



HAL
open science

RK, the first scorpion peptide with dual disintegrin activity on $\alpha 1\beta 1$ and $\alpha v\beta 3$ integrins

Oussema Khamessi, Hazem Ben Mabrouk, Houcemeddine Othman, Rym Elfessi-Magouri, Michel de Waard, Mejdoub Hafedh, Naziha Marrakchi, Najet Srairi-Abid, Riadh Kharrat

► To cite this version:

Oussema Khamessi, Hazem Ben Mabrouk, Houcemeddine Othman, Rym Elfessi-Magouri, Michel de Waard, et al.. RK, the first scorpion peptide with dual disintegrin activity on $\alpha 1\beta 1$ and $\alpha v\beta 3$ integrins. International Journal of Biological Macromolecules, 2018, 120 (B), pp.1777-1788. 10.1016/j.ijbiomac.2018.09.180 . pasteur-02003807

HAL Id: pasteur-02003807

<https://riip.hal.science/pasteur-02003807>

Submitted on 1 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

RK, the first scorpion peptide with dual disintegrin activity on $\alpha 1\beta 1$ and $\alpha \nu\beta 3$ integrins

Oussema Khamessi, Hazem Ben Mabrouk, Houcemeddine Othman, Rym ElFessi Magouri, Michel De Waard, Mejdoub Hafedh, Naziha Marrakchi, Najet Srairi-Abid, Riadh Kharrat



PII: S0141-8130(18)32661-8
DOI: [doi:10.1016/j.ijbiomac.2018.09.180](https://doi.org/10.1016/j.ijbiomac.2018.09.180)
Reference: BIOMAC 10615

To appear in: *International Journal of Biological Macromolecules*

Received date: 6 June 2018
Revised date: 7 September 2018
Accepted date: 27 September 2018

Please cite this article as: Oussema Khamessi, Hazem Ben Mabrouk, Houcemeddine Othman, Rym ElFessi Magouri, Michel De Waard, Mejdoub Hafedh, Naziha Marrakchi, Najet Srairi-Abid, Riadh Kharrat , RK, the first scorpion peptide with dual disintegrin activity on $\alpha 1\beta 1$ and $\alpha \nu\beta 3$ integrins. *Biomac* (2018), doi: [10.1016/j.ijbiomac.2018.09.180](https://doi.org/10.1016/j.ijbiomac.2018.09.180)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

RK, the first scorpion peptide with dual disintegrin activity on $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins

Oussema Khamessi¹, Hazem Ben Mabrouk¹, Houcemeddine Othman¹, Rym ElFessi Magouri¹, Michel De Waard³, Mejdoub Hafedh², Naziha Marrakchi¹, Najet Srairi-Abid¹ and Riadh Kharrat¹

¹ Université de Tunis El Manar, Institut Pasteur de Tunis, LR11IPT08 Venins et biomolécules thérapeutiques, 1002, Tunis, Tunisie.

² USCR séquenceur de protéines, faculté des sciences de Sfax, Route de Soukra, Km 3.5, BP 1171, 3000 Sfax, Tunisia.

³ Inserm U1087, Institut du Thorax, groupe Iib, Université de Nantes, 8 quai moncoussu, 44000, Nantes, France; Smartox Biotechnology, 570 rue de la chimie, bâtiment Nanobio, 38700, Saint Martin d'Hères, France.

Correspondance to: RIADH KHARRAT. Laboratoire des Venins et Biomolécules thérapeutiques Institut Pasteur de Tun. 13, Place Pasteur, BP74. 1002, Tunis - Belvédère, Tunisia. Tel.: (216) 71 783 022. Fax: (216) 71 791 833. E-mail: riadh.kharrat@pasteur.tn

Abstract:

Scorpion peptides are well known for their pharmaceutical potential on different targets. These include mainly the ion channels which were found to be highly expressed in many diseases, including cancer, auto-immune pathologies and Alzheimer. So far, however, the disintegrin activity had only been characterized for snake venom molecules. Herein, we present the first short peptide, purified from the venom of *Buthus occitanus tunetanus*, (termed RK) able to inhibit the cell adhesion of Glioblastoma, Melanoma and Rat pheochromocytoma to different extracellular matrix (ECM) receptors. Anti-integrin antibody assay suggests that RK interacts with both $\alpha_1\beta_1$ and $\alpha_v\beta_3$ with a more pronounced effect for the former. The examination of the primary structure of RK suggests the involvement of two motifs: KSS, analogue to KTS which was characterized for $\alpha_1\beta_1$ Snake venom disintegrins, and ECD, analogue to RGD which was found to be active on $\alpha_v\beta_3$. To assess their roles in the disintegrin activity of RK, we conducted a computational analysis. The molecular docking study shows that RK involves mainly two segments to interact with the $\alpha_1\beta_1$ integrin, but the peptide does not implicate the KSS motif in the interaction. The molecular modeling study, suggests the key contribution of the ECD segment in the interaction with $\alpha_v\beta_3$ integrin.

Key words: RK peptide; Adhesion; Molecular modeling.

1 Introduction

Cancer is a bottleneck disease that implicates different physiological processes to exhibit its pathological aspect [1]. In the last decades, it was demonstrated that different types of integrins, metalloproteinases and ion channels are highly expressed by many types of tumors. As a result, these proteins emerged as interesting biomarkers to detect cancers and as potential pharmacological targets to treat them [2]. Particularly, integrins, transmembrane heterodimeric proteins of non covalently associated α and β subunits, are implicated in all the processes of carcinogenesis [3]. The association between their α and β subunits forms adhesion receptors which binds to the extracellular matrix and provide critical adhesive and signalling functions [4]. Integrins are capable to affect cellular functions [5] such as cytoskeleton organization, transduction of intracellular signals [6], cell differentiation, growth, and apoptosis [7-8]. Different subtypes of integrins have been characterized on the basis of their interactions with different motifs of the extracellular matrix implicated in cell adhesion [9]. These motifs coordinate the biological responses between endothelial cells, tumor cells and the extracellular matrix [10]. In fact, these integrins were first identified on the basis of their ability to recognize the RGD amino acid sequence [11-12], and are implicated in cell proliferation, invasion and viability. Their ligand binding function is also dependent on the presence of the metal ion (Mg^{2+} , Mn^{2+} and Ca^{2+}) [13].

Integrins bind to collagen by using their αI Domain, $\alpha 1\beta 1$ and $\alpha 2\beta 1$ represent the most known collagen receptors, which are members of integrin family and are structurally very similar [14]. Their ligand binds to a Mg^{2+} ion in the Metal Ion-Dependent Adhesion Site (MIDAS) [15]. $\alpha 1\beta 1$ integrin is the principal collagen IV receptor [16] but its contribution in tumor formation and progression is poorly defined compared to $\alpha v\beta 3$ and $\alpha 5\beta 1$. The critical role of $\alpha 1\beta 1$ integrin in tumorigenicity was demonstrated by works elaborated in colon cancer cells that associate with talin and paxillin, resulting in the promotion of cancer cell invasion [17]. The ability of multiple polypeptides to inhibit the integrin activity was previously discovered for different molecules purified mainly from snake venoms called disintegrins. Disintegrins were proven to have anti tumour effects involving angiogenesis and cancer metastatic dissemination. The functional classification of disintegrins depends on their ability to interact with specific integrins [18], which is determined by the presence of a particular integrin-binding motif localized in the hairpin loop, unless they are present in the same fold. Functionally, disintegrins can be divided into three classes containing RGD, MLD, and R/KTS motifs have been identified [19]. RGD-disintegrins block $\alpha v\beta 3$, $\alpha v\beta 1$, $\alpha 5\beta 1$ and $\alpha 8\beta 1$ integrins. MLD inhibits the physiological functions of $\alpha 3\beta 1$, $\alpha 6\beta 1$, $\alpha v\beta 6$, $\alpha 7\beta 1$, $\alpha 4\beta 1$, $\alpha 4\beta 7$ and $\alpha 9\beta 1$ integrins. R/KTS-disintegrins are a potent and selective inhibitors of $\alpha 1\beta 1$ integrin [20]. The extensive screening of many scorpion venoms showed a great diversity in their composition with different pharmacological potentials. Recently, and for therapeutic purposes, special emphasis was given to bioactive peptides which have anticancer activity.

Only a few purified toxins seem to be endowed by the anticancer effects. The Chlorotoxin (CTX) isolated from *Leiurus quinquestriatus* scorpion venom [21], that have been shown to bind specifically to glioma cell surfaces as a specific chloride channel blocker and is currently in phase II of human trials [22]. Today there are nearly 970 scorpion peptides (available in public databases) described by the animal toxin annotation project *uniprot* [23]. Scorpion peptides are disulfide bond rich molecules presenting a sequence length ranging between 23 and about 80 amino acids. The majority of scorpion peptides have the sequence signature of the cystine stabilized α/β (CS α/β) motif [24]. Some scorpion CS α/β peptides display a remarkable specificity for certain subtypes of ion channels [25]. Scorpion venom contains also a significant number of peptides, without the disulfide bridge, that exhibit antimicrobial, immunomodulatory, Bradykinin-potentiating and/or hemolytic activities [26]. We have recently described the first short 14 amino acid peptide (RK1 peptide), from *Buthus occitanus tunetanus* venom that inhibits tumor cell migration, proliferation and angiogenesis [27]. In this work, we isolated a new disintegrin-like homologous peptide from the same venom, targeting cancer cell adhesion by acting on integrins. We investigated its anti-tumoral activity and its potential interaction with different integrins both *in vitro* and *in silico*.

2 Materials and methods:

2.1 Scorpion venom

Buthus occitanus tunetanus venom was provided in a liquid state by electric stimulation of the post-abdomen of the scorpion, bred in the Beni Khedach area (Tunisia). The pooled venom is kept frozen at -20 °C in its crude form until use. All reagents were purchased from Sigma- Aldrich® Chemical company.

2.2 Purification of RK

Purified RK was obtained from the scorpion venom *Buthus occitanus tunetanus* by gel filtration G50 followed by HPLC. Crude venom was dissolved in water and loaded onto a sephadex G50 column equilibrated with 0.1M ammonium acetate, pH 8.5. Different fractions were eluted and tested for their toxicity on mice. Only a fraction (BotG50) showing a toxic activity [28-29] was then applied onto semi preparative reversed-phase HPLC C8 column (10 mm x 250mm, 5 μ m, Beckman Fullerton) equilibrated in 0.1% trifluoroacetic acid in water, at a flow rate of 1ml/min. HPLC purification of the non-retained fraction was performed using an analytical C18 reversed-phase HPLC column (4.6mm x 250 mm, 5 microns Beckman). Detection was monitored at 214 nm.

2.3 Mass spectrometry

RK was analyzed by MALDI-TOF mass spectrometer (Perspective Biosystems, Inc., Framingham, MA). The sample was dissolved in CH₃CN/H₂O (30/70) with 0.3% trifluoroacetic acid to obtain a concentration of 1-10 pmol.μL⁻¹.

The matrix was prepared as follows: alpha-cyanohydroxycinnamic acid was dissolved in 50% CH₃CN in 0.3% trifluoroacetic acid/H₂O to obtain a saturated solution at 10 μg μL⁻¹. A 0.5 μL peptide solution was placed on the sample plate and 0.5 μL of the matrix solution was added. This mixture was allowed to dry. Mass spectra were recorded in linear mode, calibrated with suitable standards and were analyzed by GRAMS/386 software.

