as alarmins through activation of $\alpha_M\beta_2$.

SUPPORTING INFORMATION

Five supplementary figures and two supplementary tables are available. This material is available free of charge via the Internet at http://pubs.acs.org.

ACKNOWLEDGEMENTS

This work was supported by NIH grant HL 63199. We thank Dr. V. Yakubenko and Peter Ryabokon for technical assistance with screening phage libraries, and Dr. Xu Wang, Department of Chemistry and Biochemistry, Arizona State University for help with analyses of protein structures.

REFERENCES

1. Coxon, A., Rieu, P., Barkalow, F. J., Askari, S., Sharpe, A. H., Von Andrian, U. H., Arnaout, M. A., and Mayadas, T. N. (1996) A novel role for the β_2 integrin CD11b/CD18 in neutrophil apoptosis: a homeostatic mechanism in inflammation, *Immunity* 5, 653-666

2. Ding, Z. M., Babensee, J. E., Simon, S. I., Lu, H. F., Perrard, J. L., Bullard, D. C., Dai, X. Y., Bromley, S. K., Dustin, M. L., Entman, M. L., Smith, C. W., and Ballantyne, C. M. (1999) Relative contribution of LFA-1 and Mac-1 to neutrophil adhesion and migration, *J.Immunol.* 163, 5029-5038

3. Prince, J. E., Brayton, C. F., Fosset, M. C., Durand, J. A., Kaplan, S. L., Smith, C. W., and Ballantyne, C. M. (2001) The differential roles of LFA-1 and Mac-1 in host defense against systemic infection with Streptococcus pneumoniae, *J.Immunol.* 166, 7362-7369

4. Davis, G. E. (1992) The Mac-1 and p150,95 β_2 integrins bind denatured proteins to mediate leukocyte cell-substrate adhesion, *Exp.Cell Res.* 200, 242-252

5. Yakubenko, V. P., Lishko, V. K., Lam, S. C. T., and Ugarova, T. P. (2002) A molecular basis for integrin $\alpha_M\beta_2$ ligand binding promiscuity, *J.Biol.Chem.* 277, 48635-48642

6. Diamond, M. S., Staunton, D. E., de Fougerolles, A. R., Stacker, S. A., Garcia-Aguilar, J., Hibbs, M. L., and Springer, T. A. (1990) ICAM-1 (CD54): a counter-receptor for Mac-1 (CD11b/CD18), *J.Cell Biol.* 111, 3129-3139

7. Simon, D. I., Chen, Z. P., Xu, H., Li, C. Q., Dong, J. F., McIntire, L. V., Ballantyne, C. M., Zhang, L., Furman, M. I., Berndt, M. C., and López, J. A. (2000) Platelet glycoprotein Ibα is a counterreceptor for the leukocyte integrin Mac-1 (CD11b/CD18), *J.Exp.Med.* 192, 193-204

8. Santoso, S., Sachs, U. J., Kroll, H., Linder, M., Ruf, A., Preissner, K. T., and Chavakis, T. (2002) The junctional adhesion molecule 3 (JAM-3) on human platelets is a counterreceptor for the leukocyte integrin Mac-1, *J.Exp.Med* 196, 679-691

9. Cai, T. Q. and Wright, S. D. (1996) Human leukocyte elastase is an endogenous ligand for the integrin CR3 (CD11b/CD18, Mac-1, $\alpha_M\beta_2$) and modulates polymorphonuclear leukocyte adhesion, *J.Exp.Med.* 184, 1213-1223

10. Johansson, M. W., Patarroyo, M., Oberg, F., Siegbahn, a., and Nilsson, K. (1997) Myeloperoxidase mediates cell adhesion via the $\alpha_M\beta_2$ integrin (Mac-1, CD11b/CD18), *J.Cell Sci.* 110, 1133-1139

11. Lishko, V. K., Novokhatny, V., Yakubenko, V. P., Skomorovska-Prokvolit, H., and Ugarova, T. P. (2004) Characterization of plasminogen as an adhesive ligand for integrins $\alpha_M\beta_2$(Mac-1) and $\alpha_5\beta_1$(VLA-5), *Blood* 104, 719-726

12. Zirlik, A., Maier, C., Gerdes, N., MacFarlane, L., Soosairajah, J., Bavendiek, U., Ahrens, I., Ernst, S., Bassler, N., Missiou, A., Patko, Z., Aikawa, M., Schonbeck, U.,

Bode, C., Libby P., and Peter, K. (2007) CD40 Ligand mediates inflammation independently of CD40 by interaction with Mac-1, *Circulation* 115, 1571-1580

13. Orlova, V. V, Choi, E.Y., Xie, C., Chavakis, E., Bierhaus, A., Ihanus, E., Ballantyne, C. M., Gahmberg, C. G., Bianchi, M. E., Nawroth, P. P., and Chavakis, T. (2007) A novel pathway of HMGB1-mediated inflammatory cell recruitment that requires Mac-1 integrin, *EMBO J.* 26, 1129-1139

14. Wang, H, Zhang, M., Andersson, U., Vishnubhakat, J. M., Ombrellino, M, Che, J., Frazier, A., Yang, H, Ivanova, S., Borovikova, L., Monogue, K. R., Faist, E., Abraham, E., Andersson, J, Molina, P. E., Abumrad, N. N., Sama, A., and Tracey, K. J. (1999) HMG-1 as a late mediator of endotoxin lethality in mice, *Science* 285, 248-251

15. Scaffidi, P., Misteli, T., and Bianchi, M. E. (2002) Release of chromatin protein HMGB1 by necrotic cells triggers inflammation, *Nature* 418, 191-195

16. Lotze, M. T. and Tracey, K. J. (2005) High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal, *Nat.Rev.Immunol.* 5, 331-342

17. Diamond, M. S., Garcia-Aguilar, J., Bickford, J. K., Corbí, A. L., and Springer, T. A. (1993) The I domain is a major recognition site on the leukocyte integrin Mac-1 (CD11b/CD18) for four distinct adhesion ligands, *J.Cell Biol.* 120, 1031-1043

18. Yalamanchili, P., Lu, C. F., Oxvig, C., and Springer, T. A. (2000) Folding and function of I domain-deleted Mac-1 and lymphocyte function-associated antigen-1, *J.Biol.Chem.* 275, 21877-21882

19. Zhang, L. and Plow, E. F. (1997) Identification and reconstruction of the binding pocket within $\alpha_M\beta_2$ for a specific and high affinity ligand, NIF, *J.Biol.Chem.* 272, 17558-17564

20. Zhang, L. and Plow, E. F. (1999) Amino acid sequences within the α subunit of integrin $\alpha_M\beta_2$ (Mac-1) critical for specific recognition of C3bi, *Biochemistry* 38, 8064-8071

21. Yakubenko, V. P., Solovjov, D. A., Zhang, L., Yee, V. C., Plow, E. F., and Ugarova, T. P. (2001) Identification of the binding site for fibrinogen recognition peptide γ383-395 within the α_M I-domain of integrin $\alpha_M\beta_2$, *J.Biol.Chem.* 275, 13995-14003

22. Schober, J. M., Chen, N., Grzeszkiewicz, T., Emeson, E. E., Ugarova, T. P., Ye, R. D., Lau, L. F., and Lam, S. C. T. (2002) Identification of integrin $\alpha_M\beta_2$ as an adhesion receptor on peripheral blood monocytes for Cyr61 and connective tissue growth factor, immediate-early gene products expressed in atherosclerotic lesions, *Blood* 99, 4457-4465

23. Altieri, D. C., Plescia, J., and Plow, E. F. (1993) The structural motif glycine 190-valine 202 of the fibrinogen gamma chain interacts with CD11b/CD18 integrin $\alpha_M\beta_2$, Mac-1) and promotes leukocyte adhesion, *J.Biol.Chem.* 268, 1847-1853

24. Ugarova, T. P., Solovjov, D. A., Zhang, L., Loukinov, D. I., Yee, V. C., Medved, L. V., and Plow, E. F. (1998) Identification of a novel recognition sequence for integrin $\alpha_M\beta_2$ within the gamma-chain of fibrinogen, *J.Biol.Chem.* 273, 22519-22527

25. Schober, J. M., Lau, L. F., Ugarova, T. P., and Lam, S. C. (2003) Identification of a novel integrin $\alpha_M\beta_2$ binding site in CCN1 (Cyr61), a matricellular protein expressed in healing wounds and atherosclerotic lesions, *J.Biol.Chem.* 278, 25808-25815

