



- (51) **International Patent Classification:**
C07K 7/06 (2006.01)
- (21) **International Application Number:**
PCT/EP2015/054397
- (22) **International Filing Date:**
3 March 2015 (03.03.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/947,292 3 March 2014 (03.03.2014) US
- (63) **Related by continuation (CON) or continuation-in-part (CIP) to earlier application:**
US 61/947,292 (CON)
Filed on 3 March 2014 (03.03.2014)
- (71) **Applicants:** **VRIJE UNIVERSITEIT BRUSSEL (VUB)** [BE/BE]; Pleinlaan 2, B-1050 BRUSSEL (BE). **BELGIAN MINISTRY OF DEFENSE** [BE/BE]; Rue d'Evere 1, B-1140 Bruxelles (BE).
- (72) **Inventors:** **ZIZI, Martin**; Rue du Bois Colau 1, B-1350 Enines (BE). **FIGYS, Jurgen**; Brusselsesteenweg 340, B-2800 Mechelen (BE).
- (74) **Agent:** **LC PATENTS**; Crutzenstraat 24, B-3500 Hasselt (BE).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to the identity of the inventor (Rule 4.17(i))
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) **Title:** SEQUAMERS - A NOVEL NON-NATURAL TYPE OF PEPTIDE LIBRARIES RANDOMLY CONSTRAINED BY THEIR PRIMARY SEQUENCES

(57) **Abstract:** The present application belongs to the field of functional peptides and more particularly to the field of non-naturally occurring random peptides that are constrained by their primary sequences. The invention discloses molecules of a peptide structure as defined in the claims and methods of using such molecules for therapeutic applications and for diagnostic uses, as well as in other applications such as in the agbio field and in industrial biotechnology. The present invention also provides random peptide libraries of non-naturally occurring peptides.



**SEQUAMERS – A NOVEL NON-NATURAL TYPE OF PEPTIDE LIBRARIES RANDOMLY
CONSTRAINED BY THEIR PRIMARY SEQUENCES**

FIELD OF THE INVENTION

5 The present invention is in the field of molecular biology and is related to non-naturally occurring random peptides that are constrained by their primary sequences. The invention discloses molecules of a peptide structure and methods of using such molecules for therapeutic applications and for diagnostic uses, as well as in other applications such as in the agbio field and in industrial biotechnology. The present invention also provides random peptide
10 libraries of non-naturally occurring peptides.

BACKGROUND TO THE INVENTION

High throughput screenings consist in the directed exploration of repertoires. Various methods exist, making use of several types of libraries/repertoires. Those include classical display
15 methods (such as 2-hybrid, phage display) or improved methods such as ribosome display. Regardless of the methodologies used, the key to the successful screening of random libraries requires the need to have the largest sequence space available to explore and harness the power of randomness and in addition to have some structural shape (rigidity, stiffness, stability, scaffolding potential) to be able to compete with natural binders/competitors or in vivo
20 conditions. Most randomized repertoires (from natural or synthetic origin, concatenated or not,...) have shown promising results, but one has to conclude that the chemical variation (sequence diversity) contains mostly sequences without structural constraints and that are hence a priori not useful as binders. There is thus a need to build libraries/repertoires that are more relevant to binding by focusing not on the sequence space but on the potential shape
25 space or space-filling space. Short peptide sequences can be highly effective at binding a receptor, like some neurotransmitters or peptides hormones (such as oxytocin, vasopressin, angiotensin, pentagastrin, substance P), which have obviously solved said structural-selectivity trade –off. There are already phage display-derived peptides that are available as drugs on the market¹⁵, but their numbers remain limited. Most of them are also derived from natural
30 sequences (see ¹⁶ for review). Because to date, short peptides (obtained by screening) even when possessing high in vitro affinities (very low Kd's) perform poorly when tested against their molecular target in the presence of other molecules or competitors, and their performance can even be worse in vivo. This is presumably due to the fact that their secondary and/or tertiary structures are not stable enough and may change too much depending on the
35 conditions in complex solutions like the cell and all body fluids are. Generally speaking peptides can thus be defined as linear – i.e. without stable or unequivocal structure – or constrained – i.e. where some stable functional structures are imposed on the peptide. These latter being considered more effective in vivo binders as the flexibility of linear peptides ought to lower their affinity towards the target (for a general discussion see also ¹⁷ and ¹⁸). There are

-2-

some studies however reporting that sometimes linear peptides display the highest affinities and activities ^{19,20}, these studies however compare *identical* sequences being either constrained by cyclization or their linear equivalent. Such comparison between linear and cyclized peptides having same length and same sequence is not fully valid as i) one would expect that - by cyclization - some key residue may end up not being correctly positioned because the sequences were initially obtained for the linear peptide, and ii) in some studies it is the constrained peptide that is more effective ²¹. There exist numerous strategies or manners to constrain peptides but all strategies have their inherent deficiencies. One strategy would alter the conformational properties of the peptides without any modification like in the case of cationic peptides that are being considered as possible antibiotics. In this case, it is their amphiphilic character and the presence of some key charged residues that are relevant as their effects on bacteria are mostly biophysical, indeed most of those peptides (magainins, cecropins, and perforins) act by creating pores into- or disturbing the structures of- the bacterial membranes. Other strategies involve modifying the peptide. One such strategy involves cyclizing the peptides, cyclization is a rather old approach ²², and by reducing the conformational freedom of such molecules, it could increase affinity by decreasing unfavorable entropic effects. It is worth mentioning that oxytocin and vasopressin are cyclized short peptides. It lead to early successes in obtaining hormones or neurotransmitters analogues like enkephalin, somatostatin a.o. ^{22,23}, but did not impose itself as a generic tool to obtain high affinity peptide effective in vivo. Several ways are used to obtain cyclic peptide: closure of the peptide by a peptidic bond like in macrocyclic compounds ²⁴, a lactam bridging method ²⁵, the closure of loop(s) by disulfide bridging like in ²⁶ and ²⁷ or by various crosslinkers. Numerous studies have presented results demonstrating useful affinities for cyclic-constrained peptides around nanomolar levels for macrocyclic inhibitors of HIV-1 protease ²⁸, for the hepatitis B surface and core antigens ^{29,30}, for the CK2 protein kinase ³¹ or micromolar for the Src SH2 domain ³² and for a cardiac Na⁺/Ca²⁺ exchanger ³³. Sometimes cyclization - combined with other constraints like in the case of disulfide bridging or the use of either the L-form or the D-form of the cysteine - led to an additional constraint in the form of a right handed or left handed double helix ^{34,35}. The combination of cyclization and side chain constraints was also used ^{27,36}. The screening of cyclic peptides was also proposed as a first screening step for raw screening, prior a refined screening of linear peptides based on the cyclic ones and opened at various positions ¹⁹.