2.4 Amino acid sequence determination and peptide synthesis

Reduction and alkylation of peptides and sequence determination of native and S-alkylated peptide were performed as described by Khammessi et al [27-30]. The RK peptide was synthesized by Genosphere Biotechnologies (Paris, France). Automated synthesizerd by the Fmoc method on solid phase (Merrifield, R. B. 1969. Solid-phase peptide synthesis. *Adv. Enzymol. Relat. Areas Mol.Biol.*32:221.6). Analytical RP-HPLC and MS confirmed the purity (≥ 95%) and molecular mass of the synthesized peptide. The peptide stock was diluted in distilled water and kept at -20 °C, and quantified by a Nano DropND-1000 V3.5.2 Spectrophotometer at 280 nm. The reduced peptide was solubilized in 0.2 M Tris/HCl, pH 8, at 5 mM and stirred under air to allow folding (48 h, 25 °C). The target oxidized product was purified to homogeneity by reversed-phase liquid chromatography. The chemical identity of synthetic RK peptide was confirmed by mass determination and the elution with natural RK peptide after simultaneous injection of both products on analytical reversed-phase HPLC.

2.5 In vivo toxicity test

The in vivo toxicity of RK was tested on 20±2 g C57/BL6 male mice by intracerebro-ventricular (I.C.V) route considered to be the most sensitive route to mammals for scorpion toxins [31]. RK was diluted in 0.1% (w/v) BSA solutions containing increasing amounts of the peptides in a volume 5μL (1, 10, 50, 100 μg/kg). Six mice were used for each dose; two control mice were injected with only 0.1% BSA in water, to be sure that symptoms are not due to experimental conditions. Effects were monitored during 24 hours. All procedures met with the approval of the Institutional Research Board of the Pasteur Institute of Tunis. I.C.V administration was performed under ether anesthesia [32].

2.6 Cell culture

RPMI 1640 medium, Dulbecco's modified Eagle's medium (DMEM), fetal calf serum (FCS), and modified Eagle's medium (MEM) were purchased from Lonza (Walkersville, MD). The human

fibrinogen, laminin-1 and p-bromophenacyl bromide are from Sigma (Mannheim, Germany). Rat type I and IV collagen are obtained from Upstate (Lake Placid, NY) and human fibronectin and fibrinogen from Chemicon (Temecula, CA). Human vitronectin was purified according to Yatogho et al [32]. Rabbit anti-rat was purchased from Sigma. Cells were first cultured in MEM medium supplemented with 10% FBS containing 100 IU/ml penicillin. All cell lines were maintained at 37°C in 5% CO₂. Rat pheochromocytoma (PC12) cells were cultured in DMEM supplemented with 5% FCS and 10% heat inactivated horse serum. Human leukemia (K562) cells were cultured in RPMI 1640 medium containing 10% FCS. The human cell lines derived from fibrosarcoma (HT1080), melanoma (IGR39) and Human colon adenocarcinoma (HT29-D4) were routinely cultured in DMEM containing 10% FCS. All cell lines were graciously provided by José Luis (Laboratory of Cell Biology, Faculty of Pharmacy, Marseille).

2.7 Cell viability test

Cell viability was assessed by MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay [33]. MTT solution (500 µg/ml) was added to the culture medium 4 h before the end of treatment. Cells were incubated with RK for 72 h, then fixed with 1% glutaraldehyde, stained with a solution of 0.1% crystal violet, and lysed with 1% SDS. Absorbance was then measured at 560 nm. A control was used in the same conditions, but without RK.

2.8 Cell Adhesion

Cells, in single cell suspension, were added in 96-well plates coated with 10 µg/ml of fibronectin, type IV collagen or with 50 µg/ml fibrinogen, for 5 h at 37 °C, as described previously [34-35]. They were allowed to adhere to the substrate for 90 min (for U87, PC12, HT1080 and IGR39 cells) or 2 h (for HT29-D4 and K562 cells) at 37 °C. Adhesion of K562 cells was performed in the presence of 1mM MnCl₂ and 100 nM phorbol 12-myristate 13-acetate (PMA), in order to activate $\alpha_5\beta_1$ integrin. After the achievement of the cell adhesion to the plate wells, we tested the adhesion activity in the presence and in the absence of the RK peptide. After washing, adherent cells were fixed, stained with 0.1% crystal violet and lysed with 1% SDS. Absorbance was then measured at 560 nm. For adhesion assays, using blocking antibodies, 96-well plates were coated with 50 µl of rabbit anti-rat IgG (50 µg/ml), overnight at 4 °C. Wells were washed once with PBS and 50µl of anti-integrin blocking antibodies (10 µg/ml) were added for 5 h at 37°C. Then, wells were blocked with a solution of PBS/0.5% BSA and adhesion assay was performed as above.

2.9 RK model building

The primary structure of the RK peptide was submitted to PEP-FOLD server [36] to establish a de novo structure model for the peptide. We also parameterized the building to include the disulfide bond formation between C₃ and C₁₂ residues. The INPUT sequence was then submitted and the returned results were visually explored using PyMOL molecular Viewer [37].

2.10 Docking ensemble preparation

Two conformational ensembles of the RK peptide (ligand) and the domain I of $\alpha 1$ integrin (DI $\alpha 1$) as a receptor were prepared for the cross-docking. The conformational ensemble of the ligand was constructed by selecting the structures corresponding to the lowest energies belonging to predefined local minimum. All the structures within a difference of 1 kcal/mol of free energy are then selected. Clustering analysis applying a hierarchical algorithm was then applied in partitioning the structures according to their conformational similarity using a cutoff of 0.15 nm. Each center of these clusters is then retained to undergo the docking study. To construct the conformational ensemble for the receptor, we selected two structures of the DI $\alpha 1$ domain corresponding respectively to an X-ray crystallography structure representing the unbound state [13-38], and the NMR conformers of the bound state for a collagen mimetic peptide (2M32) [39]. For the latter, we removed the coordinates of the ligand prior to the docking. The crystal structure was submitted to the EnCoM web server in order to improve the exploring of the receptor flexibility [40]. The returned structure were submitted to a clustering analysis using a hierarchical algorithm with 0.2 nm cutoff.

2.11 Peptide-protein docking

The peptide-protein docking was effected using ZDOCK 3.0.1 [41]. For each docking, we retained the 20000 putative complexes ranked by the default scoring function. All of these underwent a re-scoring with ZRANK and we selected the best 2000 solutions [42]. Subsequently, we clustered the docking poses according to the coordinates of the ligands. Only the clusters with more than 20 members were retained for further refinement using FoldX [43]. Then the free energy of interaction for each putative solution was estimated. The best 20 solutions were retained to proceed to 20 ns molecular dynamics simulations using the same protocol used previously for RK peptide. The most stable complexes according to the RMSD time evolution (i.e those where the system is equilibrated) were then selected. We proceeded then to an additional 8 ns simulation and we re-evaluate the stability of the complexes from which the best solution was selected.

2.12 Molecular dynamics

The molecular dynamics simulation was run using Gromacs 5.1. The Gromos 53a6 force field [44] was used to assign the molecular mechanics parameters to the RK peptide atoms. The simulation

system consists of an octahedron box in which RK peptide occupies its center with a 12 Å cutoff between any solute atoms and the edge of the box. The system was solvating with SPC water type and was brought to neutrality using an appropriate number of counter-ions. To remove the bad contacts resulting from constructing the system, we apply an energy minimization stage in conjunction with positional restraints on the heavy atoms of the system. It consists of 1000 steps of steepest descent minimization. All the positional restraints were then removed before running a second stage of conjugate gradient energy minimization of 200 steps. Particle Mesh Ewald approximation was applied and a cutoff of 12 Å was used to reduce the computation concerning the non bonded interactions. The hydrogen atoms bonds were modeled under the LINCS containing algorithm to allow the use of a 2 fs (femtosecond) time step in running the molecular dynamics. The system was then equilibrated under NVT conditions were position restraints were applied for 40 ps (picosecond). They were then removed and the simulation continued for another 20 ps until the desired temperature was reached (300 K). The simulation was then resumed in the NPT ensemble by activating the Parrinello-Rahman pressure barostat with a coupling time of 2 ps during 100 ps. This was sufficient for the system to reach the desired pressure of 1 bar. The production stage was then carried for 300 ns time to generate the trajectory.

2.13 Statistical analysis:

All values are expressed as mean \pm standard deviation (SD). All the statistical significance of differential findings between experimental and control groups was determined by 2way ANOVA test except for the dose/response assay of RK disintegrin activity on the cell lines which was computed using a Student's t test. The analysis was made using Graph Pad Prism 5.0 software. $P < 0.05$ was considered statistically significant and is indicated with asterisks over the value (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$).

3 RESULTS

3.1 Purification of the peptide

RK was substantially purified from the *Buthus occitanus tunetanus* scorpion venom by a first stage of gel filtration on a Sephadex G-50 column chromatography as previously described [29]. The toxic fraction (BotG50) obtained from this separation was purified by high performance liquid chromatography (HPLC) using a semi preparative C8 column. The fraction eluting at 19-22 min (Fig. 1A), was further purified (using an analytical C18 reversed-phase HPLC column. The component eluted at 33.5 min (Fig. 1B) was homogeneous, as indicated by mass spectrometry analysis and amino acid

sequencing. It is designated RK. An analytical HPLC run of RK showed a single symmetric peak (Fig. 1C).

3.2 Sequence determination of RK and peptide synthesis

(RKs) The automatic Edman degradation of 25 pmol of native RK was performed with a reproducible yield of 95% during 17 cycles. 50 pmol of S-alkylated-protein was used to confirm without ambiguity the determined sequence, and to identify cystein positions. RK is a short peptide composed of 17 amino acid residues containing two cystein residues (Fig. 1D). The experimental molecular mass of native RK (1780.03 Da) obtained by MALDI-Tof mass spectrometry (supplementary data 1), is nearly identical to the average theoretical molecular mass calculated for the peptide with Oxidize cysteines form (1780.021 Da). This shows that the two cysteine residues at position 3 and 12 are involved in intramolecular disulfide bonds. Due to its very low concentration in venom (0.01% of the proteins), RK was chemically synthesized. After renaturation, the synthetic RK (RKs) was eluted with natural RK on analytical C18 reversed-phase HPLC. Mass spectrometry RKs gave an experimental value of 1780.05 Da, which is in agreement with the natural RK molecular mass.

3.3 In vivo toxicity and tumor cells viability

Evaluation of the toxicity of RK and RKs revealed that both peptides do not exhibit any toxicity up to 100 µg/kg body weight as determined by intracerebro-ventricular injection during 24 hours of monitoring the mortality and the side effects. RKs was tested on the viability of glioblastoma (U87), melanoma (IGR-39) and Rat pheochromocytoma (PC12) cell lines using the MTT assay. At 100 µM concentration, RKs does not affect the viability of all three cell lines including after 24h, 48h or 72h incubation time (Fig 2).

3.4 RKs inhibits cell adhesion

In order to investigate the effects of RKs, we performed cell adhesion assays by using a large array of purified ECM proteins (collagen IV (Coll IV), fibronectin (Fn), fibrinogen (Fg), laminin-1 (Lam)) and the non-specific substratum Poly-L-Lysine (PLL). RKs blocked notably the adhesion of human glioblastoma cells U87 (Fig. 3A) and melanoma cells IGR39 (Fig. 3B) to fibrinogen and fibronectin, while no effect could be observed on collagen IV or laminin. RKs was able to inhibit adhesion of Rat pheochromocytoma cells (PC12) only to type IV collagen (Fig. 3C). Moreover, no inhibition could be observed on the integrin-independent substratum, poly-L-lysine, suggesting that the effect of RKs may involve the integrin family as adhesion receptors. As shown in Fig. 4A, RKs has the highest activity on PC12 cell adhesion to collagen IV. Inhibition of adhesion occurs in a dose-dependent manner with an

IC₅₀ value of 4.84 μ M. RKs blocked notably the adhesion of human glioblastoma cells U87 cells to fibrinogen and fibronectin with IC₅₀ values of 15 μ M and 15.33 μ M respectively (Fig. 4B). However the inhibitory effect of RKs on IGR39 cells on fibrinogen and fibronectin does not exceed 40% (Fig. 4C).