26. Forsyth, C. B., Solovjov, D. A., Ugarova, T. P., and Plow, E. F. (2001) Integrin $\alpha_M\beta_2$-mediated cell migration to fibrinogen and its recognition peptides, *J.Exp.Med.* 193, 1123-1133

27. Lishko, V. K., Podolnikova, N. P., Yakubenko, V. P., Yakovlev, S., Medved, L., Yadav, S. P., and Ugarova, T. P. (2004) Multiple binding sites in fibrinogen for integrin $\alpha_M\beta_2$ (Mac-1), *J.Biol.Chem.* 279, 44897-44906

28. Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries, *EMBO J.* 16, 1501-1507

29. Kramer, A., Keitel, T., Winkler, K., Stocklein, W., Hohne, W., and Schneider-Mergener, J. (1997) Molecualr basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody, *Cell* 91, 799-809

30. Knoblauch, N. T. M., Rudiger, S., Schonfeld, H.-J., Driessen, A. J. M., Schneider-Mergener, J., and Bukau, B. (1999) Substrate specificity of the SecB Chaperone, *J.Biol.Chem.* 274, 34219-34225

31. Ugarova, T. P. and Budzynski, A. Z. (1992) Interaction between complementary polymerization sites in the structural D and E domains of human fibrin, *J.Biol.Chem.* 267, 13687-13693

32. Kramer, A. and Schneider-Mergener, J. (1998) Synthesis and screening of peptide libraries on continuous cellulose membrane supports, *Methods in Molecular Biology* 87, 25-39

33. Lishko V.K., Yakubenko, V. P., and Ugarova, T. P. (2003) The interplay between Integrins $\alpha_M\beta_2$ and $\alpha_5\beta_1$ during cell migration to fibronectin, *Exp.Cell Res.* 283, 116-126

34. Ugarova, T. P., Lishko, V. K., Podolnikova, N. P., Okumura, N., Merkulov, S., Yakubenko, V. P., Yee, V. C., Lord, S. T., and Haas, T. A. (2003) Sequence 377-395 (P2), but not γ190-202(P1), is the binding dite for the α_MI-domain of integrin $\alpha_M\beta_2$ in the γC-domain of fibrinogen, *Biochemistry* 42, 9365-9373

35. Zanetti.M. (2005) The role of cathelicidins in the innate host defenses of mammals, *Antimicrobial peptides in human health and disease*. R.L. Gallo, editor. Horizon Bioschence, Norfolk, U.K. 15-50

36. Scott, M. G. and Hancock, R. E. W. (2000) Cationic antimicrobial peptides and their multifunctional role in the immune system, *Crit.Rev.Immunology* 20, 407-431

37. Finlay, B. B. and Hancock, R. E. W. (2004) Can innate immunity be enhanced to treat microbial infections? *Nature Reviews Microbiology* 2, 497-504

38. Mookherjee, N. and Hancock, R. E. W. (2007) Cationic host defense peptides: Innate immune regulatory peptides as a novel approach for treating infections, *Cell.Mol.Life Sci.* 64, 922-933

39. Oxvig, C., Lu, C., and Springer, T. A (1999) Conformational changes in tertiary structure near the ligand binding site of an integrin I domain, *Proc.Natl.Acad.Sci.USA* 96, 2215-2220

40. Mesri, M., Plescia, J., and Altieri, D. C. (1998) Dual regulation of ligand binding by CD11b I domain. Inhibition of intercellular adhesion and monocyte procoagulant activity by a factor X-derived peptide, *J.Biol.Chem.* 273, 744-748

41. Koivunen, E., Ranta, T. M., Annaila, A., Taube, S., Uppala, A., Jokinen, M., van Willigen, G, and Gahmberg, C. G. (2001) Inhibition of β_2 integrin-mediated leukocyte adhesion by Leucine-Leucine-glycine motif-containing peptides, *J.Cell Biol.* 153, 905-915

42. Lishko, V. K., Kudryk, B., Yakubenko, V. P., Yee, V. C., and Ugarova, T. P. (2002) Regulated unmasking of the cryptic binding site for integrin $\alpha_M\beta_2$ in the γC-domain of fibrinogen, *Biochemistry* 41, 12942-12951

43. Scott, E. A. and Elbert, D. L. (2008) Mass spectrometric mapping of fibrinogen conformations at poly(ethylene terephtalate) interfaces, *Biomaterials* 28, 3904-3917

44. Rudiger, S., Schneider-Mergener, J., and Bukau, B. (2001) Its substrate specificity characterizes the Dna co-chaperone as a scanning factor for the DnaK chaperone, *EMBO J.* 20, 1042-1050

45. Gragerov, A., Zeng, L., Zhao, X., Burkholder, W., and Gottesman, M. E. (1994) Specificity of DnaK-peptide binding, *J.Mol.Biol.* 235, 848-854

46. Rudiger, S., Buchberger, A., and Bukau, B. (1997) Interaction of Hsp70 chaperones with substrates, *Nat.Struct.Biol.* 4, 342-349

47. Flick, M. J., Du, X., Witte, D. P., Jirouskova, M., Soloviev, D. A., Plow, E. F., and Degen, J. L. (2004) Leukocyte engagement of fibrin(ogen) via the integrin receptor $\alpha_M\beta_2$/Mac-1 is critical for host inflammatory response in vivo, *J.Clin.Invest.* 113, 1596-1606

48. Flick, M. J., LaJeunesse, C. M, Talmage, K. E., Witte, D. P., Palumbo, J. S., Pinkerton, M. D., Thornton, S., and Degen, J. L. (2007) Fibrin(ogen) exacerbates inflammatory join disease through a mechanism linked to the integrin $\alpha_M\beta_2$ binding motif, *J.Clin.Invest.* 117, 3224-3235

49. Lau, D., Mollnau, H., Eiserich, J., Freeman, B. A., Daiber, A., Gehling, U. M., Brummer, J., Rudolph, V., Munzel, T., Heitzer, T., Meinertz, T., and Baldus, S. (2005) Myeloperoxidase mediates neutrophil activation by association with CD11b/CD18 integrins, *Proc.Natl.Acad.Sci.USA* 102, 431-436

50. Park, J. Y., Arnaout, M. A., and Gupta, V. (2007) A simple, no wash cell adhesion-based high-throughput assay for the discovery of small-molecule regulators of the integrin CD11b/CD18, *Journal of Biomolecular Screening* 12, 406-417

51. The antimicrobial peptide database. *http://aps.unmc.edu/AP/main.html*

52. Yang, D., Chen, Q., Schmidt, A. P., Anderson, G. M., Wang, J. M., Wooters, J., Oppenheim, J. J., and Chertov, O. (2000) LL-37, the neutrophil granule- and epithelial cell-derived cathelicidin, utilizes formyl peptide receptor-like 1 (FPRL1) as a receptor to chemoattract human peripheral blood neutrophils, monocytes, and T cells, *J.Exp.Med* 192, 1069-1074

53. Matzinger, P. (2002) The danger model: a renewed sense of self, *Science* 296, 301-305

54. Seong, S.-Y. and Matzinger, P. (2004) Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses, *Nature Reviews Immunology* 4, 469-478

55. Oppenheim, J. J. and Yang, D. (2005) Alarmins: chemotactic activators of immune responses, *Current Opinion in Immunology* 17, 359-365

56. Soehnlein, O., Xie, X., Ulbrich, H., Kenne, E., Rotzius, P., Flodgaard, H., Eriksson, E. E., and Lindbom, L. (2005) Neutrophil-derived heparin-binidng protein (HBP/CAP37) deposited on endothelium enhances monocyte arrest under flow conditions, *J.Immunol.* 174, 6399-6405

57. Yang, D., Postnikov, Y. V., Li, Y., Tewary, P., de la Rosa, G., Wei, F., Klinman, D., Giovanni, T., Weiss, J. P., Furusawa, T., Bustin, M., and Oppenheim, J. J. (2012) High-mobility group nucleosome-bindign protein 1 acts as an alarmin and is critical for lipopolysaccharide-induced immune responses, *J.Exp.Med.* 209, 157-171

58. Klapper, M. H. (1977) The independent distribution of amino acid near neighbor pairs into polypeptides, *Biochem.Biophys.Res.Commun.* 78, 1018-1024

59. Coeytaux, K. Poupon A. (2005) Prediction of unfolded segments in a protein sequence based on amino acid composition, *Bioinformatics* 21, 1891-1900