In addition, the use of helix favoring amino-acid like α - amino isobutyric acid (Aib) or alanine like in ³⁷ is also reported. Likewise helicomimetics – based on the NR box, a structural element derived from steroid nuclear receptors – may constrain peptides within a LXXLL motif ³⁸. In this context, of using mimetics, some efforts were also made to control the orientation of the side chains of the 2 AA's - ASP and ARG - to understand (via the analysis of a pseudo-dihedral angle) the selectivity of the RGD motif interactions between integrins ^{39, 40,41}. This tripeptide

(Arg, Gly, Asp) can be used as a circular 'pharmacophore' – a presenting vehicle. Linear versions were not pursued. Those molecules are in fact peptidomimetics and, not pure peptides *per se* ⁴⁰. Other molecules like cyclopropane, a.o. were used to obtain peptidomimetics ⁴². The affinities obtained with such peptidomimetics strategies were in the micromolar range with some exception in the nanomolar range like a back-to-back α -helix mimetic against a virus ⁴³. Disulfide bridges are known to be essential for structure stability and biological activities in numerous proteins or enzymes. Numerous authors applied this strategy. This strategy was also used either with intra-chain or inter-chain bridging. Intra-chain bridging can be used for cyclization – partial or circular. It is accepted in the Art that disulfide constrained peptides - as such - give generally weak binders ⁴⁴, but still some developments are worth mentioning. Disulfide bridges were hence used to build a structurally-biased peptide β -turn library to obtain a hairpin scaffold (rotamers) ⁴⁵. A well-placed Trp residue was even found essential in generating a stabilized hairpin structure. But given the need for the S-S bridging (generated by having twice a CT sequence), the fixed position Trp, it led to scaffolds of the type XCTWX₄LTCX (X being any random AA) ⁴⁵. Another disulfide strategy has been to use a constrained hairpin loop stabilized by 2 Cys residues as a very tight folding unit to keep a loop (randomized or not) inserted between both ends in a tight turn, hence constraining it indirectly by a clamping effect ⁴⁶. Template-constrained or non-peptide structures serving as scaffold for peptide ended up to be used to enhance bioavailability, and reached nanomolar affinities when applied to somatostatin analogues ^{47,48}. Other templating strategies copied/mimicked a given pattern, that is naturally occurring (like in a toxin, conotoxin or the insect antimicrobial peptides apidaecins and lebecins), and used this basis to obtain coactamer libraries. Those natural templates are either rich in Cys or in Pro. But in both cases, they provide a fixed backbone that is position-dependent ⁴⁹. Templating was also pursued with Zinc finger and Kunitz domains, avidins, anticalins,... ^{50,51}. Here again the skilled person may select from random peptides constrained within a well-defined fold, where fixed positions of crucial AA's are necessary to keep the fold. A stable fold based on a highly hydrophobic cluster, comprising a solvent-shielded Tryptophan (TRP) surrounded by polyprolines and known as a Trp-cage has been proposed as a stable scaffold of 20 AA's. This mimics the template of exendin-4, a peptide isolated from *Heloderma suspectum* ⁴⁴, it is thus once again not position-independent as the relative positions of the key AA's are fixed. It is neither length-independent as 20 AA's are minimally required to enable the stable folding ^{44,52,53}. In yet another strategy, a library of phylomers – i.e. a repertoire of naturally-occurring protein sub domains found in various bacterial species – has been proposed as a library of folds ¹⁶. The minimal sizes are around 23 AA's and often larger (around 40). Such sizes are needed to encode those simple folds. The genomic distance from humans renders this approach very difficult given the multiple immunogenic problems. It must also be noted that such strategy only allows for naturally occurring folds, those natural folds may at times not be

-4-

the best as evidenced by the classical finding of a more effective non-natural ATP-binding site in ribozyme using directed evolution, which found both the naturally occurring and the non-natural one. Concatenation – besides reducing total sequence space while keeping diversity – could also theoretically induce some tertiary structure constraints in the form of salt bridges and/or hydrophobic surfaces that could sterically stabilize the repeats towards each other. Until today, there was no successful attempt at generating the needed constraints from the primary sequences alone in a length and position independent manner so there is a need to develop such novel binders. We termed such linear primary-structure constrained random peptides sequamers.

The present invention provides a novel class of non-naturally occurring peptides with a range between 6-15 amino acids. These peptides have a specific density of proline and tryptophane residues in the sequence structure.

BRIEF DESCRIPTION OF THE DRAWINGS

With specific reference now to the figures, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the different embodiments of the present invention only. They are presented in the cause of providing what is believed to be the most useful and readily description of the principles and conceptual aspects of the invention. In this regard no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention. The description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

Fig. 1: shows the ELISA values of 254 unique binder sequences against 2 isoforms of the mitochondrial membrane channel VDAC. The ELISA values are expressed as a function of the density of Pro in those sequences, regardless of the PRO position and of the length of those sequences. The lengths were between 6 and 48 AA's. No optimal Pro content can be seen.

Fig. 2: shows the same ELISA data expressed as the densities of TRP within the primary sequences, the precise position of the TRP is not relevant, nor is the length of the peptides. Once again, no advantage of the number of TRP can be seen.