3.5 RKs activity is selective to $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins

To identify the possible targeted integrins, we checked the RKs effect on various cell/ECM protein pairs involving in each case a unique integrin:

$\alpha_1\beta_1$ (PC12/type I collagen), $\alpha_2\beta_1$ (HT1080/ type I collagen), $\alpha_5\beta_1$ (K562/fibronectin), $\alpha_v\beta_5$ (HT29-D4/vitronectin), $\alpha_v\beta_6$ (HT29-D4/fibronectin), $\alpha_6\beta_4$ (HT29-D4/laminin) and $\alpha_v\beta_3$ (HT29-D4 transfected by 3 subunit/fibrinogen). As illustrated in Fig. 5A, RKs was not able to alter cell adhesion through $\alpha_2\beta_1$, $\alpha_v\beta_6$, $\alpha_v\beta_5$, $\alpha_5\beta_1$ and $\alpha_6\beta_4$ integrins, but significantly reduced the adhesive function of $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins. We tested the effect of function-blocking antibodies against integrins $\alpha_1\beta_1$, $\alpha_v\beta_3$ and $\alpha_5\beta_1$ on U87 cell adhesion to RKs used as a substrate. As shown in Fig. 5B, only antibodies against $\alpha_1\beta_1$ (60%), and $\alpha_v\beta_3$ (20%) abolished U87 cells attachment. Interestingly, the synthetic RGD peptide (1 mM) was able to inhibit more than 80% of human glioblastoma cell adhesion on RKs (Fig. 5B). These results suggest that the interaction between RKs and cells might involve an RGD-like motif.

3.6 Computational study

To understand the structural behavior of the RK peptide, we constructed a tridimensionnel structure model using PEP-FOLD server which applies a de novo approach. Since, we were not able to detect any significant homologous structure in the Protein Data Bank. PEP-FOLD server returned five different structures for RK. We selected the one which presents a disulfide bond closure between residues C3-C12. The C α (The alpha carbon) distance between each of these amino acids is about 5.9 Å (Angstrom) which still in the cutoff range of disulfide bond formation (4.4 to 6.8 Å). The structure of the RK peptide consists of 3 sub-segments. The first one corresponds to the dipeptide I₁D₂. The second one is the segment flanked by the two cystein residues (C₃GTVMIPSEC₁₂) which harbors a one-turn α helix consisting of residues 9-12. The third sub-segment is the C-terminal end corresponding to the DPKSS sequence (Fig. 6A). Except for the two cystein residues, all the amino acids of RK are solvent exposed. The Accessible surface area calculated for each residue using GETAREA shows a minimum of exposure for E₁₁ (53%) and a maximum (100%) for T₅, P₈, S₁₆ and S₁₇.

3.7 Molecular dynamics of the RK peptide

To gain more insight of the conformational properties of RK, we run a molecular dynamics simulation of 300 ns (nanosecond). The RMSD (Root Mean Square Deviation) of the peptide C α atoms shows

broadly three phases on the plot. The first phase is very short, consisting of an equilibration ending at the 8th nanosecond of the simulation. The second plateau phase lasts between the 8th and the 52nd nanosecond where the RMSD fluctuate approximately around an average value of 3.8 Å relative to the starting structure. Until the end of this phase, the peptide preserves the overall conformation of the initial structure, mainly the one turn α helix. The following phase lasts for the rest of the trajectory where the RMSD values increased significantly to reach an average of 6.8 Å but the values can reach more than 7 Å. In fact, The RMSD at this phase is not as stable as its predecessor. We also noticed that RK peptide loses the overall initial conformation at this phase, in which the segment between the two cysteins shows a disordered structure (Fig. 6B). To characterize the flexibility of RK, we calculated the RMSF values per C α atom. The C-terminal end of the peptide consisting in S₁₆ and S₁₇ residues is the most flexible segment of RK. These amino acids show an RMSF values of 4.8 and 7.0 Å, respectively (Fig. 6B). Residues 12-15 are the least flexible with RMSF (Root Mean Square Fluctuation) values ranging between 2.7 and 4.8 Å. The RMSF values for the residues of central segment, 4-12, seem to be more uniform which could be the result of a collective behavior of these amino acids. In addition to the previous analysis, we established the energy landscape of the peptide based on the radius of gyration and the RMSD values as reaction coordinates. The Fig. 6C shows that the peptide describes mainly two low energy wells. The first one is the largest in which the conformation with the lowest energy preserves the one turn α helix. The well is highly populated and corresponds to a global minima. The second well is not as much populated as the first well in which the conformations present low energy values and corresponds to local minima. The two wells are separated by two other metastable local minima. The well representing the global minima of the free energy landscape of RK structure served to construct the docking ensemble of the ligand.

3.8 Peptide-protein docking

The clustering of the RK structure models leads to an ensemble of 13 structures. The receptor structure used in this study is the domain I of α 1 integrin (DI α 1). It consists of an inserted segment near the N-terminus found only in α 1, α 2 and α 12 integrin sequences. For instance, in contrast to α _v β ₃ the MIDAS (metal-ion-dependent adhesion site) is harbored in DI α 1 for the α 1 integrin by which it interacts with the extracellular collagen molecules. There are eleven structures in NMR ensemble of DI α 1 (PDB code: 2M32). To these, we added three other conformations resulting from the output of ENCoM server after submitting the crystal unbound structure of DI α 1 (PDB code: 1PT6). As a result, a total number of 14 receptor conformers were processed in this study. Consequently, we affected a number of 182 (13 for the ligand vs. 14 for the receptor) docking and a total of 364000 complexes were evaluated using the ZRANK scoring function. We end up with 14 conformers. At the final stage of the docking study we run molecular dynamics simulations for nine complexes of 10 ns. Only three

of them show a stable RMSD time evolution. The free energy of interaction, estimated by PRODIGY web server [45], are of -7.0, -7.3 and -8.4 kcal/mol. The docking solution with the lowest binding energy shows also a very stable profile of RMSD time evolution. In fact, the complex C α RMSD fluctuates around 2.5 Å (Fig. 7A) starting from the second ns to the end of the simulation. We propose this docking solution as a model of interaction between the RK peptide and (DI α 1). The interaction interface between the peptide and the receptor covers the largest part of the binding site of the collagen. The I₂ and D₂ residues interact with the C Loop of DI α 1. M₇ of RK seems to play an important role in the interaction. Its side chain occupies a small cavity situated away from the collagen interaction site. The cavity results from the 3D arrangement of two segments: 181-187 and 243-247 of the integrin (Notice that the amino acids numbering is based on the *Uniprot* accession P56199). W188 residues forms the interior of the cavity interacting with the methyl group of M₇.

4 Discussion

In this study, we isolated and characterized RK, from the *Buthus occitanus tunetanus* scorpion venom targeting the cell adhesion activity by acting on integrins. To our knowledge, this is the the first disintegrin-like peptide purified from a scorpion venom. Interestingly, RK peptide is unique among the variety of known peptides purified from scorpion venoms in terms of its structure. It is a small peptide of 17 amino acids (**IDCGTVMIPSECDPKSS**) containing a single disulfide bond.

Compared to our previously characterized peptide RK1 (**IDCSKVNLTAECS**) [27], RK has a sequence identity of 47%. The conserved residues are localized mainly at the C and N terminal ends including the disulfide bridge forming cysteins. With such similarity, it is likely that both peptides share a homology relationship. However, more evidences would be required to confirm this. RK may represent the first member of a new group of scorpion peptides. Interestingly RK did not show any toxicity a dose of 100 μ g per kg mice body weight. It is also non cytotoxic until 100 μ M concentration. RK cause the detachment of three cell lines on different ECM proteins. The ineffectiveness of the RK on the non specific substratum poly-L-lysine coated plates suggests that the inhibition effect of tumoral cells adhesion, by RK peptide, is integrin dependent. High levels of $\alpha_v\beta_3$, $\alpha_5\beta_1$ and $\alpha_3\beta_1$ are expressed in U87 [46], while IGR-39 highly expresses $\alpha_2\beta_1$ and $\alpha_v\beta_5$ integrins and modestly expresses $\alpha_v\beta_3$ [47]. The disintegrin activity of RK on U87 and IGR-39 would be linked to an effect on $\alpha_v\beta_3$ integrin. This could explain why a more important effect is exhibited by the peptide on U87 compared to IGR39. The lack of a significant activity on Fibronectin coated plates on IGR39 seems to be unlikely at first glance as we expected for the inhibition effect of RK to be similar on both Fibrinogen and Fibronectin like we did observe in U87. However these results might be explained by the fact that IGR39 expresses high levels of $\alpha_v\beta_5$ which can bind to Fibronectin though it is more specific to

vitronectin [48] and for such reason, the cell can still adhere to the extracellular matrix even in the presence of RK. A snake venom anti-adhesion effect of IGR39 was previously reported for lelectin, a C-type lectin. In contrast to RK, lelectin is only active on fibronectin and fibrinogen matrix types but not on collagen I and IV [49].

RK peptide disintegrin activity on U87 was evaluated at IC₅₀ values of 15 μM and 15.33 μM receptively for Fibrinogen and Fibronectin. The assay on PC12 cell line shows an anti adhesion activity on collagen IV coated plates, while no significant effect is noted for the rest of the ECM proteins. PC12 expresses α₁β₁ and α₃β₁ integrins, both are able to bind laminin *in vitro*, but only α₁β₁ can interact with type I and type IV collagen [50]. Therefore, in addition to α_vβ₃ we suggest in our study that RK acts on α₁β₁ integrin. As a comparison, the anti-adhesion to fibrinogen and fibronectin of PIVL (7691 Da), a snake venom disintegrin targeting α_vβ₃, on U87-cells was evaluated to similar IC₅₀ of about 250 nM and 300 nM respectively [51].

The similar IC₅₀ values might suggest also that RK inhibits the activity of a common target receptor for both Fibrinogen and Fibronectin which we suggest to be the α_vβ₃ integrin. This is supported by the previous calculation of the association rate constants of fibrinogen and fibronectin to α_vβ₃ integrin which are also very similar, evaluated at $2.49 \pm 0.8 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ and $2.63 \pm 0.05 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, respectively [52]. The antibody assay, confirms that both integrins are potential targets for the scorpion peptide. To our knowledge, RK is the first biomolecule demonstrating the dual activity on α₁β₁, which binds to collagen IV, and α_vβ₃ which binds to all ECM macromolecules presenting the RGD motif (vitronectin, fibronectin, fibrinogen, osteopontin, and bone sialoprotein) [53].

We though, that such property might be caused by the presence of one or two functional motifs described for the snake venom disintegrin. RK sequence contains two potential disintegrin motifs: ECD and KSS. Indeed, the ECD motif was previously described in Acurhagin-C to be responsible for the inhibition of the α_vβ₃ mediated endothelial cell adhesion to extracellular matrix [54].