60. Diamond, M. S., Staunton, D. E., Marlin, S. D., and Springer, T. A. (1991) Binding of the integrin Mac-1 (CD11b/CD18) to the third immunoglobulin-like domain of ICAM-1 (CD54) and its regulation by glycosylation, *Cell* 65, 961-971

61. Cai, T.-Q. and Wright, S. D. (1996) Human leukocyte elastase is an endogenous ligand for the integrin CRR3 (CD11b/CD18, Mac-1, $\alpha_M\beta_2$) and modulates polymorphonuclear leukocyte adhesion, *J.Exp.Med.* 184, 1213-1223

62. Scott, M. G., Dullaghan, E., Mookherjee, N., Glavas, N., Waldbrook, M., Thompson, A., Wang, A., Lee, K., Doria, S., Hamill, P., Yu, J. J., Li, Y., Donini, O., Guarna, M. M., Finlay, B. B., North, J. R., and Hancock, R. E. W. (2007) An anti-infective peptide that selectively modulates the innate immune response, *Nat.Biotechnol.* 25, 465-472

FIGURE LEGENDS

FIGURE 1. **Screen of cellulose-bound peptide libraries spanning the sequences of selected proteins and peptides for α_MI-domain binding.** *A,* Peptide libraries derived from the sequences of human cathepsin G, elastase, azurocidin, proteinase 3, human myeloperoxidase, connective tissue growth factor (CCN1), Protein C, soy bean trypsin inhibitor (SBTI), bone syaloprotein (BSP) and ovomucoid were screened for α_MI-domain binding. The numbers indicate the peptide (spot) numbers. *B*, Peptide libraries derived from the sequences of mammalian and non-mammalian antimicrobial peptides: bovine bactenecin 5, human HNP-1, human cathelicidin LL-37, fruit fly drosocin, polyphemusin from American horseshoe crab, pig tritrpticin, and human beta defensin 1 (BD-1). IDR-1 (innate defense regulator 1) is a synthetic peptide (58). The libraries were constructed and the α_MI-domain binding examined as described in Experimental Procedures.

FIGURE 2. **Amino acid distribution in α_MI-domain binding and non-binding peptides.** *A,* Relative occurrence of all amino acids in 1,406 peptides derived from 16 protein sequences (Table I) that represent the entire library (gray bars) was compared with that in protein sequences derived from the protein databases (58,59) (black bars). The frequency of each amino acid in the group of strong α_MI-domain binders (*B*) and non-binding peptides (*C*) was calculated as a percentage of its occurrence in the whole peptide library (100%, dashed line). The data for each amino acid in strong binding and non-binding peptides were obtained for each protein shown in Table 1 and presented as mean ± SD. P values are given for comparison across the entire library. Statistical analyses were performed using Student's t test. *$p < 0.05$, **$p < 0.01$.

FIGURE 3. **Frequency of hydrophobic residues at the specified positions around basic residues in the population of strong α_MI-domain binders**. The relative occurrence of hydrophobic residues at -2, -1, +1, and +2 positions around basic residues (Arg or Lys; set as "0") in the population of strong α_MI-domain binders (196 peptides) is given as a percentage. Although Pro was slightly depleted in the population, it was frequently found in the vicinity of basic residues and therefore, was included in the analyses. The small hydrophobic amino acid Ala is not included.

FIGURE 4. **Distribution of energy values in populations of α_MI-domain binding and non-binding peptides.** Energy distribution in populations of strong (—) and non-binding (---) α_MI-domain peptides was calculated as a percentage of the total peptides in each population.

FIGURE 5. **Energy distribution within the peptide library spanning the γC domain of fibrinogen.** *A,* α_MI-domain binding to the peptide scan derived from the γC domain of fibrinogen (residues γ148-411) was described previously (27). *B*, the 9-mer peptide library of the γC domain shown in A with segments having the lowest energy values shaded in gray. *C*, The energy value for each 9-mer peptide in the library was calculated using the developed algorithm (shown on the ordinate). The position of each peptide in the library is shown on the abscissa. Peptides with negative energy values correspond to α_MI-domain binding peptides (the spots in Figure 5A) whereas those with positive energy values do not bind the α_MI-domain.

FIGURE 6. **LL-37 promotes functional responses in monocytes via integrin $\alpha_M\beta_2$.** *A*, Sequence of LL-37 with the α_MI-domain recognition motifs underlined. *B,* Migration of U937 monocytoid cells along the gradient of LL-37 in the absence (*upper panel*) or presence (*lower panel*) of function-blocking anti-α_M mAb 44a. Cells were added to the spots marked by crosses. Direction of cell migration towards the agarose gel with 1.5 μg/ml of LL-37 is shown above

the figure (arrow) and the gradient of LL-37 is shown at the bottom. The edges of the gel in *c* and *f* are marked by dashed lines. Cell migration was determined after 2 h at 37 °C. In the absence of blocking mAb 44a (a), cells moved towards the LL-37 concentration gradient (*b*) and some cells arrived at the edge of the agarose gel (c). In contrast, no cell migration was detected in the presence of anti-α_M mAb 44a (*d-f*). Photographs were taken using a 20× objective. A schematic representation of the experimental design is shown on the right. *C*, Migration of $\alpha_M\beta_2$-expressing HEK293 cells to LL-37 (0.5 μg/ml) in a Transwell system. As indicated, cells in the upper wells were pretreated with 20 μg/ml anti-α_M mAb 44a, anti-β_2 mAb IB4, or mAb OKM1. Migration data are expressed as mean cells/view ± SE from three or more experiments. *** denotes medium alone. *D,* Dose-dependent migration of $\alpha_M\beta_2$-expressing HEK293 cells towards LL-37. *E*, Migration of thioglycollate-elicited macrophages isolated from the peritoneum of wild-type and $\alpha_M\beta_2$-deficient mice to LL-37 (1 μg/ml) and fMLP (100 nM). *F*, LL-37-induced expression of the activation-dependent epitope in the α_MI-domain. Neutrophils were pretreated with fMLP (200 nM) or different concentrations of LL-37 and then incubated with a conformation-dependent mAb CBRM1/5. Epitope expression was assessed by flow cytometry. Selected cells were pretreated with NIF (1 μg/ml) before addition of LL-37 (4 and 10 μg/ml) followed by mAb CBRM1/5.

TABLE 1

Protein and peptide sequences that were screened for α_MI-domain binding

Proteins	Protein data bank (PDB) accession number	Number of residues	pI	PDB code for the 3D-structure	Shown to be a $\alpha_M\beta_2$ ligand, a predicted ligand, or a predicted non-ligand
Azurocidin precursor (CAP37, cationic antimicrobial protein, HBP)	P20160	222	9.5	1A7S	(8)
Bone syaloprotein (BSP)	P10451	298	4.4	NA[a]	Predicted non-ligand
Cathepsin G	P08311	235	11.4	1AU8	Predicted ligand[b]
Elastase	P08246	238	9.9	1PPG	(8)
Myeloperoxidase	P05164	697	9.3	1DNW	(*10*)
CCN1 (Cyr61)	O00622	357	8.5	NA	(*22*)
Fibrinogen A α-chain (1-611) Fibrinogen β C-domain (200-461) Fibrinogen γ C-domain (148-411)	P02671 P02675 P02679	611 261 263	7.7 7.0 6.1	1FZA 1FZA 1FIB	Predicted ligand[b] (*27*) (23,24,27)
ICAM-1 (IgG-like C2-type domain 3)	P05362	67	4.2	NA	(*60*)
Ovomucoid	P01005	186	4.8	NA	Predicted non-ligand
Pg N-terminal peptide (1-78)	Q5TEH4	78	4.7	NA	(*11*)
Protein C	P04070	419	5.6	1AUT	Predicted ligand[b] (2004)
Proteinase 3	P24158	221	7.8	1FUJ	(*61*)
Soybean Trypsin Inhibitor (SBTI)	P01071	181	4.7	1AVU	(*4*)
Antimicrobial peptides					
Cathelicidin (hCAP-18/LL-37)	P49913	37	10.6	2K6O	(Lishko et al, 2014)[c]
Bactenecin 5	P19660	43	12.5	NA	Predicted ligand[b]
HNP-1	P59665	30	8.7	3GNY	Predicted ligand
HBD-1	P60022	36	8.9	1E4S	Predicted ligand
Drosocin	P36193	19	12.0	NA	Predicted ligand
Tritrpticin	P51524	13	12.5	1D6X	Predicted ligand
Polyphemusin 1	P14215	18	10.3	1RKK	Predicted ligand
IDR-1 (innate defense regulator)	NA	13	11.0	NA	(*62*)

[a] NA, not available

[b] Supports adhesion of $\alpha_M\beta_2$-expressing cells, including $\alpha_M\beta_2$-transfected-HEK293 cells, human neutrophils, human monocytoid U937 cells, and murine IC-21 macrophages.