Fig. 3: In this figure, when the simple algorithm [W+P] is applied, regardless of the positions of those PRO and TRP within the primary sequences, and regardless of the length of the peptides, there is a clear optimal combined density of W+P. Even without the highest point (ELISA at 2), the optimum is clearly marked around 30-40 %.

Fig. 4: (A) the ELISA data for HSP70 and HSC70 vs. the density of Proline are shown, regardless of the conditions and the libraries screened. The ELISA values for HSP70 and HSC70 binders are shown as a function of either the frequency of Tryptophan (B), or the cumulative frequency of W+P in their sequences (C).

Fig. 5: Apparent K_d in planar bilayer experiments in the femtomolar range. K_d of a 7-mer, linearly constrained, peptide on the voltage gating charge of VDAC1, the mitochondrial ion channel used as target for one of the selections. The apparent affinity constant in this highly sensitive cell-free system (electrophysiological recording of ion channels reconstituted into planar lipid bilayer membranes) is extremely low. The value indicated by the fit is not totally correct because of the extremely low effective concentrations needed. It is fair and accurate to mention that the range lies in the femtomolar domain. (# of exp. =6, data+/-SEM)

Fig. 6: Sensorgram testing one ultra-high affinity linear peptide against HVDAC1 proteoliposomes immobilized into a SPR device.

Fig. 7: In A, % of total cell death as measured by Propidium iodide fluorescence in peptide-laden K562 cells using a FACS vs. the measured intracellular peptide concentrations calibrated by peptide-laden liposomes (LUV's calibrated for the cellular size). The peptides were labeled with carboxy-fluorescein. In B, the ratio of apoptotic vs. necrotic cellular in peptide-laden K562 cells using FACS. Apoptosis was quantified by flow cytometric detection of caspase-3 activity (white bars). Cells displaying PI-fluorescence on the flow cytometer without caspase-3 activation were considered as necrotic cells (black bars).

Fig. 8: illustrates the competition between 14 short peptides and a polyclonal HVDAC1 antibody (commercially obtained and tested) inside permeated and non-fixed living cells followed by FACS fluorescence. Note that this means that all the cell content containing the natural ligands and competitors for the target site are present. (A) The experimental outline of a competition experiment between a polyclonal HVDAC1-Ab (labeled with APC) and different peptides (labeled with FITC) for target HVDAC1. In blue, the predicted peptide displacement after Aby administration, seen as a decrease in FITC signal. In red, the Aby displacement after peptide administration, seen as decrease in APC-signal (for detailed procedure, see material and methods). (B) A density plot representation of peptide displacement by the HVDAC1-Ab (μM) is shown for the negative control polyG (left panel) and a positive peptide L7H (right panel), by flow cytometric analysis of the decrease in FITC-signal. All peptides and the Aby are administered at $1\mu\text{M}$ concentration. (C) Density plot representation of HVDAC1-Ab displacement by the same 2 peptides from panel B, measured by flow cytometric analysis of the decrease in APC-signal. Final concentrations are $1\mu\text{M}$ for both Aby and peptides, such

concentration, which is higher than the functional ones, was needed to ensure correct fluorescence readings. (D) Bar graph representation of the results of all 14 peptides, analyzed like mentioned in panel (B) and (C). In grey the peptide displacement by the HVDAC1-Ab, in black the HVDAC1-Ab displacement by the different peptides.

5

Fig. 9: K7H decreases the intracellular ATP-concentration in intact cells. Intracellular ATP concentrations were followed in time after mock and K7H electroporation of K562 cells. The relative values as % of the control are shown. Data are presented as means $\pm$ se (n=4). Statistical analysis was determined by a paired Student t test ($P<0.05$). The P value is mentioned when statistically significant.

10

Fig. 10: Cell death induction by K7H electroporation followed in time. K562 cells were either mock electroporated, electroporated with 100 μ M polyG or with 100 μ M K7H (both N-term FITC-labeled). Cell death modalities were determined by FACS at different time points. Apoptosis was measured through caspase-3 activity. Necrosis was the percentage Propidium iodide (PI) positive cells decreased with the percentage apoptotic cells. No distinction could be made between early and late apoptosis due to the FITC labeling of the peptides and thus the lack of a good double staining procedure. Mock electroporated cells were used as negative control. PolyG was a control peptide, which has no binding and no effect when inside the cells. Data are presented as means $\pm$ se (n=3). Statistical analysis determined by a paired Student t test confirmed that the differences between the K7H effect and the controls were significant ($P<0.05$).

15

20

Fig. 11: Modulation of cell death types after VDAC closure in relationship to the intracellular ATP availability. K562 cells were treated with PBS (-) or 800 μ M adenosine (+) for 24 hours. Cells were then mock electroporated and electroporated with 10 μ M of the mentioned peptide. Adenosine pre-treated cells were further incubated with 800 μ M adenosine. After electroporation the cell death types were determined by FACS at 24 hours (first panel) and 48 hours (second panel). Apoptosis was measured through caspase-3 activity. Necrosis was the percentage Propidium iodide (PI) positive cells decreased with the percentage apoptotic cells. Data are presented as means $\pm$ se (n=6 for A and n=3 for B and C). Statistical analysis was determined by a paired Student t test ($P<0.05$). The P value was mentioned when statistically significant.

25

30

Fig. 12: The different peptides have different balancing effects towards apoptosis and necrosis.

35

Fig. 13: The cell death induction after HVDAC1 closure by K7H depends on the metabolic state of cells. (A) Quiescent immature murine DC cells were incubated with LPS for

maturation. In both quiescent (left panels) and activated DC (right panels) apoptosis was measured by flow cytometric analysis of the double staining with annexin-V and PI after K7H-induced HVDAC1 closure at the indicated time points. All histograms show the difference in annexin-V and PI staining after mock electroporation (in black) and K7H electroporation (in red). Similar results were obtained in three separate experiments. (B) (B.1) K562 cells, incubated with or without 1U/ml erythropoietin, displayed differences in cell proliferation and oxygen consumption after 72 hours of incubation (left panel). After 72 hours of incubation, EPO-treated cells (with a '+' sign in the right figure) and untreated cells ('-' sign) were electroporated with nothing (mock), 100µM K7H or 100µM polyG (negative control) as described earlier and were followed for 48 hours. Apoptosis was quantified by flow cytometric detection of caspase-3 activity (white bars). Cells displaying PI-fluorescence on the flow cytometer without caspase-3 activation were seen as necrotic cells (black bars). Data are presented as means $\pm$ se (n>3). Statistical analysis was determined by a paired Student t test (P<0.05). The P value was mentioned when statistically significant.