Furthermore, the peptidomimetic of 12 amino acids, issued from the parent ECD loop carrier in Acurhagin-C purified from *Formosan Agkistrodon acutus* venom, showed similar disintegrin activity to RGD motif containing peptide [55]. ECD motif was also described in Alternagin-C, and Leberagin-C [56]. On the other hand, the KSS segment in RK could be is similar to KTS motif which is thought to be responsible for the α₁β₁ disintegrin activities for Obtustatin, Viperistatin and Lebestatin [57]. The importance of KTS motif was also corroborated in short linear peptide constructions [58]. To investigate the expected interaction of RK with α₁β₁ and α_vβ₃ integrins, we conducted computational studies. The molecular dynamics simulation of RK highlights the conformational flexibility of the peptide which allows it to adopt multiple stable states.

The free energy landscape shows a dense populated global minimum with a large bottom. Such characteristic allows a great adaptability toward the binding site of α₁β₁ integrin (**Fig. 7**). The latter

has the property to undergo a structural rearrangement upon the binding of collagen IV notably for the C loop as concluded by the crystal and NMR solution of the integrin [59-60-61-62]. The binding to $\alpha_1\beta_1$ could then be ensured favorably enough even if the peptide structure shifts significantly from the global minimum conformation which corresponds to a low drifts in the energy of the system. This property also, imply that the RK peptide must preserve a minimum specific set of contacts with high structural stability, in order to interact with the integrin.

The RMSF profile suggests that the central segment of the peptide, limited by the positions of the two cystein residues, are stable during the molecular dynamics simulation. Particularly, the segment V₆, M₇ and I₈ of the peptide might be critical to ensure the minimal contacts required for the interaction with $\alpha_1\beta_1$ regarding also the results from the docking (**Fig. 7**). The local conformation of the segment corresponding to a turn-like structure that remains stable during the 300 ns of molecular dynamics simulation. It allows to the M₇ side chain to preserve a high level of solvent exposure but also to minimize the structural constraints with other atoms of the peptide.

These properties are essential for the interaction with the small hydrophobic pocket on the surface of the integrin. The low flexibility in the 12-15 segment is also essential to stabilize the structure of K₁₅ residue. Such stability might allow the formation of the salt bridge with E₂₈₅ amino acid of DI α 1. The position of the S₁₆ and S₁₇ residues at the C-terminal end of the sequence causes a high instability. In addition, based on the protein-peptide docking study, none of the final retained solutions shows any critical implication of the KSS segment in the interaction with the integrin. Therefore, we suggest that the interaction of RK peptide with $\alpha_1\beta_1$ integrin does not implicate the KSS segment.

Our results are in accordance with those obtained with the snake venom disintegrin when mutagenesis study demonstrates that the substitution of KTS with AAA sequence in obtustatin, do not allow the total loose of its activity, compared to the wild type obtustatin [63], although this motif has been earlier proposed as the functional amino acid triplet for its activity [64]. These findings, suggest that the interaction mechanism of the snake venom disintegrin with $\alpha_1\beta_1$ might not rely only on the KTS triplet and that other critical residues might be involved in the formation of the complex. In the absence, of any interaction model of $\alpha_1\beta_1$ snake venom disintegrin, the mechanism is still to be investigated. The incubation of U87 with RGD peptide coated plates in the presence and the absence of RK suggests that both peptides share the same binding site on $\alpha_v\beta_3$. Consequently, we investigated the involvement of the ECD segment of RK in the interaction with the integrin. The rational behind this choice is the similarity with the same motif described in other snake disintegrin inhibiting the same receptor [65-55-58]. Our results showed that RK was able to sample a set of structures among the molecular dynamics ensemble for which the ECD segment has very similar geometric property compared to RGD in the co-crystal structure [66-67] (**Fig. 8A**). In addition, the conformations with the lowest clash score correspond also to ideal geometries of interactions with the $\alpha_v\beta_3$ integrin (**Fig. 8B**). This in fact suggests that RK is able

to engage the ECD segment in the interaction with RGD binding site without been hindered by the other residues of the same peptide. Therefore we suggest that, ECD might be the functional segment which controls the interaction with $\alpha_v\beta_3$ integrin (**Fig. 8C**). However, the conformations with very similar RGD geometry for the defined reactions coordinates in our computational analysis are not densely represented among the total ensemble sampled in the molecular dynamics simulation. This in fact, can explain the low inhibition activity of RK toward $\alpha_v\beta_3$ compared to $\alpha_1\beta_1$. To compare the structure-function relationship of RK peptide to our recently characterized short scorpion peptide RK1, we modeled the structure of the latter using the pep-Fold server [36].

To our surprise RK1 presents the exact same fold of RK which suggest that some functional features are shared between the two peptides (**supplementary data 2A**). Indeed, the sequence alignment revealed the presence several conserved segments, including the N-terminal IDC motif, V6, EC segment and the C-terminal SS segment. However, RK presents a DPK sequence which is absent in RK1 and several other divergent segments which harbors key amino acids in the interaction with the receptors. We generated a homology based complex of RK1 with $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins. Indeed, we found it very unlikely for the RK1 peptide to interact with alpha1 integrin in the same manner. In fact, RK1 lacks the key M7 amino acid substituted by a polar N7 residue which is incapable to establish the steric interactions with the exposed pocket of the integrin (**supplementary data 2B**). The importance of the hydrophobic interaction made with this small hydrophobic pocket was recently highlighted in Lebetin 2 which involves a Trp residue to stabilize the complex with alpha1 integrin [68]. Moreover, RK1 peptide lacks the K15 amino acids, which establishes a tight interaction with the receptor via a salt bridge formation in RK peptide. The homology based complex of RK1 with $\alpha_v\beta_3$ that the ECS segment would not replace the ECD motif in RK. The substitution of Asp by a Ser residue would cause the loss of coordinating the Mg^{++} ion in the MIDAS center (**supplementary data 2C**).

This study was the first to describe scorpion peptide with dual disintegrin activity on $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins. This double effect of RK is due to the presence of contiguous motifs ECD and KSS. It is to highlight that selective blockade of $\alpha_1\beta_1$ and $\alpha_v\beta_3$ integrins is a desirable goal for the therapy of a number of pathological conditions including essentially cancer and tumor angiogenesis.

Acknowledgements

We thank Pr Hechmi Louzir, head of Institut Pasteur de Tunis for his continued interest in this study and for his support.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This work was supported by grants from the Tunisian Ministry of Higher Education and Scientific Research (LR11IPT08) and Pasteur Institute of Tunisia (approval number: LNFP/Pro 152012).

LEGENDS TO FIGURES

Figure 1: Purification of RK from the venom *Buthus occitanus tunetanus*.

(A) Purification of BotG50 on semi preparative C8 reverse phase HPLC column. Fractions 19-22 min was collected. (B) Separation of fraction 19-22 min on C18-RP-HPLC, fraction eluted at 33.5 min was collected. (C) Final purification step to ensure the purity of peak eluting at 33.5 min (RK). (D) Amino acid sequence of RK.

Figure 2. Cell viability.

PC12 (A), IGR39 (B) and U87 (C) cells incubated for 72 h with different concentrations of RK. MTT solution was then added for 4 h. The MTT solution was removed and replaced with 100 μ L of DMSO into each well in order to dissolve the precipitated formazan crystals 100 μ l SDS 1%. Finally, the absorbance was measured at 560 nm. All data shown are mean (\pm SD) from 3 experiments (n=3) performed in triplicate at different times.

Figure 3. RK inhibits cell adhesion.

Glioblastoma cells U87 (A), Melanoma cells IGR39 (B) and Rat pheochromocytoma cells PC12 (C) Were then added to 96-well microtiter plates coated with different ECM proteins (type IV collagen, fibrinogen (Fg), fibronectin (Fn), laminin I (LnI)) or poly-L-lysine and allowed to adhere for 1 h at 37 °C. After washing, adherent cells were stained with crystal violet, solubilized by SDS and absorbance was measured at 560 nm. All data shown are mean (\pm SD) from 3 experiments (n=3) performed in triplicate at different times

Figure 4. RK inhibits integrin-mediated functions in tumour cells.

After washing, adherent cells were stained with crystal violet, solubilized by SDS and absorbance was measured at 560 nm. Glioblastoma cells U87 (A), Melanoma cells IGR39 (B) and Rat pheochromocytoma cells PC12 (C), were preincubated with 5, 10, 15, 20 μ M and 30 μ M of RK for 30 min at room temperature and then added to wells coated with 10 mg/ml fibronectin (Fn), 50 mg/ml fibrinogen (Fg) or 10 mg/ml type IV collagen, (Col IV) and allowed to attach for 1 h at 37 °C. Data shown are means (\pm SD) from 3 experiments (n=3) performed in triplicate at different times.

Figure 5. The effect of RK on various integrins in cell adhesion assays.

(A) Adhesion assays were performed with various cells/ECM protein pairs involving unique integrins: α 1 β 1 (PC12/type I collagen), α 2 β 1 (HT1080/type I collagen), α 6 β 4 (HT29-D4/laminin), α v β 6 (HT29-D4/fibronectin), α v β 3 (HT29 D4/fibrinogen), α v β 5 (HT29-D4/vitronectin) and α 5 β 1 (K562/fibronectin). Cells were preincubated without or with 15 μ M of RK for 30 min at room temperature. Cells were then added to 96-well microtiter plates coated with 5 mg/ml fibronectin, vitronectin or laminin-1, with 10 mg/ml type I collagens or with 50 mg/ml fibrinogen and allowed to adhere for 1.30 or 2 h at 37°C. After washing, adherent cells were stained with crystal violet, solubilized by SDS and absorbance was measured at 560 nm. (B) Cells incubated without or with 10 μ g/ml antibodies against α 1 β 1, α v β 3, α 5 β 1 or β 3, with 1 mM GRGDSP peptide (RGD peptide), all experiments were performed in triplicate. Data shown are means (\pm SD) from 3 experiments (n=3) performed in triplicate at different times.

Figure 6. Molecular modeling and molecular dynamics of RK peptide.

(A) We present the predicted model returned by the PEP-FOLD server. The C α atoms of each of the residues are indicated in spheres. The disulfide bond between the 3rd and the 12th amino acid is colored in red. (B) RMSD time evolution (upper panel) and RMSF profile (lower panel) of RK peptide calculating for a molecular dynamics trajectory of 300 ns. (C) Potential of mean forces of

the RK peptide calculated from the molecular dynamics simulation. The reaction coordinates consist of the radius of gyration (Rg) and the RMSD values relative to an average structure computed from the 300 ns simulation time.

Figure 7. Peptide-Protein docking of RK with $\alpha_1\beta_1$.

(A) RMSD time evolution of the best retained complex between RK and $\alpha_1\beta_1$ simulated for 10 ns. (B) Interaction of RK peptide (red) with the collagen IV binding site on the $\alpha_1\beta_1$. (C) The key interaction with the integrin is expressed in more details. The M₇ residue is capable to interact with a small hydrophobic pocket on the surface of the integrin formed by the residues in yellow. We also demonstrate the interaction of I₁ and I₂ with the C-loop and the salt bridge between K₁₅ and E₂₈₅.

Figure 8. Analysis of the interaction of RK with $\alpha_v\beta_3$.