[c] Manuscript prepared for submission

TABLE 2

Occurrence of motifs composed of basic and hydrophobic residues in α_MI-domain binding and non-binding peptides

	1	2	3	4	5	6	7	8	9	10
Motifs	0 +1 **B Hy**	-1 0 **Hy B**	0 +1 +2 **B X Hy**	-2 -1 0 **Hy X B**	-1 0 +1 **Hy B Hy**	-2 -1 0 +1 **Hy Hy B X**	-1 0 +1 +2 **X B Hy Hy**	-2 -1 0 +1 **Hy Hy B Hy**	-1 0 +1 +2 **Hy B Hy Hy**	-2 -1 0 +1 +2 **Hy Hy B Hy Hy**
Strong binders	83 (4.6)	83 (5.5)	70 (3.5)	74 (5.7)	33 (5.5)	21 (5.3)	24 (4.8)	11 (5.5)	14 (7.0)	3.1 (6.2)
Non-binders	18	15	20	13	6	4	5	2	2	0.5

Numbers shown are occurrence of a particular motif in the population. Numbers in parentheses show the increase (folds) of each motif in the population of strong binders compared to the population of non-binding peptides. **B** denotes basic residues Arg and Lys, **Hy** is any hydrophobic residue, and **X** is any residue (except Asp or Glu).

TABLE 3

Inhibitory potency of the α_MI-binding peptides derived from the peptide and phage display libraries

Peptide	Origin	IC_{50}, μM	∑ E (kJ/mol)[a]
RKLRSLWRR	Myeloperoxidase	2.5 ±0.2	-21.3
QVLRIRKRA	Protein C	8.4 ± 0.5	-15.5
LQLRFPRFV	Cathepsin G	10.5 ± 1.6	-11.0
KYRLTYAFAG	Fibrinogen	12.6 ± 1.5	-9.3
ARLPIWF	Phage display [b]	17 ± 3	-6.8
TMKIIPFNRLTIG (P2-C)	Fibrinogen	30 ± 2.5	-6.7
GRLPMPW	Phage display [b]	36 ±7	-3.8
NRLLLTG	Chaperone ligand[c]	110 ±16	-2.9
SVNKYRGTAGNA	Fibrinogen	93 ±11	-1.9
GWTVFQKRLDGS (P1)	Fibrinogen	72 ±4	-0.12
LLHNYGVYT	Protein C	25% inhibition[d]	0.14
GDDPSDKFF	Fibrinogen	No inhibition	15.1

Different concentrations of each peptide were preincubated with α_MI-domain-GST and the mixture was added to microtiter plates coated with human fibrinogen D fragment (1 μg/ml). The inhibitory potencies of the peptides are presented as molar concentration required for 50% inhibition (IC_{50}).

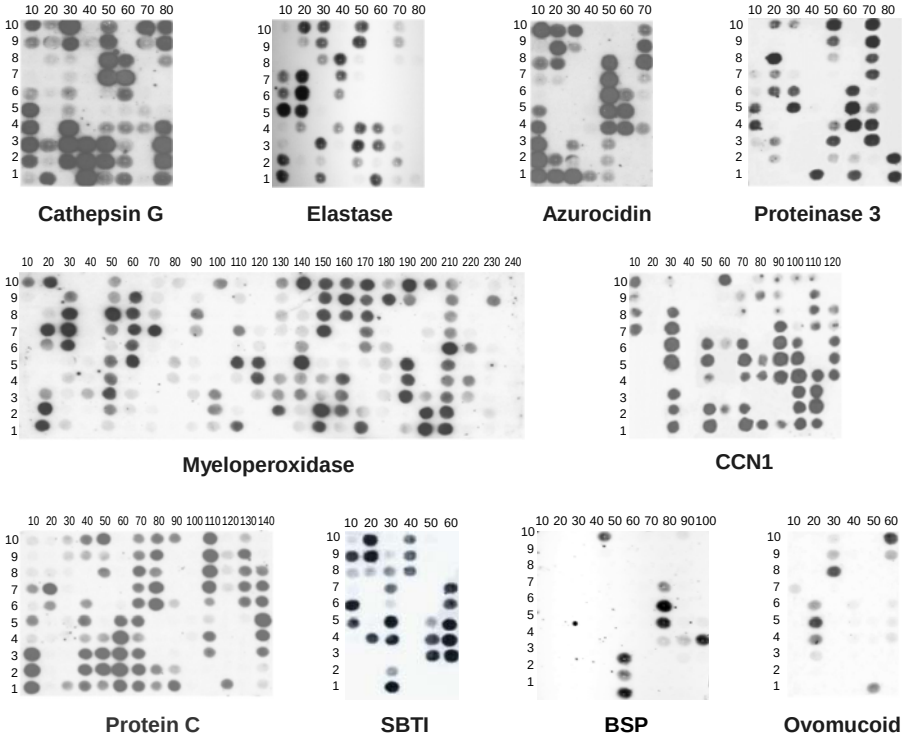
[a] The combined energy value for each sequence was calculated as described in the Experimental Procedures.

[b] Peptide sequences derived from phage display analyses.

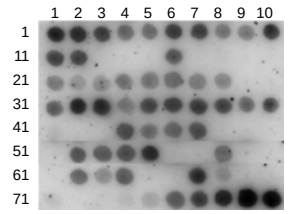
[c] A commonly used chaperone ligand (45).

[d] 25% inhibition at 200 μM.

A



B



Bactenecin 5 (spots 1-13) RFRPPIRRPPIRPPFYPPFRPPIRPPIFPPIRPPFRPPLGPFP
IDR-1 (spots 16-18) KSRIVPAIPVSL
HNP-1 (spots 21-28) ACYCRIPACIAGERRYGTCIYQGRLWAFCC
LL-37 (spots 31-41) LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES
Drosocin (spots 44-47) KPRPYSPRPTSHPRPIR
Polyphemusin 1 (spots 52-55) RRWCFRVCYRGFCYRKCR
Tritrpticin (spots 62-64) VRRFPWWWPFLRR
hBD-1 (spots 71-80) DHYNCVSSGGQCLYSACPIFTKIQGTCYRGKAKCCK