Fig. 14: Interaction between VDAC closure and DMSO-induced apoptosis. K562 cells were incubated with 3% DMSO 5 hours after either mock, polyG and K7H electroporation. Electroporations without DMSO incubation are mentioned for easy comparison. Cell death was followed in time and measured by flow cytometry. The results shown are 48 hours after electroporation. Apoptosis (grey bars) was obtained with a caspase 3-activity detection kit and necrosis (white bars) with PI minus the % apoptotic cells.

Fig. 15: Cell death measured at the indicated times in the P815 cell line by trypan blue exclusion (experiments measured by 2 persons in a blind manner). Green box the control cells (addition of milliQ water at the same volume as the tests), white bars addition of ethanol in water at 0.01%, grey bars addition of the K7H peptide at 1.10⁻⁸ M together with the ethanol at 0.01%. (N=3 exp)

DETAILED DESCRIPTION OF THE INVENTION

The present invention will be described with respect to particular embodiments and with reference to certain drawings but the invention is not limited thereto but only by the claims. Any reference signs in the claims shall not be construed as limiting the scope. The drawings described are only schematic and are non-limiting. In the drawings, the size of some of the elements may be exaggerated and not drawn on scale for illustrative purposes. Where the term "comprising" is used in the present description and claims, it does not exclude other elements or steps. Where an indefinite or definite article is used when referring to a singular noun e.g. "a" or "an", "the", this includes a plural of that noun unless something else is

specifically stated. Furthermore, the terms first, second, third and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

The following terms or definitions are provided solely to aid in the understanding of the invention. Unless specifically defined herein, all terms used herein have the same meaning as they would to one skilled in the art of the present invention. Practitioners are particularly directed to Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 4th ed., Cold Spring Harbor Press, Plainsview, New York (2012); and Ausubel et al., *current Protocols in Molecular Biology* (Supplement 100), John Wiley & Sons, New York (2012), for definitions and terms of the art. The definitions provided herein should not be construed to have a scope less than understood by a person of ordinary skill in the art.

Non-covalent interactions of proteins with other molecules are at the basis of all biological processes in life. These interactions determine the strength and specificity of molecular recognition events. Accordingly, rational application of the principles behind such interactions, in combination with biotechnological methods, may provide a means to interfere (or intervene which is an equivalent term) with biological processes, or alternatively enable detection or even the purification of specific compounds.

The present invention provides a novel class of non-naturally occurring linear peptides.

These peptides have a length of between 6 and 15 amino acid residues and the specific density of proline (P) and tryptophane (W) is determinative for their constrained character. There are inherent advantages to the small size of these constrained peptides. These peptides will first enhance bioavailability in vivo but will also avoid the trigger of an immune response. Indeed under 9 amino-acid residues (8 to 11 residues for MHC class I molecules) and 13 residues (13 to 17 residues for MHC class II)¹², those peptides would not be able to be effective epitopes or antigenic determinants. Thus these shorter peptides would be nearly transparent for the immune system; hence ease and enhanced safety of use with potentially longer periods prior immunization against the product. Moreover short chains peptides can be chemically synthesized and can be considered to be equivalent with the flexibility of chemicals. In addition, we show in the experimental data section that these novel peptides are able to bind with subnanomolar (Kd) – affinity to selected proteins.

Accordingly, the present invention provides a non-naturally occurring polypeptide consisting of between 6 and 15 contiguous amino acids which contains at least one proline (P) and at least one tryptophan (W), wherein

- the frequency of P is between 10%-25% and,

-9-

- the frequency of W is between 6.7%-33% and,
- optionally one G is present in said polypeptide.

In yet another embodiment the invention provides a non-naturally occurring polypeptide consisting of between 6 and 15 contiguous amino acids which contains at least one proline (P)

5 and at least one tryptophan (W), wherein

- the frequency of P is between 10%-25% and,
 - the frequency of W is between 6.7%-33% and,
 - optionally one G is present in said polypeptide and,
 - the frequency of I is between 5.2%-6.6% in the peptides and wherein no I is
- 10 present in sequences smaller than 14 amino acids.

In yet another embodiment the invention provides a non-naturally occurring polypeptide consisting of between 6 and 15 contiguous amino acids which is characterized as having:

- between 1 and 4 P,
- between 1 and 5 W, and
- which contains optionally one G.

15

In yet another embodiment the invention provides a non-naturally occurring polypeptide consisting of between 6 and 15 contiguous amino acids which is characterized as having:

- between 1 and 4 P,
- between 1 and 5 W, and
- which contains optionally one G and
- which contains no I in peptides smaller than 14 amino acids.

20

The non-naturally occurring polypeptides of the invention are herein further defined as "the molecules of the invention".

In yet another embodiment the invention provides a random peptide library comprising the

25

non-naturally occurring polypeptides of the invention as defined herein above.
In yet another embodiment the invention provides the use of a peptide library comprising molecules of the invention for the selection of a molecule of the invention with a specificity against a specific target.

In specific embodiments the molecules of the invention bind to a target with an affinity characterized by a dissociation constant K_d lower than 1000 micromolar, preferably lower than 100 micromolar, more preferably lower than 10 micromolar, more preferably lower than 1 micromolar, more preferably lower than 0.1 micromolar, more preferably lower than 0.01 micromolar, more preferably lower than 1 nanomolar, more preferably lower than 0.1 nanomolar, more preferably lower than 0.01 nanomolar, more preferably lower than 1

30

picomolar.