(A) We fitted the atoms of the ECD segment from each snapshot the molecular dynamics trajectory of RK for those of the RGD coordinates of the 1L5G PDB structure. The ROSETTA clash score was then calculated. The reaction coordinates consist of the distance between E₁₁ CA and D₁₃ CG, and the angle between E₁₁ CA, C₁₂ CA and D₁₃ CG. (B) The conformation of RK with the least RMSD value is presented along with RGD peptide (yellow sticks). (C) The complex of RK with $\alpha_v\beta_3$ was constructed by fitting the conformation corresponding to the lowest clash score of the ECD segment with the RGD coordinates of the cilengetide from the PDB structure 1LG5.

Figure supplementary data2. Comparative *in silico* analysis of RK (IDCGTVMIPSECDPKSS) and RK1 (IDCCKVNLTAESS) peptides.

(A) Structural superposing between the two models of RK (cyan) and RK1 (green) peptides using PEP-FOLD2 method. Side chains of RK1 residues are shown. (B) We generated a homology based peptide-protein complex between RK1 and α_1 integrin. The procedure consists of using the structure of RK in complex with the integrin as a template and a model of RK1 was generated using MODELLER. We believe that this approach is rational since we demonstrated with molecular dynamics that the structure of RK peptide is highly flexible and therefore using the structure of RK1 generated by PEP-FOLD2 would not be a better choice. The homology based method allows also to normalize the comparison relative to the RK- α_1 complex as a reference state. We demonstrate in the figure how the N7 residue is not capable to optimize the steric contact with the nearby interaction pocket. Also by the absence of K15 residue, RK1 is not capable to establish a salt bridge complex with E285 residue of α_1 integrin. (C) Using the same method described in (B) we generated a complex between RK1 and $\alpha_v\beta_3$ integrin. RK and RK1 as well as their corresponding labels are colored in yellow and cyan respectively. The S13 residue in the ECS segment is likely not able to stabilize the interaction with the MIDAS site of the integrin. The position of the Calpha atom of S13 is shifted considerably from the equivalent position of D13 of RK peptide.

Supplementary material

- Figure: The experimental molecular mass of native RK (1780.03 Da) obtained by MALDI-Tof mass spectrometry. The profile corresponds to MALDI-TOF mass spectrometry analysis of the RP-HPLC peak containing RK.
- Figure: Comparative *in silico* analysis of RK (IDCGTVMIPSECDPKSS) and RK1 (IDCSKVNLTAECS) peptides.
- The retained complex between RK peptide and DI α 1 domain: "alpha1bet1-RK.pdb"
- Description of the interactions between RK and DI α 1 domain. This file can be generated with PyMOL, version 1.7.2.1 : "interaction_d111-RK.pse"
- Trajectory file for the molecular dynamics simulation of RK. The solvent molecules and the counterions were stripped from the file to decrease its size: md_300ns_nowater.xtc
- The gromacs coordinates file of RK used to read the xtc trajectory with VMD molecular viewer: "RK.gro"
- The structure from which we simulate the molecular dynamics trajectory: "RK_start.pdb"
- The structure from which we simulate the molecular dynamics trajectory: "RK_start.pse"
- A tab separated file, used to generate the figure 8A: "avb3_interaction_study.tsv"

Supplementary data 2:

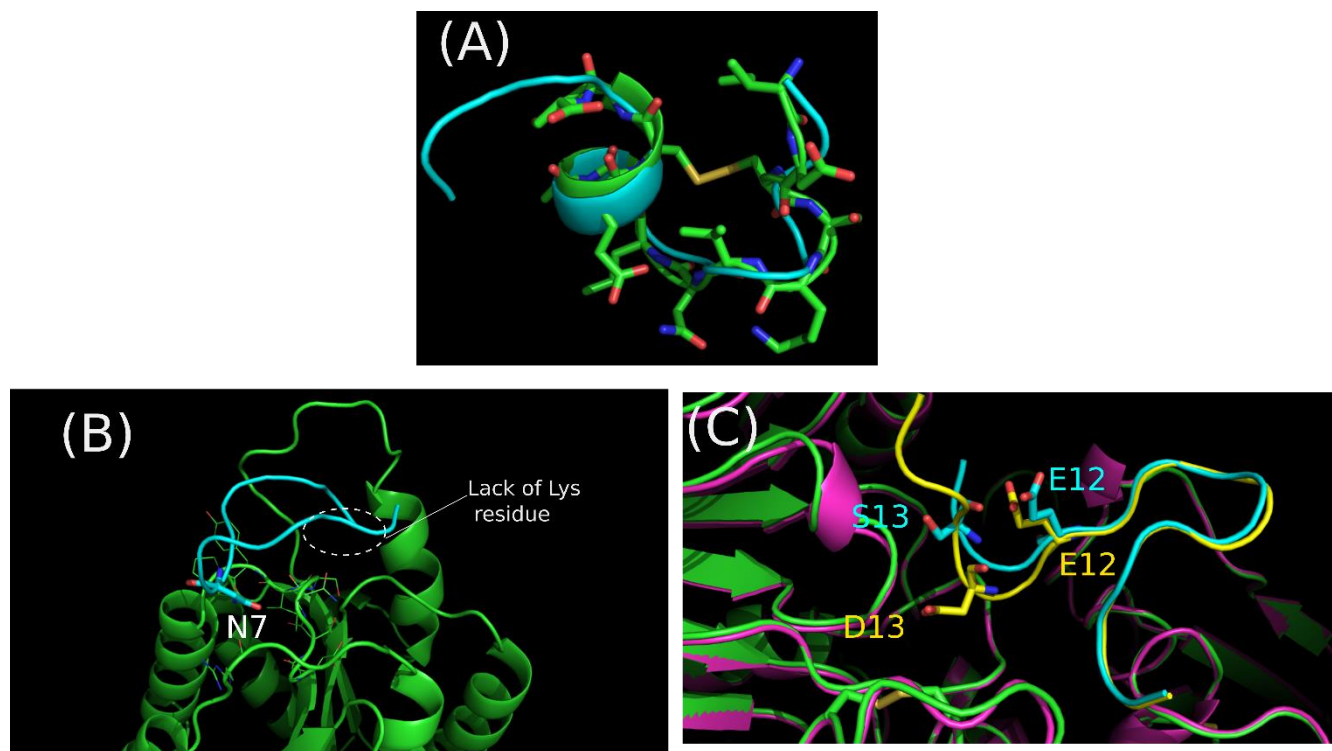


Figure supplementary data2. Comparative *in silico* analysis of RK (IDCGTVMIPSECDPKSS) and RK1 (IDCCKVNLTAESS) peptides.

(A) Structural superposing between the two models of RK (cyan) and RK1 (green) peptides using PEP-FOLD2 method. Side chains of RK1 residues are shown. (B) We generated a homology based peptide-protein complex between RK1 and α_1 integrin. The procedure consists of using the structure of RK in complex with the integrin as a template and a model of RK1 was generated using MODELLER. We believe that this approach is rational since we demonstrated with molecular dynamics that the structure of RK peptide is highly flexible and therefore using the structure of RK1 generated by PEP-FOLD2 would not be a better choice. The homology based method allows also to normalize the comparison relative to the RK- α_1 complex as a reference state. We demonstrate in the figure how the N7 residue is not capable to optimize the steric contact with the nearby interaction pocket. Also by the absence of K15 residue, RK1 is not capable to establish a salt bridge complex with E285 residue of α_1 integrin. (C) Using the same method described in (B) we generated a complex between RK1 and $\alpha_v\beta_3$ integrin. RK and RK1 as well as their corresponding labels are colored in yellow and cyan respectively. The S13 residue in the ECS segment is likely not able to stabilize the interaction with the MIDAS site of the integrin. The position of the Calpha atom of S13 is shifted considerably from the equivalent position of D13 of RK peptide.

```

RK          IDCGTVMIPSECDPKSS
RK1        IDCCKVNLTAESS---S
          ***..*  :.:**  *

```

Sequence alignment between RK and RK1 peptides

References:

- [1] Hanahan, D.; Weinberg, R.A. The Hallmarks of Cancer. *Cell* 2000, 100, 57–70.
- [2] Quan, J.; Yahata, T.; Adachi, S.; Yoshihara, K.; Tanaka, K. Identification of receptor tyrosine kinase, discoidin domain receptor 1 (DDR1), as a potential biomarker for serous ovarian cancer. *International Journal of Molecular Sciences*. 2011, 12, 971–982. doi: 10.3390/ijms12020971.
- [3] Guo, W.; Giancotti, F.G. Integrin signalling during tumour progression. *Nat Rev Mol. Cell Biol.* 2004, 5, 816–826.
- [4] Shepparda, D. In vivo functions of integrins: lessons from null mutations in mice Dean. *Matrix Biol.* 2000, 19, 203-209.
- [5] Giancotti, F.G. Integrin signaling: specificity and control of cell survival and cell cycle progression. *Curr Opin Cell Biol.* 1997, 9, 691–700.
- [6] Liddington, R.C.; Ginsberg M.H. Integrin activation takes shape. *J Cell Biol.* 2002, 158, 833–839.
- [7] Gahmberg, C.G.; Fagerholm, S.C.; Nurmi, S.M.; Chavakis, T.; Marchesan, S.; Grönholm, M. Regulation of integrin activity and signalling. *Biochimica et Biophysica Acta.* 2009, 6, 431–444. doi: 10.1016/j.bbagen.2009.03.007.
- [8] Van der Flier, A.; Sonnenberg, A. Function and interactions of integrins. *Cell Tissue Research.* 2001, 3, 285-298.
- [9] Hynes, R.O. Integrins: bidirectional, allosteric signaling machines. *Cell.* 2000, 110, 673–687.
- [10] Hynes, R.O. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell.* 1992, 69, 11–25.
- [11] Juarez, P.; Comas, I.; Gonzalez-Candelas, F.; Calvete, J.J. Evolution of snake venom disintegrins by positive Darwinian selection. *Mol Biol Evol.* 2008, 25, 2391-407. doi: 10.1093/molbev/msn179.
- [12] McLane, M.A.; Marcinkiewicz, C.; Vijay-Kumar, S.; Wierzbicka-Patynowski, I.; Niewiarowski, S. Viper venom disintegrins and related molecules. *Proc. Soc. Exp. Biol. Med.* 1998, 219, 109–119.
- [13] Springer, T.A.; Wang, J-h. The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion. *Adv Protein Chem.* 2004, 68, 29–63.
- [14] Lee, J.O.; Bankston, L.A.; Arnaout, M.A.; Liddington, R.C. Two conformations of the integrin A-domain (I-domain): A pathway for activation. *Structure.* 1995, 3, 1333–1340.
- [15] Heino, J. The collagen receptor integrins have distinct ligand recognition and signaling functions. *Matrix Biol.* 2000, 19, 319-23.
- [16] Shi, M.; Pedchenko, V.; Greer, B.H.; Van Horn, W.D.; Santoro, S.A.; Sanders, C.R.; Hudson, B.G.; Eichman, B.F.; Zent, R., Pozzi, A. Enhancing integrin $\alpha 1$ inserted (I) domain affinity to ligand potentiates integrin $\alpha 1\beta 1$ -mediated down-regulation of collagen synthesis. *J Biol Chem.* 2012, 287, 35139–35152. doi: 10.1074/jbc.M112.358648.
- [17] Renner, C.; Saccà, B.; Moroder, L. Synthetic heterotrimeric collagen peptides as mimics of cell adhesion sites of the basement membrane. *Biopolymers.* 2004, 76, 34-47.
- [18] Arruda Macedo, J.K.; Fox, J.W.; de Souza Castro, M. Disintegrins from snake venoms and their applications in cancer research and therapy. *Curr Protein Pept Sci.* 2015, 16, 532-48.
- [19] Marcinkiewicz, C. Functional characteristic of snake venom disintegrins: Potential therapeutic implication. *Current Pharmaceutical Design.* 2005, 11, 815–27.
- [20] Calvete, J.J.; Marcinkiewicz, C.; Monleón, D.; Esteve, V.; Celda, B.; Juárez, P.; Sanz, L. Snake venom disintegrins: evolution of structure and function. *Toxicon.* 2005, 45, 1063-1074.
- [21] Debin, J.A.; Maggio, J.E.; Strichartz, G.R. Purification and characterization of chlorotoxin, a chloride channel ligand from the venom of the scorpion. *Am J Physiol.* 1993, 264, 361-9.