Fig. 1

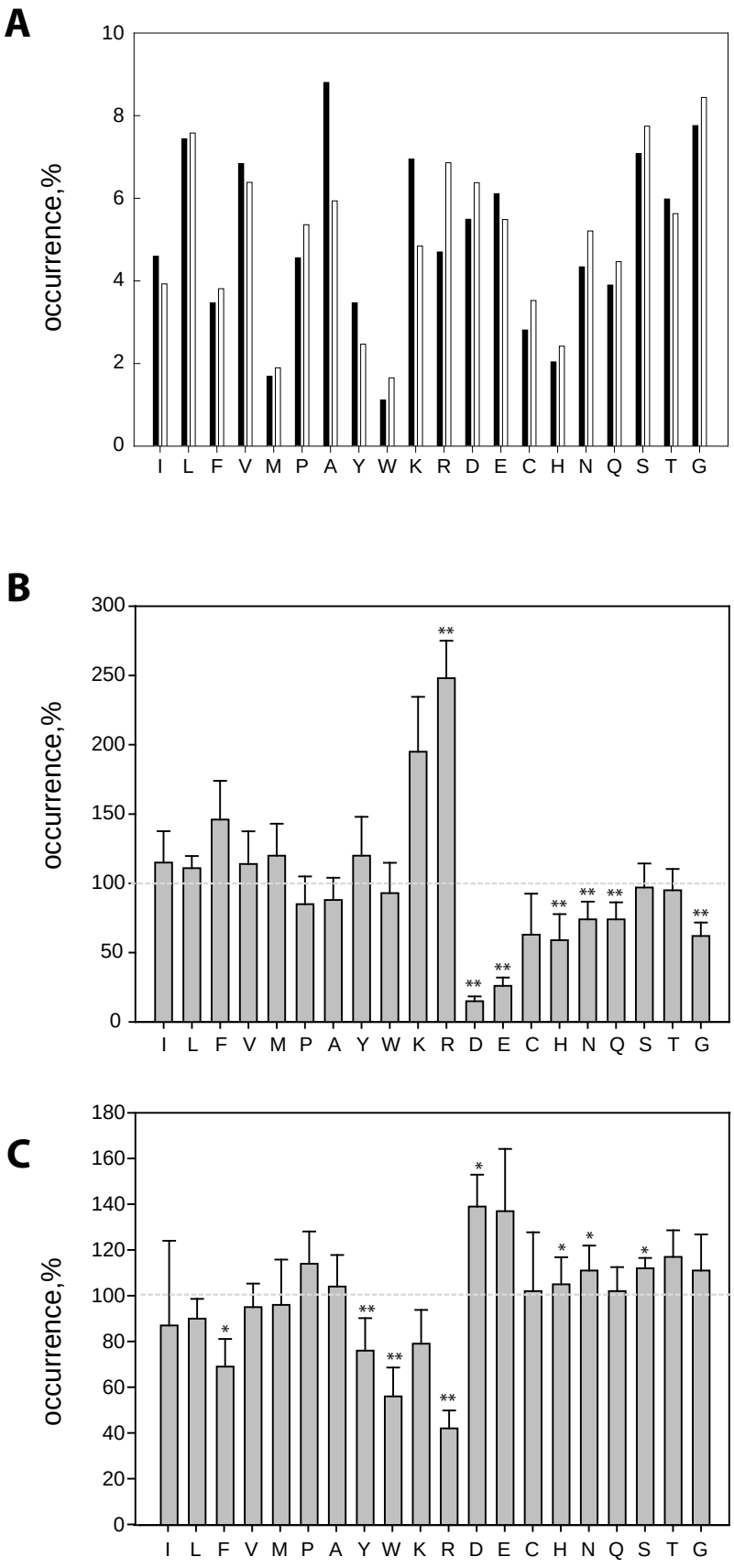

A
occurrence,%
0
2
4
6
8
10
I L F V M P A Y W K R D E C H N Q S T G
B
occurrence,%
0
50
100
150
200
250
300
**
I L F V M P A Y W K R D E C H N Q S T G
C
occurrence,%
0
20
40
60
80
100
120
140
160
180
*
**
I L F V M P A Y W K R D E C H N Q S T G

Fig. 2

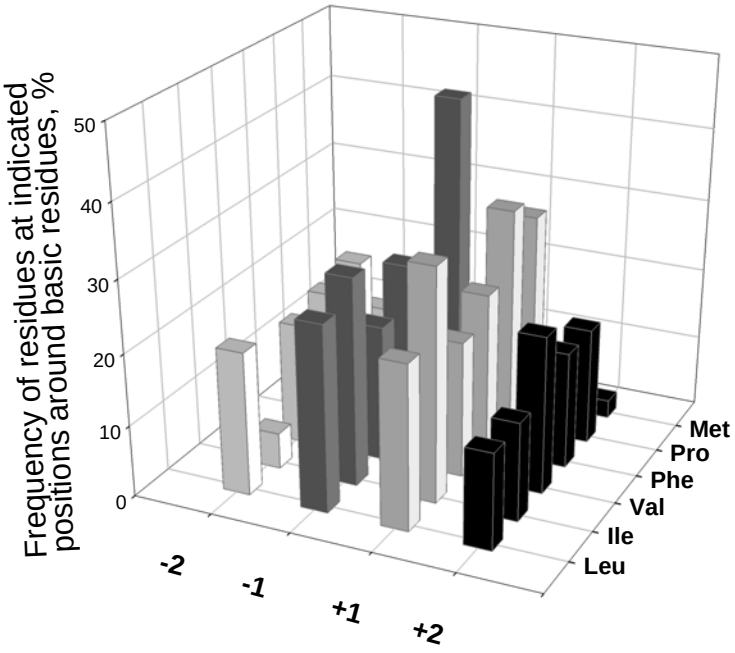

Frequency of residues at indicated positions around basic residues, %
50
40
30
20
10
0
-2
-1
+1
+2
Met
Pro
Phe
Val
Ile
Leu

Fig. 3

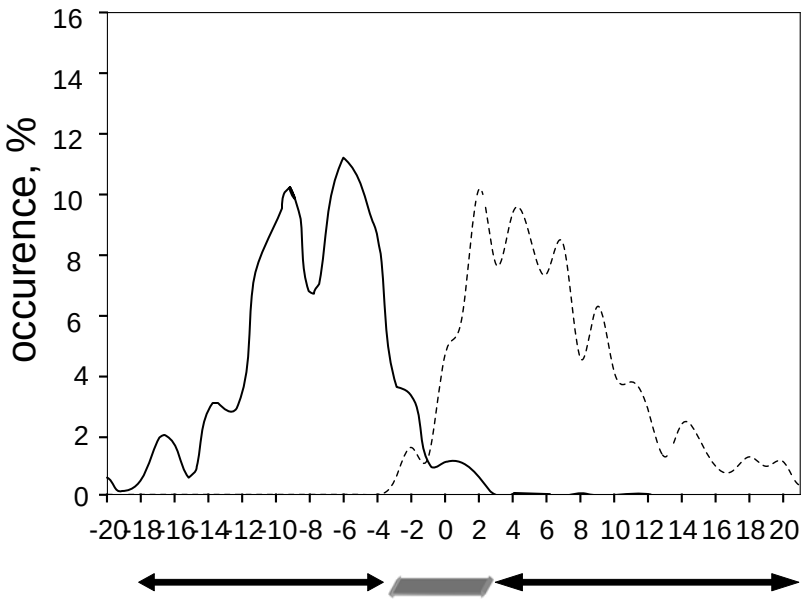
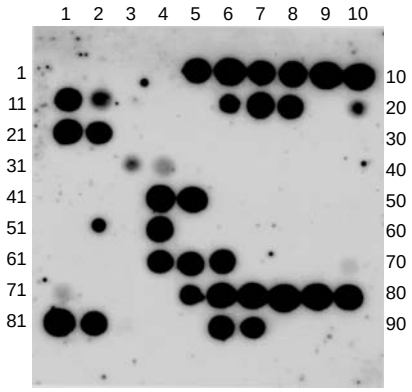

occurence, %
16
14
12
10
8
6
4
2
0
-20 -18 -16 -14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14 16 18 20

Fig. 4

A



B

1.	VQIHDIT	31.	IHLISTQ	61.	EGNCAEQ
2.	HDITGKD	32.	ISTQSAI	62.	CAEQDGS
3.	TGKDCQD	33.	QSAIPYA	63.	QDGSGWW
4.	DCQDIAN	34.	IPYALRV	64.	SGWWMNK
5.	DIANKGA	35.	ALRVELE	65.	WMNKCHA
6.	NKGAKQS	36.	VELEDWN	66.	KCHAGHL
7.	AKQSGLY	37.	EDWNGRT	67.	AGHLNGV
8.	SGLYFIK	38.	NGRTSTA	68.	LNGVYYQ
9.	YFIKPLK	39.	TSTADYA	69.	VYYQGGT
10.	KPLKANQ	40.	ADYAMFK	70.	QGGTYSK
11.	KANQQFL	41.	AMFKVGP	71.	TYSKAST
12.	QQFLVYC	42.	KVGPEAD	72.	KASTPNG
13.	LVYCEID	43.	PEADKYR	73.	TPNGYDN
14.	CEIDGSG	44.	DKYRLTY	74.	GYDNGII
15.	DGSGNGW	45.	RLTYAYF	75.	NGIIWAT
16.	GNGWTVF	46.	YAYFAGG	76.	IWATWKT
17.	WTVFQKR	47.	FAGGDAG	77.	TWKTRWY
18.	FQKRLDG	48.	GDAGDAF	78.	TRWYSMK
19.	RLDGSVD	49.	GDAFDGF	79.	YSMKKTT
20.	GSVDFKK	50.	FDGFDFG	80.	KKTTMKI
21.	DFKKNWI	51.	FDFGDDP	81.	TMKIIPF
22.	KNWIQYK	52.	GDDPSDK	82.	IIPFNRL
23.	IQYKEGF	53.	PSDKFFT	83.	FNRLTIG
24.	KEGFGHL	54.	KFFTSHN	84.	LTIGEGQ
25.	FGHLSPT	55.	TSHNGMQ	85.	GEGQQHH
26.	LSPTGTT	56.	NGMQFST	86.	QQHHLGG
27.	TGTTEFW	57.	QFSTWDN	87.	HLGGAKQ
28.	TEFWLGN	58.	TWDNDND	88.	GAKQAGD
29.	WLGNEKI	59.	NDNDKFE	89.	QAGDVAA
30.	NEKIHLI	60.	DKFEGNC		