35

In a particular embodiment the invention is also related to a method to produce a random

peptide library as defined herein above wherein said method comprises the following steps: i) designing the specific amino acid sequence for the peptides applying the rules for the frequency of P, W, G and I as defined herein above, ii) producing the designed peptides by chemical synthesis and iii) obtaining said random peptide library.

- 5 In specific embodiments the peptide library is a random peptide library. A suitable random peptide library of the invention comprises more than 1000, more than 10.000, more than 100.000, more than 10^6 , more than 10^7 molecules of the invention.

Throughout the application, the standard one letter notation of amino acids will be used. Typically, the term "amino acid" will refer to "proteinogenic amino acid", i.e. those amino acids
10 that are naturally present in proteins. Most particularly, the amino acids are in the L isomeric form. D amino acids are also envisaged.

In the present invention it is envisaged that in selected binders particular amino acids (amino acid residues, residues) can be substituted (mutated, changed, varied) without compromising the essential properties of the binder. Non-naturally peptide molecules of the invention can be
15 mutated so as to allow them to bind (bind to, bind with) with specific target molecules, but one can exactly identify which amino acid residues can (and which cannot) be varied for this purpose. The molecules of the invention are able to interfere with (influence, modify) biological processes through impeding (blocking, inhibiting) natural or synthetic chemical or enzymatic reactions or natural molecular recognition events, or through creation of non-natural molecular
20 recognition events. Some aspects of the present invention relate to methods to generate or optimize such compounds possible to detect molecular compounds of interest in a study sample and possibly to isolate molecular compounds of interest from this study sample.

The molecules of the invention can be used across a whole range of fields, including white biotechnology (or industrial biotechnology), red or medical biotechnology, green or agricultural
25 biotechnology, blue (or aquatic) biotechnology. They can be used to inhibit proteins, as well as to detect proteins, and this in all of these fields.

Instances of biological interference with the molecules of the invention include, without limitation, blocking of human receptors, binding to pathogenic species, and binding to disease- or disorder-related proteins. Such type of biological interference is typically intended to
30 diagnose or curate severe diseases or disorders. These applications belong to the field of therapeutic applications. Instances wherein specific probe molecules (probes) are applied to detect the presence of an analyte of interest (target analyte) in a given sample of interest (study sample), include, without limitation, experimental analyses of samples of human, animal, plant, bacterial, viral, biotechnological or synthetic origin. Such samples typically
35 contain biomolecules (e.g., polypeptides, polynucleotides, polysaccharides, hormones, vitamins or lipids, or derivatives thereof) that can interact specifically with a selected probe

molecule. The latter interaction typically gives rise to a characteristic (e.g., spectroscopic or radioactive) signal, indicative of the presence of said target analyte in this study sample. These applications belong to the field of analytical research and development and diagnostics.

5 The number of combinations of different types of probes and targets that are effectively used in medical and biotechnological applications is virtually unlimited. In view of the continuous evolution in these areas, there is an ongoing need for new analytical tools (e.g., probes) with desired physico-chemical properties (e.g., specificity, affinity, stability, solubility), as well as improved methods for the production, purification, testing and optimization of such compounds. Instances wherein specific ligand molecules (ligands) are applied to retain
10 (extract, isolate, purify, filter) other molecules of interest (targets, target analytes) in a given sample of interest (crude sample) include, without limitation, samples of human, animal, plant, bacterial, viral, biotechnological or synthetic origin containing biomolecules (e.g., polypeptides, polynucleotides, polysaccharides, hormones, vitamins or lipids, or derivatives thereof) that can interact (associate) with high specificity with selected ligand molecules, where the latter are
15 separated, or can be separated, from the crude sample (e.g., by attachment onto a solid support or by precipitation) for the purpose of co-separating the target molecules from the crude sample. These applications belong to the field of purification technology.

More specific examples of purification methods wherein the molecules of the invention can be used include affinity chromatography and immunoprecipitation. In view of the continuous
20 evolution in these areas, there is an ongoing need for new molecules for purification with desired physico-chemical properties (e.g., specificity, affinity, stability, solubility), as well as improved methods for the production, purification, testing and optimization of such compounds. Immunoglobulin molecules (antibodies, including homologs and derivatives) are widely used in all of the aforementioned fields. They can recognize a diverse repertoire of
25 target antigens and bind with great specificity. However, they suffer from many disadvantages including (i) the requirement of laboratory animals for the production of polyclonal antibodies (immunization technology), (ii) the requirement of complicated methods to derive monoclonal antibodies from polyclonal ensembles (hybridoma technology), (iii) the non-human nature of antibodies obtained through immunization of animal vertebrates (causing potential problems
30 related to immunogenicity), (iv) difficulties to convert non-human antibodies into human or humanized variants (e.g., causing affinity loss), (v) alternative production methods based on protein library display and selection usually do not yield high-affinity products, thereby requiring additional affinity enhancement steps, (vi) all production methods are time-consuming and require highly specialized researchers, (vii) standard immunoglobulins may
35 experience steric difficulties to reach their target binding sites in vivo, as opposed to in vitro test systems, (viii) native and, to a lesser extent, engineered antibodies may have suboptimal properties relating to hydrophobicity, immunogenicity, bivalency or effector function, (ix)

therapeutic antibodies must be stored at near freezing temperatures, (x) immunoglobulin products are generally digested in the gut and must therefore be administered via injection or infusion, (xi) antibodies experience difficulties to permeate the blood-brain barrier.

It is implicit that linearly constrained peptides can display very high affinities (Fig. 3 and Table 3). Such high affinities render them ideally suitable to be used as capture agents or scavengers in numerous analytical processes such as ELISA, solid-surface filtration, immuno-based columns and others. Indeed the performance limits of said analytical measurements methods are defined by the affinities of their capturing agents. Likewise, catalysis can also use high affinity chelating agents. An advantage of said short sequence peptides is that they can be chemically synthesized and hence would be cost-effective even in the case they would be considered as 'single use'.