- [22] Dardevet, L.; Rani, D.; Aziz, T.A.; Bazin, I.; Sabatier, J.M.; Fadl, M.; Brambilla, E.; De Waard, M. Chlorotoxin: a helpful natural scorpion peptide to diagnose glioma and fight tumor invasion. *Toxins (Basel)*. 2015, 7, 1079–101. doi: 10.3390/toxins7041079.
- [23] He, Q.Y.; He, Q.Z.; Deng, X.C.; Yao, L.; Meng, E.; Liu, Z.H.; Liang, S.P. ATDB: a uni-database platform for animal toxins. *Nucleic Acids Res.* 2008, 36 (Database issue), D293–D297. doi: 10.1093/nar/gkm832.
- [24] Harvey, A.L. Toxins and drug discovery. *Toxicon*. 2014, 92, 193–200. doi: 10.1016/j.toxicon.2014.10.020.
- [25] Vita, C.; Roumestand, C.; Toma, F.; Ménez, A. Scorpion toxins as natural scaffolds for protein engineering. *Biochemistry*. 1995, 92, 6404–6408.
- [26] Zeng, X.C.; Corzo, G.; Harin, R. Scorpion venom peptides without disulfide bridges. *IUBMB Life*. 2005, 57, 13–21.
- [27] Khammessi, O.; Ben Mabrouk, H.; ElFessi-Magouri, R., Kharrat.; R. RK1, The first very short peptide from *Buthus occitanus tunetanus* inhibits tumor cell migration, proliferation and angiogenesis. *Biochem Biophys Res Commun*. 2018 Jan 20. pii: S0006-291X (18)30148-7. doi: 10.1016/j.bbrc.2018.01.133.
- [28] Mahjoubi-Boubaker, B.; Crest, M.; Ben Khalifa, R.; El Ayeb, M.; Kharrat, R. Kbot1, a three disulfide bridges toxin from *Buthus occitanus tunetanus* venom highly active on both SK and Kv channels. *Peptides*. 2004, 25, 637–45.
- [29] ElFessi-Magouri, R.; Peigneur, S.; Khamessi, O.; Srairi-Abid, N.; El ayeb, M.; Mille, B.G.; Cuypers, E.; Tytgat, J.; Kharrat, R. Kbot55, purified from *Buthus occitanus tunetanus* venom, represents the first member of a novel α -KTx subfamily. *Peptides*. 2016, 80, 4–8. doi: 10.1016/j.peptides.2015.05.015.
- [30] ElFessi-Magouri, R.; Peigneur, S.; Khamessi, O.; Srairi-Abid, N.; El ayeb, M.; Mille, B.G.; Cuypers, E.; Tytgat, J.; Kharrat, R. Kbot55, purified from *Buthus occitanus tunetanus* venom, represents the first member of a novel α -KTx subfamily. *Peptides*. 2016, 80, 4–8. doi: 10.1016/j.peptides.2015.05.015.
- [31] Galeotti, N.; Bartolini, A.; Ghelardini, C. Diphenhydramine-induced amnesia is mediated by Gi-protein activation. *Neuroscience*. 2003, 122, 471–478.
- [32] Yatogho, T.; Izumi, M.; Kashiwagi, H.; Hayashi, M. Novel purification of vitronectin from human plasma by heparin affinity chromatography. *Cell. Struct. Funct.* 1988, 13, 281–292.
- [33] Mosmann, T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods*. 1983, 65, 55–63.
- [34] Kadi, A.; Pichard, V.; Lehmann, M.; Briand, C.; Braguer, D.; Marvaldi, J.; Rognoni, J.B.; Luis, J. Effect of microtubule disruption on cell adhesion and spreading. *Biochem. Biophys. Res. Commun.* 1998, 246, 690–695.
- [35] Defilles, C.; Lissitzky, J.C.; Montero, M.P.; Andre, F.; Prevot, C.; Delamarre, E.; Marrakchi, N.; Luis, J.; Rigot, V. α 5 β 1 integrin suppression leads to a stimulation of α 2 β 1 dependent cell migration resistant to PI3K/Akt inhibition. *Exp. Cell Res.* 2009, 315, 1840–1849.
- [36] Thévenet, P.; Shen, Y.; Maupetit, J.; Guyon, F.; Derreumaux, P.; Tufféry, P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic. Acids. Res.* 2012, 40, W288–293. doi: 10.1093/nar/gks419.
- [37] Seeliger, D.; de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des.* 2010, 24, 417–22. doi: 10.1007/s10822-010-9352-6.
- [38] Luo, B-H.; Carman, C.V.; Springer, T.A. Structural basis of integrin regulation and signaling. *Annu Rev Immunol.* 2007, 25, 619–647.
- [39] Chin, Y.K.; Headey, S.J.; Mohanty, B.; Patil, R.; McEwan, P.A.; Swarbrick, J.D.; Mulhem, T.D.; Emsley, J.; Simpson, J.S.; Scanlon, M.J. The structure of integrin α II domain in complex with a collagen-mimetic peptide. *J Biol Chem.* 2013, 288, 36796–809. doi: 10.1074/jbc.M113.480251.

- [40] Frappier, V.; Chartier, M.; Najmanovich, R.J. ENCoM server: exploring protein conformational space and the effect of mutations on protein function and stability. *Nucleic Acids Res.* 2015, 43, 395-400. doi: 10.1093/nar/gkv343.
- [41] Chen, R.; Li, L.; Weng, Z. ZDOCK: an initial-stage protein-docking algorithm. *Proteins.* 2003, 52, 80-7. doi:10.1002/prot.10389
- [42] Pierce, B.; Weng, Z. ZRANK: reranking protein docking predictions with an optimized energy function. *Proteins.* 2007, 67, 1078-86. doi: 10.1002/prot.21373
- [43] Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R., Rousseau, F.; Serrano, L. The FoldX web server: an online force field. *Nucleic Acids Res.* 2005, 33(Web Server issue), W382-8. Doi: 10.1093/nar/gki387.
- [44] Oostenbrink, C.; Villa, A.; Mark, A.E.; Van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem.* 2004, 25, 1656-76. doi: 10.1002/jcc.20090
- [45] Xue, L.C.; Rodrigues, J.P.; Kastriitis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics.* 2016, 32, 3676-3678. doi:10.1093/bioinformatics/btw514.
- [46] Malric, L.; Monferran, S.; Gilhodes, J.; Boyrie, S.; Dahan, P.; Skuli, N.; Sesen, J.; Filleron, T.; Kowalski-Chauvel, A.; Cohen-Jonathan Moyal, E.; Toulas, C.; Lemarié, A. Interest of integrins targeting in glioblastoma according to tumor heterogeneity and cancer stem cell paradigm: an update. *Oncotarget.* 2017, 8:86947-86968. doi.org/10.8632/oncotarget.20372
- [47] Siret, C.; Terciolo, C.; Dobric, A.; Habib, M.C.; germain, S.; Bonnier, R.; Lombardo, D.; Rigot, V.; André, F. Interplay between cadherins and $\alpha 2\beta 1$ integrin differentially regulates melanoma cell invasion. *Br J Cancer.* 2015, 113, 1445-53.
- [48] Eva-maria,F.;Gustav-Paul,W.; Ute, K.B.; Bodo, H.; Simon, L.G.; Jan-Dirk, R. Immunohistochemical analysis of integrins $\alpha v\beta 3$, $\alpha v\beta 5$ and $\alpha 5\beta 1$, and their ligands, fibrinogen, fibronectin, osteopontin and vitronectin, in frozen sections of human oral head and neck squamous cell carcinomas. *Exp Ther Med.* 2011, 2, 9–19.
- [49] Sarray, S.; Berthet, V.; Calvete, J.J.; Secchi, J.; Marvaldi, J.; El-Ayeb, M.; Marrakchi, N.; Luis, J. Lebectin, a novel C-type lectin from *Macrovipera lebetina* venom, inhibits integrin-mediated adhesion, migration and invasion of human tumour cells. *Lab. Invest.* 2004, 84, 573–581.
- [50] Tomaselli, K.J.; Hall, D.E.; Flier, L.A.; Gehlsen, K.R.; Turner, D.C.; Carbonetto, S.; Reichardt, L.F. A neuronal cell line (PC12) expresses two beta 1-class integrins-alpha 1 beta 1 and alpha 3 beta 1-that recognize different neurite outgrowth-promoting domains in laminin. *Neuron.* 1990, 5, 651-62.
- [51] Morjen, M.; Kallech-Ziri, O.; Bazaa, A.; Othman, H.; Mabrouk, K.; Zouari-Kessentini, R.; Sanz, L.; Calvete, J.J.; Srairi-Abid, N.; El Ayeb, M.; Luis, J.; Marrakchi, N. PIVL, a new serine protease inhibitor from *Macrovipera lebetina* transmediterranea venom, impairs motility of human glioblastoma cells. *Matrix Biol.* 2013, 32, 52-62. doi: 10.1016/j.matbio.2012.11.015.
- [52] Smith, J.W., Piotrowicz, R.S., Mathis, D. A mechanism for divalent cation regulation of beta 3-integrins. *J Biol Chem.* 1994, 269, 960-7.
- [53] Arnaout, M.A.; Mahalingam, B.; Xiong, J.P. Integrin structure, allostery, and bidirectional signaling. *Annu Rev Cell Dev Biol.* 2005, 21, 381-410.
- [54] Wang, W.J. Acurhagin-C, an ECD disintegrin, inhibits integrin $\alpha v\beta 3$ -mediated human endothelial cell functions by inducing apoptosis via caspase-3 activation. *Br J Pharmacol.* 2010, 160, 1338-51.
- [55] Selistre-de-Araujo, H.S.; Pontes, C.L.; Montenegro, C.F.; Martin, A.C. Snake venom disintegrins and cell migration. *Toxins (Basel).* 2010, 2, 2606-21.
- [56] Limam, I.; Bazaa, A.; Srairi-Abid, N.; Taboubi, S.; Jebali, J.; Zouari-Kessentini, R.; Kallech-Ziri, O.; Mejdoub, H.; Hammami, A.; El Ayeb, M.; Luis, J.; Marrakchi, N. Leberagin-C, A

disintegrin-like/cysteine-rich protein from *Macrovipera lebetina transmediterranea* venom, inhibits alphavbeta3 integrin-mediated cell adhesion. *Matrix Biol.* 2010, 29, 117–126.

[57] Olfa, K.Z.; Jose, L.; Salma, D.; Amine, B.; Najet, S.A.; Nicolas, A.; Maxime, L.; Raoudha, Z.; Kamel, M.; Jacques, M., Jean-Marc, S.; Mohamed el, A.; Naziha, M. Lebestatin, a disintegrin from *Macrovipera* venom, inhibits integrin-mediated cell adhesion, migration and angiogenesis. *Lab. Invest.* 2005, 85, 1507–1516. doi: 10.1038/labinvest.3700350.