C

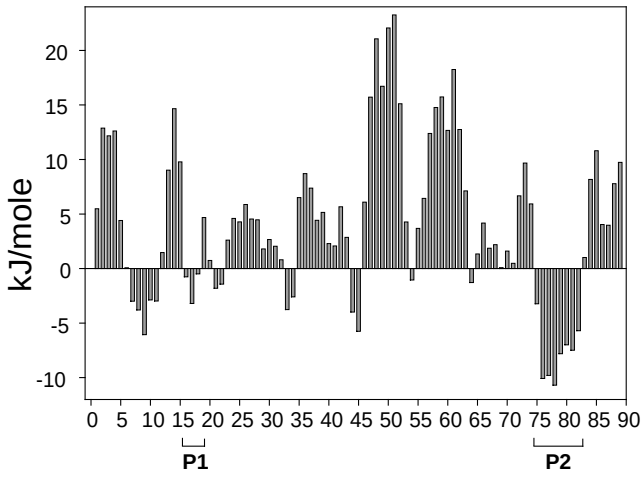
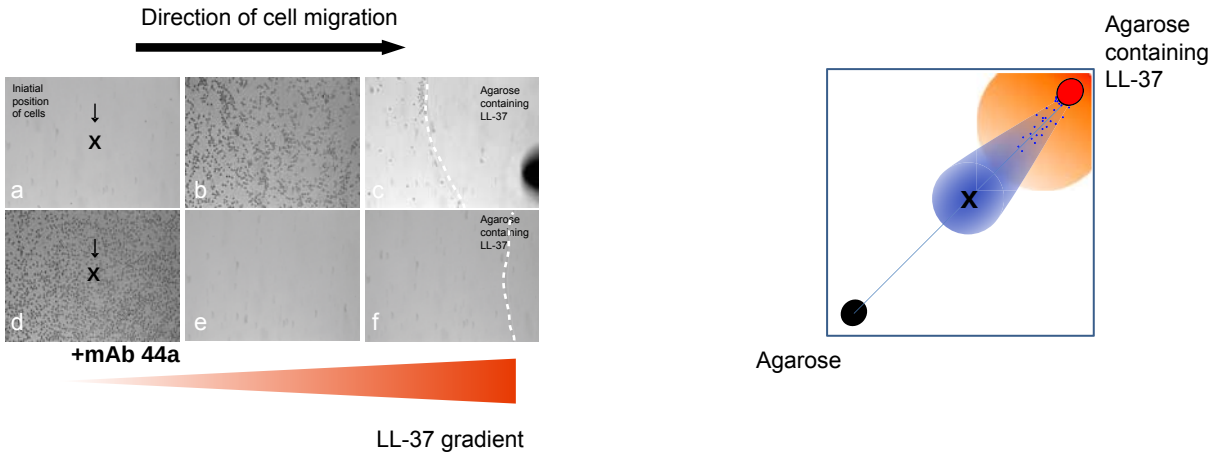


Fig. 5

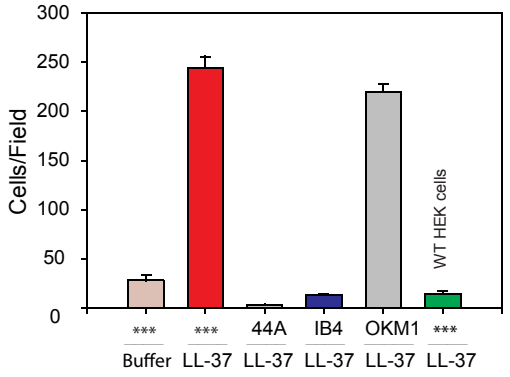
A

^{1}LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES^{37}

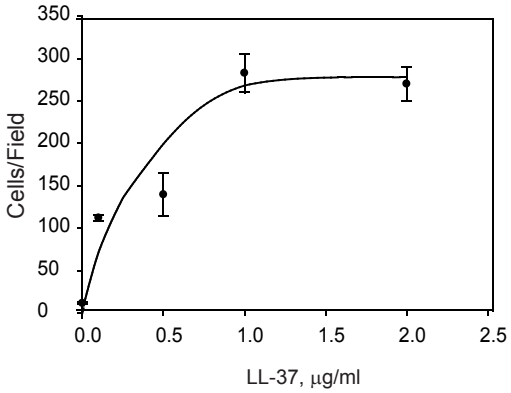
B



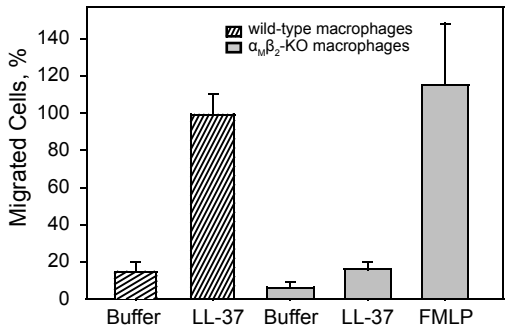
C



D



E



F

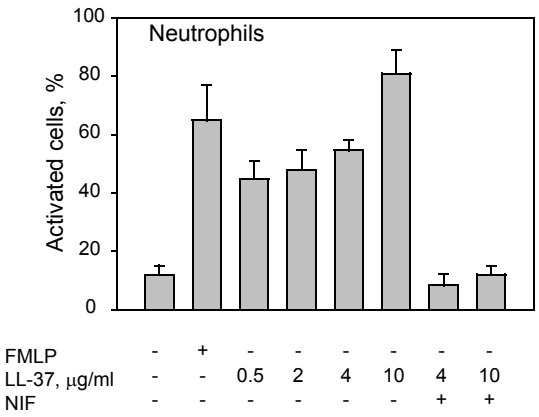


Fig. 6

Nataly P. Podolnikova, Andriy V. Podolnikov, Thomas A. Haas, Valeryi K. Lishko, and Tatiana P. Ugarova

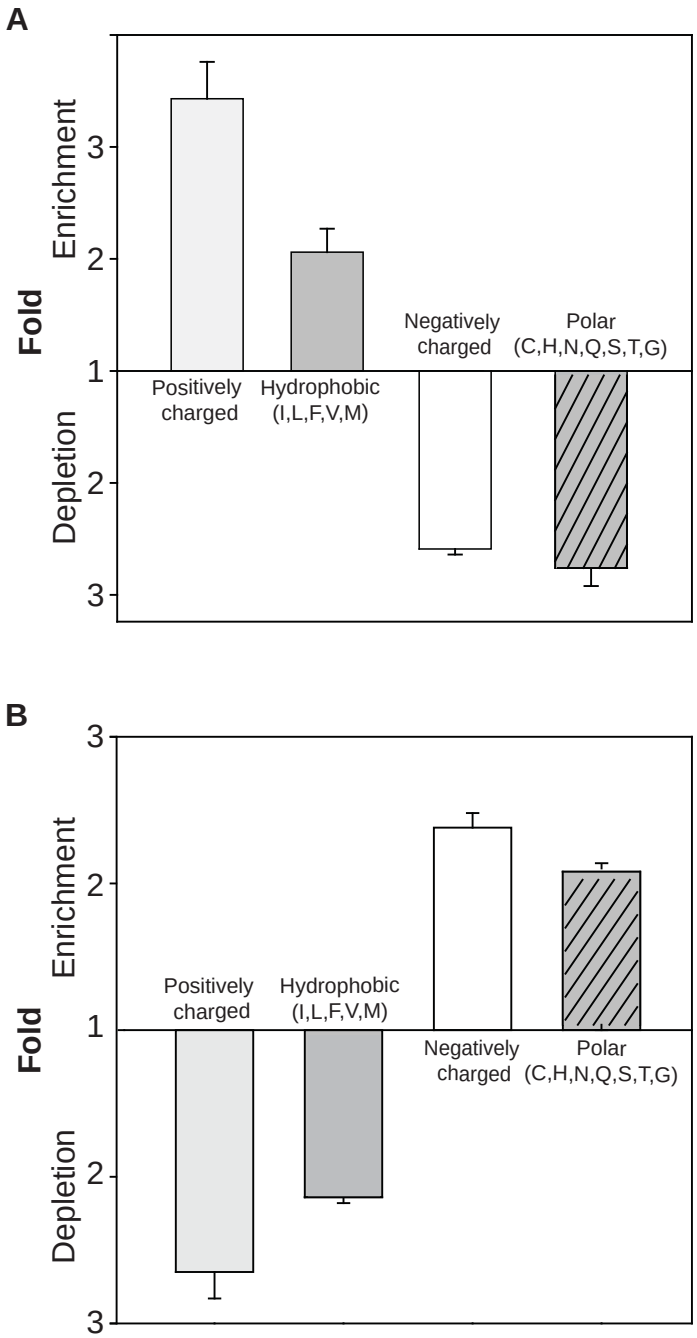


Figure S1. Enrichment or depletion of amino acids in strong binding (A) and non-binding (B) peptides, grouped by the chemical nature of their side chains. Graphs were constructed based on the data shown in Fig. 2.