In other embodiments the present invention also envisages the coupling of the non-naturally occurring molecules of the invention to different molecules (or agents). Such coupling is particularly carried out with a suitable linker moiety. The nature of such linker moieties is not vital to the invention, although long flexible linkers are typically not used. According to particular embodiments, each linker is independently selected from stretch of between 0 and 20 identical or non-identical units, wherein a unit is an amino acid, a monosaccharide, a nucleotide or a monomer. Non-identical units can be non-identical units of the same nature (e.g. different amino acids, or some copolymers). They can also be non-identical units of a different nature, e.g. a linker with amino acid and nucleotide units, or a heteropolymer (copolymer) comprising two or more different monomeric species. According to particular embodiments, the length of at least one, and particularly each linker is at least 1 unit. According to other particular embodiments no linker is provided. According to particular embodiments, all molecules in the linker are identical. Amino acids, monosaccharides and nucleotides and monomers have the same meaning as in the art. Note that particular examples of monomers include mimetics of natural monomers, e.g. non-proteinogenic or non-naturally occurring amino acids (e.g. carnitine, GABA, and L-DOPA, hydroxyproline and selenomethionine), peptide nucleic acid monomers, and the like. Examples of other suitable monomers include, but are not limited to, ethylene oxide, vinyl chloride, isoprene, lactic acid, olefins such as ethylene, propylene, amides occurring in polymers (e.g. acrylamide), acrylonitrile-butadiene-styrene monomers, ethylene vinyl acetate, and other organic molecules that are capable of polymer formation.

According to alternative embodiments, the linker units are chemical linkers, such as those generated by carbodiimide coupling. Examples of suitable carbodiimides include, but are not limited to, 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), N,N'-Diisopropylcarbodiimide

(DIC), and Dicyclohexylcarbodiimide (DCC). Another particularly envisaged chemical linker is 4, 7, 10-trioxatridecan-succinic acid (sometimes also designated as 4, 7, 10-trioxatridecan-succinamic acid) or Ttds.

5 In yet another particular embodiment the non-naturally occurring molecules of the invention can further comprise a detectable label. The detectable label can be N- or C-terminally (e.g. through the linker, or the linker itself can be used as the detectable label). Alternatively, the detectable label can refer to the use of one or more labeled amino acids (e.g. fluorescently or radioactively labeled amino acids). Note that in principle any known label for molecules of
10 proteinaceous nature can be used, as long as the label can be detected. Particularly envisaged labels include, but are not limited to, tags, fluorescent labels, enzyme substrates, enzymes, quantum dots, nanoparticles which may be (para)magnetic, radiolabels, optical labels and the like.

As with other moieties, since the molecules have two ends, it is envisaged that the molecules
15 will be fused to another moiety (e.g. a label) at both its N- and C-terminus. These two labels can be identical (yielding a stronger signal) or different (for different detection purposes). According to particular embodiments, the detectable label is GFP or biotin.

According to still other particular embodiments, the molecules of the invention may be fused to other moieties, e.g. to extend their half-life in vivo. Apart from increasing stability, such
20 moieties may also increase solubility of the molecule they are fused to. A well-known example of such moiety is PEG (polyethylene glycol). This moiety is particularly envisaged, as it can be used as linker as well as solubilizing moiety. Other examples include peptides and proteins or protein domains, or even whole proteins (e.g. GFP). In this regard, it should be noted that, like PEG, one moiety can have different functions or effects. For instance, a flag tag (sequence
25 DYKDDDDK is a peptide moiety that can be used as a label, but due to its charge density, it will also enhance solubilisation. PEGylation has already often been demonstrated to increase solubility of biopharmaceuticals (e.g. Veronese and Mero, *BioDrugs*. 2008; 22(5):315-29). Adding a peptide, polypeptide, protein or protein domain tag to a molecule of interest has been extensively described in the art. Examples include, but are not limited to, peptides derived from
30 synuclein (e.g. Park et al., *Protein Eng. Des. Sel.* 2004; 17:251-260), SET (solubility enhancing tag, Zhang et al., *Protein Expr Purif* 2004; 36:207-216), thioredoxin (TRX), Glutathione-S-transferase (GST), Maltose-binding protein (MBP), N-Utilization substance (NusA), small ubiquitin-like modifier (SUMO), ubiquitin (Ub), disulfide bond C (DsbC), Seventeen kilodalton protein (Skp), Phage T7 protein kinase fragment (T7PK), Protein G B1 domain, Protein A IgG
35 ZZ repeat domain, and bacterial immunoglobulin binding domains (Hutt et al., *J Biol Chem.*; 287(7):4462-9, 2012). The nature of the tag will depend on the application, as can be determined by the skilled person.

Apart from extending half-life, molecules may be fused to moieties that alter other or additional pharmacokinetic and pharmacodynamic properties. For instance, it is known that fusion with

albumin (e.g. human serum albumin), albumin-binding domain or a synthetic albumin-binding peptide improves pharmacokinetics and pharmacodynamics of different therapeutic proteins (Langenheim and Chen, *Endocrinol.*; 203(3):375-87, 2009). Another moiety that is often used is a fragment crystallizable region (Fc) of an antibody. The nature of these moieties is not vital to the invention and can be determined by the person skilled in the art depending on the application.

According to particular embodiments, the molecules are fused to an agarose bead, a latex bead, a cellulose bead, a magnetic bead, a silica bead, a polyacrylamide bead, a microsphere, a glass bead or any solid support (e.g. polystyrene, plastic, nitrocellulose membrane, glass).