[58] Kisiel, D.G.; Calvete, J.J.; Katzhendler, J.; Fertala, A.; Lazarovici, P.; Marcinkiewicz, C. Structural determinants of the selectivity of KTS-disintegrins for the alpha1beta1 integrin. *FEBS Lett.* 2004, 577, 478-82. doi: 10.1016/j.febslet.2004.10.050.

[59] Nolte, M.; Pepinsky, R.B.; Venyaminov, S.Y.; Kotliansky, V.; Gotwals, P.J.; Karpusas, M. Crystal structure of the $\alpha 1\beta 1$ integrin I-domain: insights into integrin I-domain function. *FEBS Lett.* 1999, 452, 379-385.

[60] Rich, R.L.; Deivanayagam, C.C.; Owens, R.T.; Carson, M.; Hook, A.; Moore, D.; Symersky, J.; Yang, V.W.; Narayana, S.V.; Hook, M. Trench- shaped binding sites promote multiple classes of interactions between collagen and the adherence receptors, $\alpha 1\beta 1$ integrin and *Staphylococcus aureus* cna MSCRAMM. *J Biol Chem.* 1999, 274, 24906-24913.

[61] Salminen, T.A.; Nymalm, Y.; Kankare, J.; Kapyla, J.; Heino, J.; Johnson, M.S. Production, crystallization and preliminary X-ray analysis of the human integrin $\alpha 1$ I domain. *Acta Crystallogr. D Biol. Crystallogr.* 1999, 55, 1365-1367.

[62] Nymalm, Y.; Puranen, J.S.; Nyholm, T.K.; Kapyla, J.; Kidron, H.; Pentikainen, O. T.; Airene, T.T.; Heino, J., Slotte, J.P.; Johnson, M.S.; Salminen, T.A. Jararhagin-derived RKKH peptides induce structural changes in $\alpha 1$ I domain of human integrin $\alpha 1\beta 1$. *J. Biol. Chem.* 2004, 279, 7962-7970.

[63] Brown, M.C.; Eble, J.A.; Calvete, J.J.; Marcinkiewicz. Structural requirements of KTS-disintegrins for inhibition of alpha(1)beta(1) integrin. *Biochem J.* 2009, 417, 95-101. doi: 10.1042/BJ20081403.

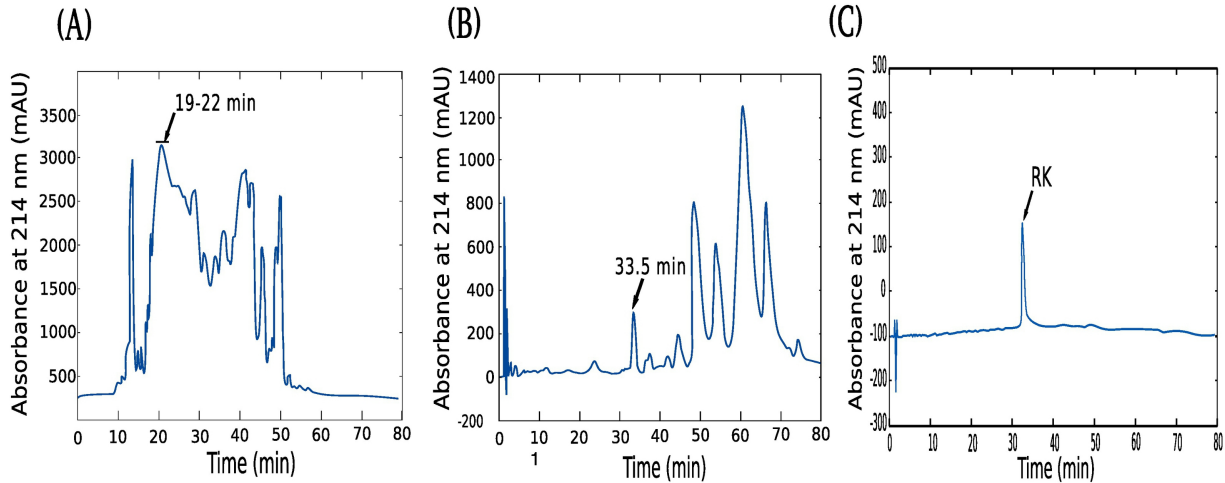
[64] Moreno-Murciano, P.; Daniel Monleon, D.; Calvete, J.J.; Celda, B.; Marcinkiewicz, C. Amino acid sequence and homology modeling of obtustatin, a novel non-RGD-containing short disintegrin isolated from the venom of *Vipera lebetina obtusa*. *Protein Sci.* 2003, 12, 366–371. doi: 10.1110./ps.0230203.

[65] Shih, C.H.; Chiang, T.B.; Wang, W.J. Inhibition of integrins $\alpha v/\alpha 5$ -dependent functions in melanoma cells by an ECD-disintegrin acurhagin-C. *Matrix Biol.* 2013, 32, 152-159.

[66] Xiong, J.P.; Stehle, T.; Diefenbach, B.; Zhang, R.; Dunker, R.; Scott, D.L.; Joachimiak, A.; Goodman, S.L.; Arnaout, M.A. Crystal structure of the extracellular segment of integrin $\alpha V\beta 3$. *Science.* 2001, 294, 339–45. doi 10.1126/science1064535

[67] Xiong, J.P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S.L.; Arnaout, M.A. Crystal structure of the extracellular segment of integrin $\alpha V\beta 3$ in complex with an Arg-Gly-Asp ligand. *Science (NY).* 2002, 296, 151–155.

[68] Morjen, M.; Othman, H.; Abdelkafi-Koubaa Z.; Messad, E.; Jebali J, El Ayeb M, Abid NS, Luis J, Marrakchi N. Targeting $\alpha 1$ inserted domain (I) of $\alpha 1\beta 1$ integrin by Lebetin 2 from *M. lebetina transmediterranea* venom decreased tumorigenesis and angiogenesis. *Int J Biol Macromol.* 2018 Oct 1;117:790-799. doi: 10.1016/j.ijbiomac.2018.05.230. Epub 2018 Jun 2.



(D) **IDCGTVMIPSECDPKSS**

Figure 1

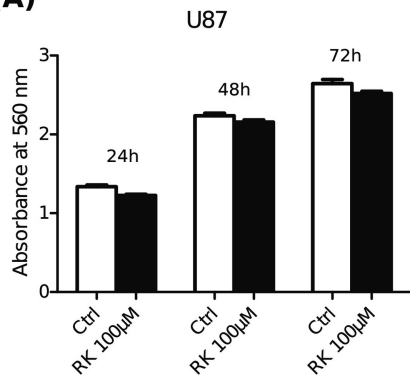
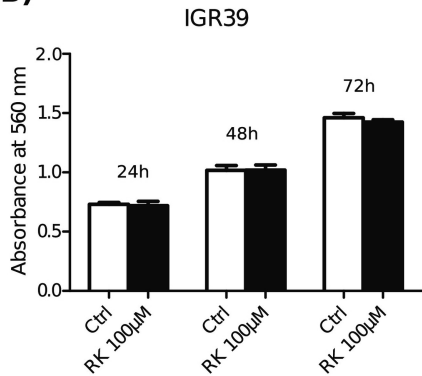
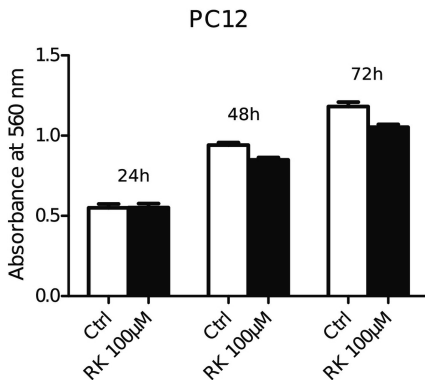
(A)**(B)****(C)**

Figure 2

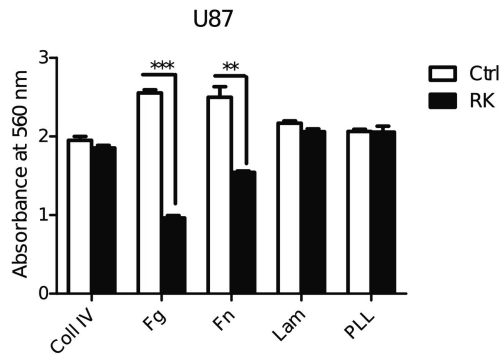
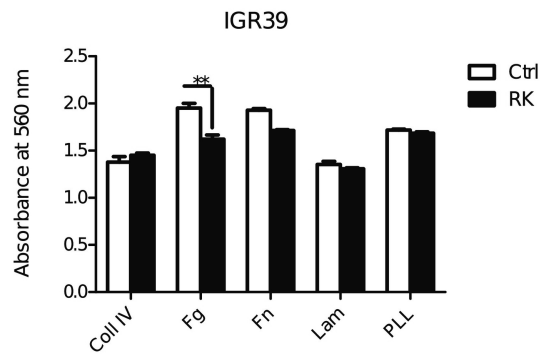
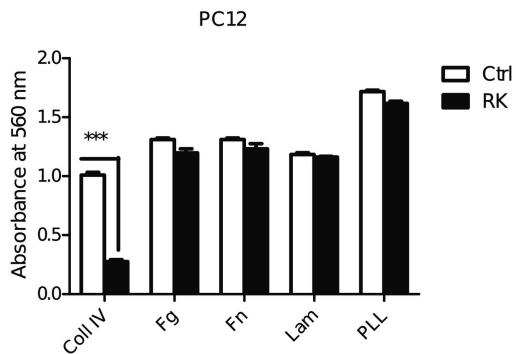
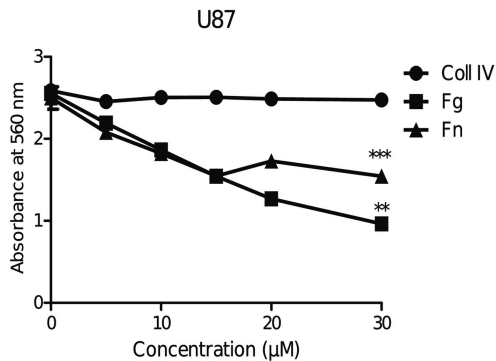
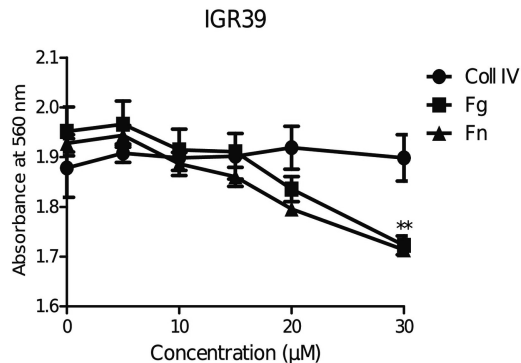
(A)**(B)****(C)**

Figure 3

(A)



(B)



(C)