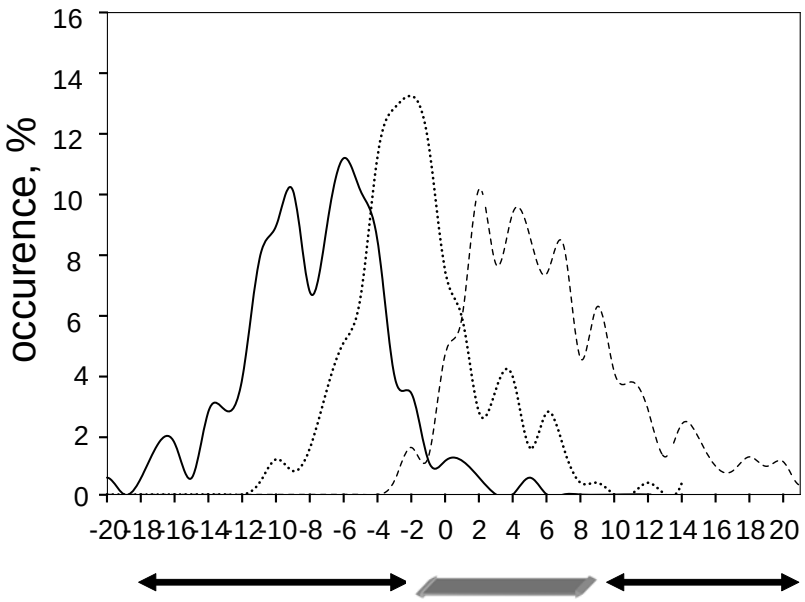


Figure S2. Distribution of energy values in populations of α_MI-domain binding and non-binding peptides. Energy distribution in populations of strong (—), good (....), and non-binding (---) α_MI-domain peptides was calculated as a percentage of the total peptides in each population.

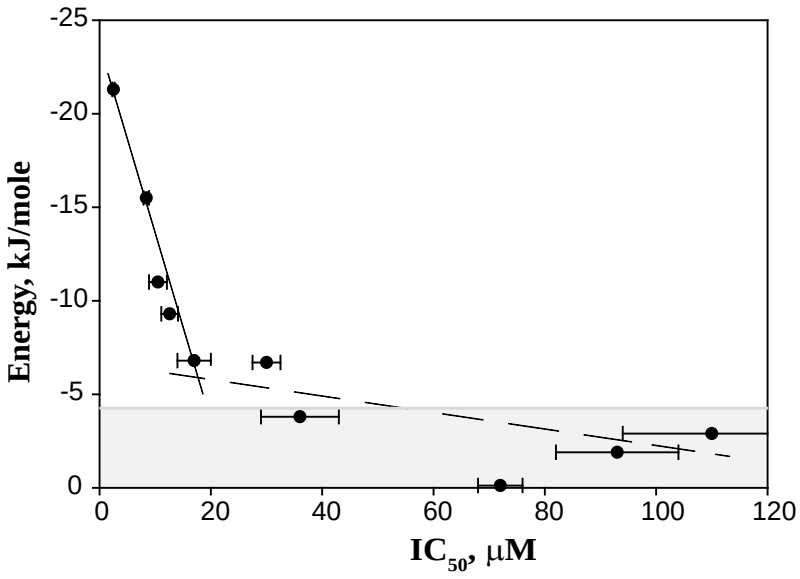


Figure S3. Correlation between the potency of the peptides (Table 3) to inhibit binding of the α_M-I domain to the D_{100} fragment and their energy scores.

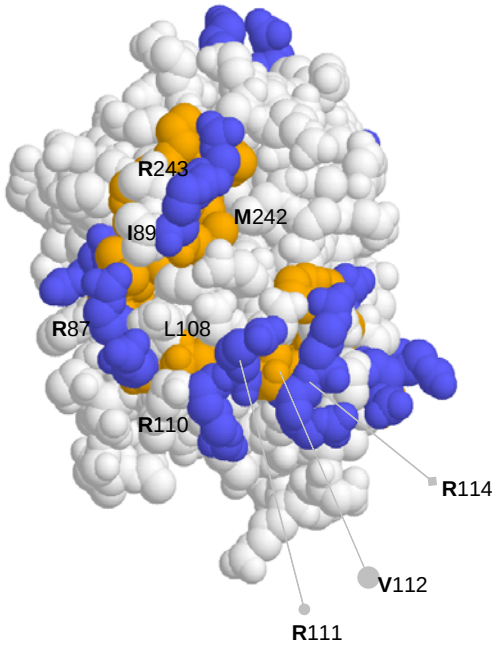


Figure S4. Position of the α_MI-domain binding regions in the 3D structure of cathepsin G. The α_MI-domain binding sequences identified in the peptide scan of human cathepsin G (Fig. 1) were mapped onto its crystal structure (PDB ID 1AU8) shown in a space-filling representation. Only the overlapping sequences within contiguous peptides which bind the α_MI-domain strongly ("strong binders") were selected for presentation. Basic and hydrophobic residues in the α_MI-domain binding clusters are colored in blue and yellow, respectively. Many residues in the α_MI-domain binding sites are exposed on the surface of cathepsin G. The side chains of selected arginines and hydrophobic residues within three segments, ARRAIR (spots 23-24), LSRRVRRNVP (spots 31-33) and IRMTTMR (spots 72-73), are shown. The figure was constructed using the Entrez's 3D-structure database (Chen, J., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He S., Hurwitz, D.I., Jackson, J.D., Jacobs, A.R., Lanczycki, C.J. etal., (2003). MMDB: Entrez's 3D-structure database. Nucleic Acids Res. 31, 474-477).

A

```
  1 MPKRKVSSAE GAAKEEPKRR SARLSAKPPA KVEAKPKKAA AKDKSSDKKV
 51 QTKGKRGAKG KQAEVANQET KEDLPAENGE TKTEESPASD EAGEKEAKSD
```

B

1	MPKRKVSSA	-7.7180
2	RKVSSAEGA	-0.2486
3	SSAEGAAKE	8.5381
4	EGAAKEEPK	9.1882
5	AKEEPKRRS	-4.3215
6	EPKRRSARL	-10.3250
7	RRSARLSAK	-13.7548
8	ARLSAKPPA	-3.6490
9	SAKPPAKVE	1.4631
10	PPAKVEAKP	1.7598
11	KVEAKPKKA	-4.3931
12	AKPKKAAAK	-6.6349
13	KKAAAKDKS	-1.5805
14	AAKDKSSDK	6.5503
15	DKSSDKKVQ	5.8487
16	SDKKVQTKG	1.3628
17	KVQTKGKR	-7.8982
18	TKGKRGAKG	-6.3565
19	KRGAKGKQA	-7.0198
20	AKGKQAEVA	1.9357
21	KQAEVANQE	7.6634
22	EVANQETKE	10.6196
23	NQETKEDLP	13.1137
24	TKEDLPAEN	12.9239
25	DLPAENGET	16.5559
26	AENGETKTE	12.1743
27	GETKTEESP	11.9938
28	KTEESPASD	13.3466
29	ESPASDEAG	16.9656
30	ASDEAGEKE	17.0343