Other moieties which are also envisaged in combination with the molecules described herein are targeting moieties. For instance, the molecules may be fused to e.g. an antibody, a peptide or a small molecule with a specificity for a given target. This is similar to the strategy which is outlined in WO2008148751. An extensive list of possible target moieties (also designated as 'binding regions' or 'binding domains' in WO2008148751) which can be combined with the molecules of the invention is described in WO2008148751 on page 3 (starting on line 26) and page 4 (ending on line 34): the term 'binding region' or 'binding domain' typically refers to a molecule that interacts with the target protein. In certain cases a binding domain is a chemical compound (e.g. a small compound with an affinity for at least one target protein) and in certain other cases a binding domain is a polypeptide, in certain other cases a binding domain is a protein domain. A protein binding domain is an element of overall protein structure that is self-stabilizing and often folds independently of the rest of the protein chain. Binding domains vary in length from between about 25 amino acids up to 500 amino acids and more. Many binding domains can be classified into folds and are recognizable, identifiable, 3-D structures. Some folds are so common in many different proteins that they are given special names. Non-limiting examples are Rossmann folds, TIM barrels, armadillo repeats, leucine zippers, cadherin domains, death effector domains, immunoglobulin-like domains, phosphotyrosine-binding domain, pleckstrin homology domain, src homology 2 domain, the BRCT domain of BRCA1, G-protein binding domains, the Eps 15 homology (EH) domain and the protein-binding domain of p53. Antibodies are the natural prototype of specifically binding proteins with specificity mediated through hypervariable loop regions, so called complementary determining regions (CDR). Although in general, antibody-like scaffolds have proven to work well as specific binders, it has become apparent that it is not compulsory to stick strictly to the paradigm of a rigid scaffold that displays CDR-like loops. In addition to antibodies, many other natural proteins mediate specific high-affinity interactions between domains. Alternatives to immunoglobulins have provided attractive starting points for the design of novel binding (recognition) molecules. The term scaffold, as used in this invention, refers to a protein framework that can carry altered amino acids or sequence insertions that confer binding to specific target proteins. Engineering scaffolds and designing libraries are mutually interdependent processes. In order to obtain specific binders, a combinatorial library of the

scaffold has to be generated. This is usually done at the DNA level by randomizing the codons at appropriate amino acid positions, by using either degenerate codons or trinucleotides. A wide range of different non- immunoglobulin scaffolds with widely diverse origins and characteristics are currently used for combinatorial library display. Some of them are comparable in size to a scFv of an antibody (about 30 kDa), while the majority of them are much smaller. Modular scaffolds based on repeat proteins vary in size depending on the number of repetitive units. A non-limiting list of examples comprise binders based on the human 10th fibronectin type III domain, binders based on lipocalins, binders based on SH3 domains, binders based on members of the knottin family, binders based on CTLA-4, T-cell receptors, neocarzinostatin, carbohydrate binding module 4-2, tendamistat, kunitz domain inhibitors, PDZ domains, Src homology domain (SH2), scorpion toxins, insect defensin A, plant homeodomain finger proteins, bacterial enzyme TEM- 1 beta-lactamase, Ig-binding domain of *Staphylococcus aureus* protein A, *E. coli* colicin E7 immunity protein, *E. coli* cytochrome b562, ankyrin repeat domains. Also included as binding domains are compounds with a specificity for a given target protein, cyclic and linear peptide binders, peptide aptamers, multivalent avimer proteins or small modular immunopharmaceutical drugs, ligands with a specificity for a receptor or a co-receptor, protein binding partners identified in a two-hybrid analysis, binding domains based on the specificity of the biotin-avidin high affinity interaction, binding domains based on the specificity of cyclophilin-FK506 binding proteins. Also included are lectins with an affinity for a specific carbohydrate structure.

According to yet other embodiments, the molecules of the invention can further comprise a sequence which mediates cell penetration (or cell translocation), i.e. the molecules are further modified through the recombinant or synthetic attachment of a cell penetration sequence. As such the molecules of the invention may be further fused or chemically coupled to a sequence facilitating transduction of the fusion or chemical coupled proteins into prokaryotic or eukaryotic cells. Cell-penetrating peptides (CPP) or protein transduction domain (PTD) sequences are well known in the art and include, but are not limited to the HIV TAT protein, a polyarginine sequence, penetratin and pep-1. Still other commonly used cell-permeable peptides (both natural and artificial peptides) are disclosed e.g. in Sawant and Torchilin, *Mol Biosyst.* 6(4):628-40, 2010; Noguchi et al., *Cell Transplant.* 19(6):649-54, 2010 and Lindgren and Langel, *Methods Mol Biol.* 683:3-19, 2011. Typical for CPP is their charge, so it is possible that some charged molecules described herein do not need a CPP to enter a cell. Indeed, as will be shown in the examples, it is possible to target signal peptides or intracellular regions, which require that the molecules are taken up by the cell, and this happens without fusion to a CPP.

In those instances where other moieties are fused to the molecules of the invention, it is envisaged in particular embodiments that these moieties can be removed from the molecule. Typically, this will be done through incorporating a specific protease cleavage site or an equivalent approach. This is particularly the case where the moiety is a large protein: in such

cases, the moiety may be cleaved off prior to using the molecule in any of the methods described herein (e.g. during purification of the molecules). The cleavage site may be incorporated separately or may be an integral part of the external linker.