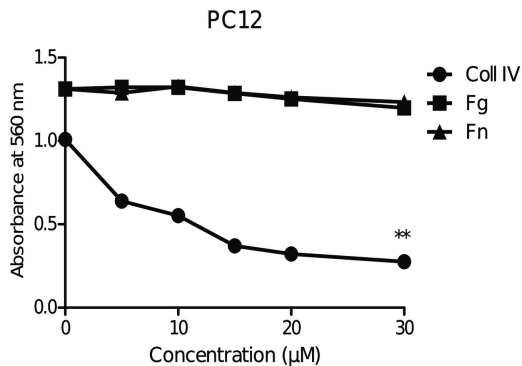


Figure 4

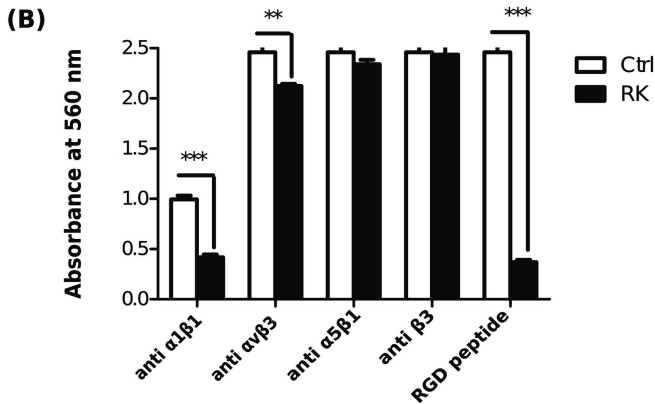
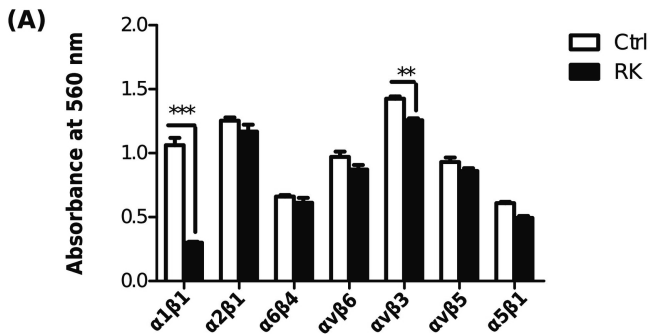


Figure 5

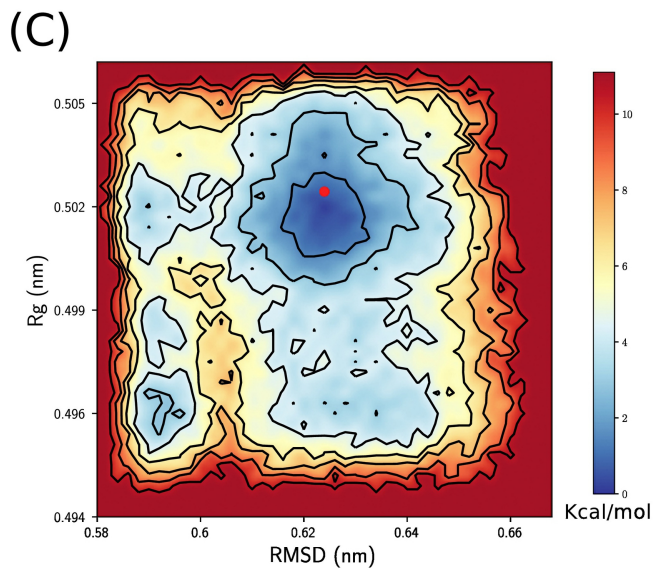
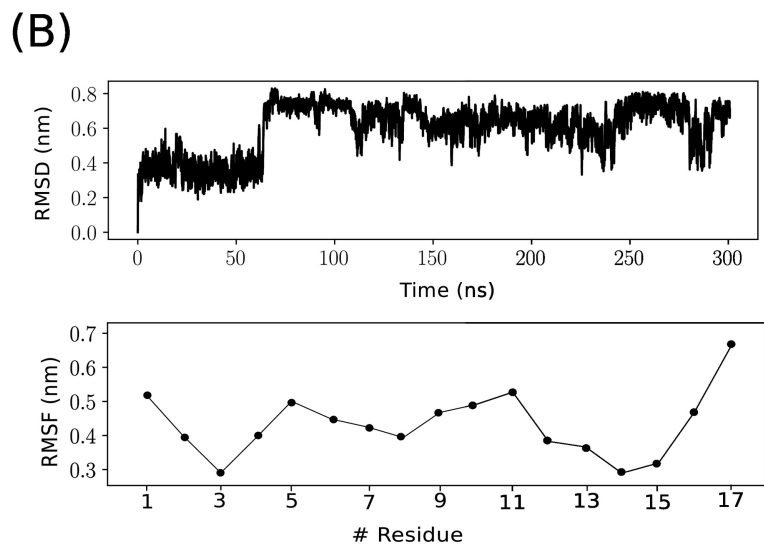
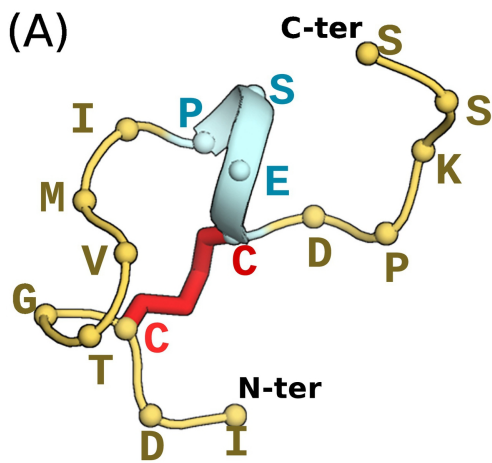


Figure 6

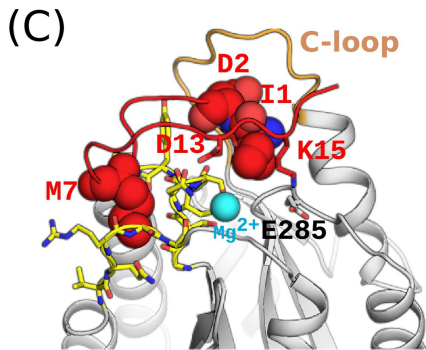
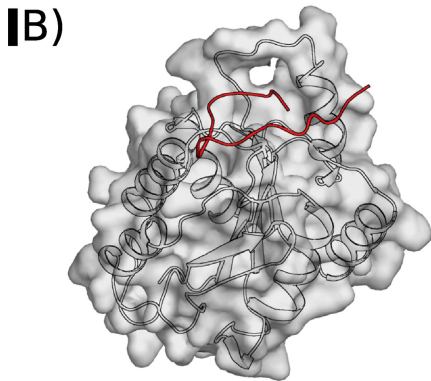
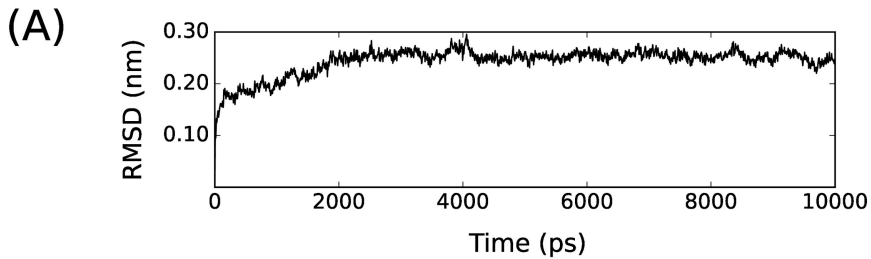


Figure 7

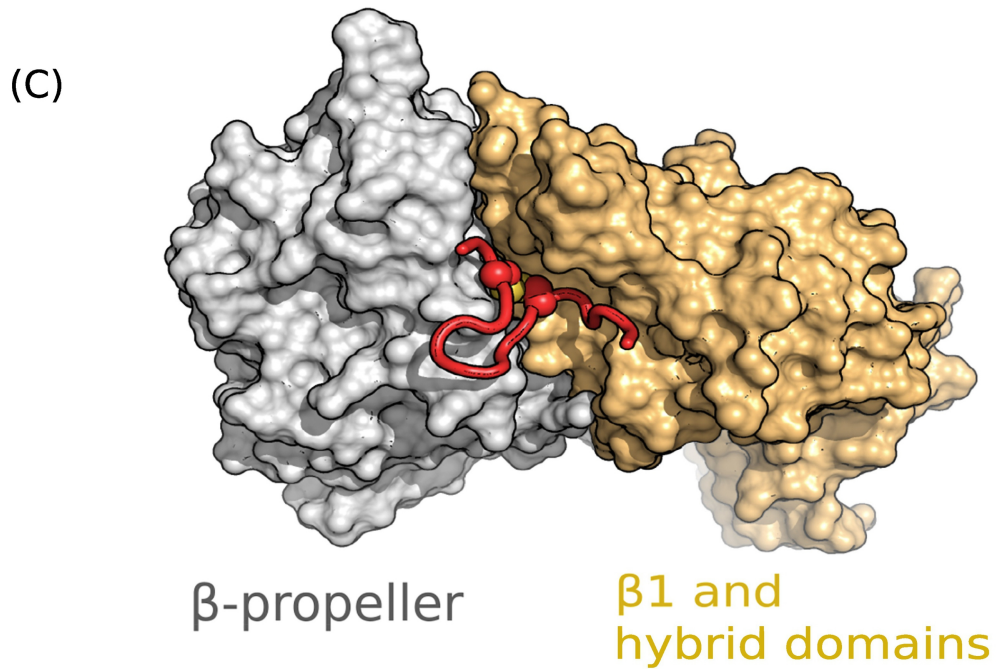
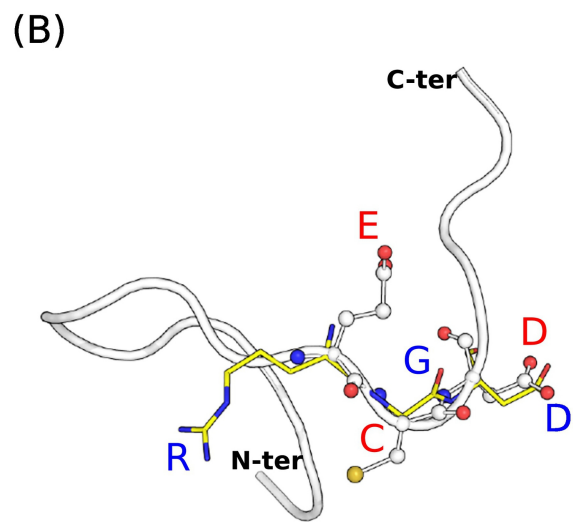
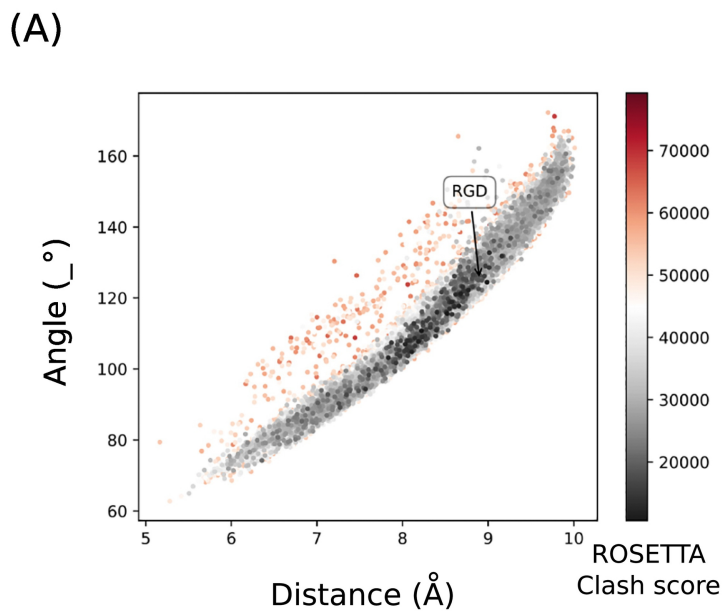


Figure 8