C

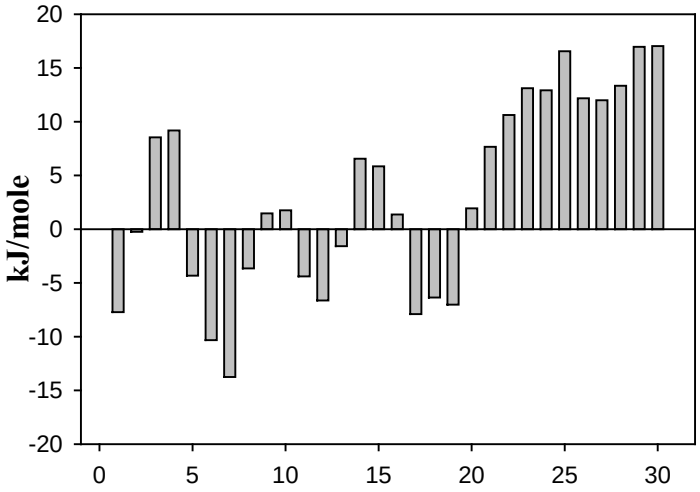


Figure S5. Application of the developed algorithm (IRMA) to predict the α_MI-domain binding sites in HMGN1. *A*, The sequence of human HMGN1. *B*, The peptide library derived from the sequence of HMGN1 consisting of 9-mer peptides. The numbers on the right are the energies (kJ/mole) calculated using the computer program IRMA. *C*, The energy value for each 9-mer peptide in the library is shown on the ordinate. The position of each peptide in the library is shown on the abscissa. Peptides with negative energy values are the most probable α_MI-domain binding sites.

TABLE S1

Sequences of strong α_MI-domain binders obtained from peptide libraries derived from 15 protein sequences

PROTEIN	
N-terminal Pg peptide	SLFSVTKKQ
	SVTKKQLGA
	RKSSIIIRM
	SIIIRMRDV
	VLFEKKVYL
Fibrinogen Aα chain	RGPRVVERH
	LRSRIEVLK
	RIEVLKRKV
	VLKRKVIEK
	IDIKIRSCR
	KIRSCRGSR
	STTTTRRSC
	TTRRSCSKT
	RSCSKTVTK
	GTLDGFRHR
	FPSRGKSSS
	RGKSSSYSK
	SSSYSKQFT
	YSKQFTSST
	STKRGHAKS
	RGHAKSRPV
	AKSRPVRGI
	RPVRGIHTS
Fibrinogen βC-domain	YQISVNKYR
	SWYSMRKMS
	SMRKMSMKI
	KMSMKIRPF
	MKIRPFFPQ
	RPFFPQQAA
Fibrinogen γC-domain	NKGAKQSGL
	YFIKPLKAN
	KPLKANQQF
	WTVFQKRLD
	DFKKNWIQY
	DKYRLTYAY
	RLTYAYFAG
	WMNKCHAGH
	IWATWKTRW
	TWKTRWYSM
	TRWYSMKKT
	YSMKKTTMK
	KKTTMKIIP
	TMKIIPFNR
	IIPFNRLTI
CCN1	TALKGICRA
	GRPCEYNSR
	CEYNSRIYQ
	NSRIYQNGE
	CPNPRLVKV
	PRLVKVTGQ
	GKGSSLKRL
	PRILYNPLQ
	QPVYSSLKK
	YSSLKKGKK
	LKKGKKCSK
	GKKCSKTKK
	PVRFTYAGC
	FTYAGCLSV
	AGCLSVKKY
	LSVKKYRPK
	KKYRPKYCG
	PQLTRTVKM
	TRTVKMRFR
	VKMRFRCED
	RFRCEDGET

PROTEIN	
Myeloperoxidase	RRESIKQRL
	SIKQRLRSG
	LLSYFKQPV
	LLERKLRSL
	RKLRSLWRR
	RSLWRRPFN
	ITGMCNNRR
	SNRAFVRWL
	GWTPGVKRN
	KRNGFPVAL
	RSLMFMQWG
	LLAVNQRFQ
	LTNRSARIP
	RSARIPCFL
	TSMHTLLLR
	AMVQIITYR
	LGPTAMRKY
	RKYLPTYRS
	PRIANVFTN
	ANVFTNAFR
	FTNAFRYGH
	AFRYGHTLI
	PNPRVPLSR
	RVPLSRVFF
	LSRVFFASW
	VFFASWRVV
	RGLMATPAK
	MATPAKLNR
	PAKLNRQNQ
	LNRQNQIAV
	PALNMQRSR
	LPGYNAWRR
	YNAWRRFCG
	WRRFCGLPQ
	LGTVLRNLK
	VLRNLKLAR
	NLKLARKLM
	PLKRKGRVG
	RKGRVGPLL
	IGTQFRKLR
Elastase	VGGRRARPH
	RRARPHAWP
	FMVSLQLRG
	ANVNVRAVR
	NVRAVRVVL
	AVRVVLGAH
	NLSRREPTR
	IVILQLNGS
	GRRLGNGVQ
	GLLGRNRGI
	GRNRGIASV
	TVVTSLCRR
	TSLCRRSNV
	CRRSNVCTL
	CTLVRGRQA
Cathepsin G	RPHSRPYMA
	HITARRAIR
	ARRAIRHPQ
	IMLLQLSRR
	LQLSRRVRR
	SRRVRRNRN
	VRRNRNVNP
	GRVSMRRGT
	REVQLRVQR
	QLRVQRDRQ
	RRERKAAFK
	LPWIRTTMR
	IRTTMRSFK

PROTEIN	
Azurocidin	IVGGRKARP
	GRKARPRQF
	ARPRQFPFL
	GALIHARFV
	IHARFVMTA
	LGAYDLRRR
	RRRERQSRQ
	SQRSGGRLS
	SGGRLSRFP
	RLSRFPRFV
	RFPRFVNVT
	RFVNVTVTP
	CTGVLTRRG
	VLTRRGGIC
	FFTRVALFR
Protein C	RAHQVLRIR
	QVLRIRKRA
	RIRKRANSF
	LAFWSKHVD
	VGWRRCSCA
	RRCSCAPGY
	VKFPCGRPW
	PCGRPWKRM
	RPWKRMEKK
	KRMEKKRSH
	EKKRSHLKR
	ESKKLLVRL
	KLLVRLGEY
	KEAKRNRTF
	KRNRTFVLN
	RTFVLNFIK
	VLNFIKIPV
	NYGVYTKVS
	VYTKVSRYL
ICAM-1	SFSAKASVS
	GTQRLTCAV
	RLTCAVILG
BSP	VYGLRSKSK
	LRSKSKKFR
	KFRRPDIQY
	SHKQSRLYK
	QSRLYKRKA
	KHLKFRISH
Proteinase 3	RLVNVVLGA
	VTVVTFFCR
	ICTFVPRRK
	FVPRRKAGI
	FFTRVALYV
	RSTLRAAAA
FALL-39	FALLGDFFR
	LGDFFRKSK
	FFRKSKEKI
	FKRIVQRIK
	RIKDFLRNL
	DFLRNLVPR
SBTI	GTIISSPFR
	ISSPFRIRF
	PFRIRFIAE
	LRLKFDSFA
	GTRRLVVSK
	RLVVSKNKP
	VSKNKPLVV

Strong α_MI-domain binders were identified based on densitometry analyses.

TABLE S2

Data used for construction of the α_MI-domain binding site prediction algorithm

Amino acid	Content in strong binders, %	Content in non-binders, %	ΔG_K (kJ/mole)
A	5.1	6.2	0.47
C	2.6	3.2	0.54
D	0.8	8.7	5.82
E	1.7	7.6	3.61
F	4.7	2.7	-1.29
G	5.5	9.3	1.30
H	1.6	2.5	1.18
I	4.2	3.3	-0.63
K	9.4	3.6	-2.35
L	8.2	6.9	-0.42
N	3.9	5.7	0.93
M	2.5	2.0	-0.52
P	4.5	6.1	0.74
Q	3.6	4.7	0.66
R	16.9	2.9	-4.29
S	7.6	9.1	0.46
T	5.5	6.6	0.49
V	6.9	5.8	-0.42
W	1.8	1.0	-1.42
Y	3.2	1.9	-1.24

The relative occurrence of amino acids in the α_MI-domain binding (196) and non-binding (748) peptide populations was used to calculate the statistical energy that each residue contributes to the α_MI-domain binding region (ΔG_K). Energy contributions of each amino acid values were determined according to $\Delta G_K = -RT[\ln Pb/Pn]$, where Pb and Pn are the relative occurrence of each amino acid in the binding and non-binding peptides calculated as described previously (Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* 16, 1501-1507).