- 5 According to specific embodiments, the total length of the molecules of the invention (further modified) described herein does not exceed 100 amino acids. More particularly, the length does not exceed 80, 70, 60, 50, 40 amino acids, 30 amino acids, 25 amino acids or even does not exceed 20 amino acids. Thus, if a cleavage site has been built in the molecule, the length restriction typically applies to the length after cleavage.
- 10 In particular embodiments the molecules of the invention may also be manufactured using suitable expression systems comprising bacterial cells, yeast cells, animal cells, insect cells, plant cells or transgenic animals or plants. The recombinant molecules of the invention may be purified by any conventional protein or peptide purification procedure close to homogeneity and/or be mixed with additives. In yet another embodiment said molecule of the invention is a
- 15 chemically modified polypeptide. Chemical synthesis enables the conjugation of other small molecules or incorporation of non-natural amino acids by design. In a particular embodiment the conjugation of small molecules to the molecule of the invention might lead to a potential application of these molecules in the growing area of targeted cytotoxic agents for anti-tumor therapy. Incorporation of non-natural amino acids into the peptide opens up the
- 20 possibility for greater chemical diversity, analogous to small-molecule medicinal chemistry approaches for developing high-affinity, high-specificity molecular recognition. Non-natural amino acids can also prevent rapid degradation of the molecules of the invention by rendering the molecule unrecognizable to proteases (e.g. serum or stomach). In yet another embodiment the molecules of the invention comprise modified amino acids such as a D-amino acid or a
- 25 chemically modified amino acid. In yet another embodiment said molecule consists of a mixture of natural amino acids and unnatural amino acids. In yet another embodiment the half-life of a peptide can be extended by modifications such as glycosylation (Haubner R. *et al* (2001) *J. Nucl. Med.* 42, 326-336), conjugation with polyethylene glycol (PEGylation, see Kim TH *et al* (2002) *Biomaterials* 23, 2311-2317), or engineering the peptide to associate with
- 30 serum albumin (see Koehler MF *et al* (2002) *Bioorg. Med. Chem. Lett.* 12, 2883-2886). The administration of a pharmaceutical composition comprising a molecule may be by way of oral, inhaled, transdermal or parenteral (including intravenous, intraperitoneal, intramuscular, intracavity, intrathecal, and subcutaneous) administration. Particularly preferred examples of delivery methods for molecules are a transdermal patch (Henry S *et al* (1998) *J. Pharm. Sci.*
- 35 87, 922-925), iontophoresis (Suzuki Y *et al* (2002) *J. Pharm. Sci.* 91, 350-361), sonophoresis (Boucaud A *et al* (2002) *J. Control. Release* 81, 113-119), aerosols (Duddu SP *et al* (2002) *Pharm. Res.* 19, 689-695), transfersomes or liposomes (Guo J *et al* (2000) *Drug Deliv.* 7, 113-116). The molecules may be administered alone or preferably formulated as a pharmaceutical composition. (this means methods are provided comprising administering the molecules alone,

or formulated as pharmaceutical composition).

In a specific embodiment the invention provides the molecules of the invention for the use as a medicament.

In yet another specific embodiment the invention provides the molecules of the invention as defined herein above but wherein the length is between 6 and 60 amino acids, for the use as a medicament.

In yet another specific embodiment the invention provides the molecules of the invention for the use as anti-cancer agents.

In yet another specific embodiment the invention provides the human VDAC1 binders as specified in Table 5 for the use of anti-cancer agents.

According to further embodiments, methods to screen for new compounds in cell lines (including, but not limited to, human, mammalian, insect and plant cell lines), pathogens or microbial organisms are provided. In these embodiments the molecules of the invention are applied as if they were a classical compound library in a use for screening. These methods allow rapid identification of compounds which have effect on growth, reproduction or survival of the cell line or organism under study, or of compounds which inhibit protein function of proteins in said cell line or protein, even without prior knowledge of the target. Thus, not only can new compounds be obtained using these screening methods, they also allow identification of new drug targets. For this reason, it may also be particularly interesting to use cell lines that model disease (such as e.g. cancer cell lines, or even cells directly isolated from a tumor). In order to make sure that the identified compound (i.e. for example one molecule of the invention selected from a random peptide library as described herein) is effective and without side effects, toxicity of the compound should be tested, particularly on vertebrate, most particularly in mammalian systems.

In yet another embodiment the present invention also includes isotopically labelled molecules, which are identical to those defined herein, but for the fact that one or more atoms are replaced by an atom having an atomic mass or mass number different from the atomic mass or mass number usually found in nature. Examples of isotopes that may be incorporated into molecules of the present invention include isotopes of hydrogen, carbon, nitrogen, oxygen, phosphorous, sulfur, fluorine and chlorine, such as ^2H , ^3H , ^{13}C , ^{11}C , ^{14}C , ^{15}N , ^{18}O , ^{17}O , ^{31}P , ^{32}P , ^{35}S , ^{18}F , and ^{36}Cl , respectively. Molecules of the present invention and pharmaceutically acceptable salts of said molecules or which contain the aforementioned isotopes and/or other isotopes of other atoms are within the scope of this invention. Certain isotopically labeled molecules of the present invention, for example those into which radioactive isotopes such as ^3H and ^{14}C are incorporated, are useful in drug and/or substrate tissue distribution assays. Tritiated, i.e., ^3H , and carbon-14, i.e., ^{14}C , isotopes are particularly preferred for their ease of preparation and detectability. Further, substitution with heavier isotopes such as deuterium, i.e., ^2H , may afford certain therapeutic advantages resulting from greater metabolic stability, for

example increased in vivo half-life or reduced dosage requirements and, hence, may be preferred in some circumstances, Isotopically labelled molecules of this invention may generally be prepared by carrying out the procedures disclosed in the Examples below, by substituting a readily available isotopically labelled reagent for a non-isotopically labelled reagent.

It is to be understood that although particular embodiments, specific configurations as well as materials and/or molecules, have been discussed herein for cells and methods according to the present invention, various changes or modifications in form and detail may be made without departing from the scope and spirit of this invention. The following examples are provided to better illustrate particular embodiments, and they should not be considered limiting the application. The application is limited only by the claims.

Examples

1. Identification of a novel class of linear constrained peptides

Phage-display was carried out against 3 different proteins in 4 different conditions. Panning performed using 5 different random linear epitopes libraries (7-mer, 6-mer, 12-mer, 15-mer, 60-mer). Most of these phage display libraries were commercially available. The selecting targets were: YVDAC1, HVDAC1, HSC70, HSP70.

The yeast and human isoforms of the mitochondrial Voltage-dependent anion channel (YVDAC1 or HVDAC1) are known to have quite diverging sequences even if they can both form a β -barrel inside phospholipids membranes. The conditions used were physiological pH 7.2 versus low pH 5.8 combined with room temperature (22 °C) or physiological temperature (37°C), giving thus 4 selection conditions 22°/pH7.2, 22°/pH5.8, 37°/pH7.2, 37°/pH5.8. Those differences in pH are also reported to have major effects on the VDAC channel structures, the channel entering super-closed states.

HSP70 and HSC70 are different proteins, while being both chaperones. Here various conditions were also used during panning: temperature at 22 °C, 37°C and 42°C.

This makes in total (14X5) different pannings. When analyzing all the obtained sequences, and focusing only on the highest binders (i.e. binders having ELISA values at least 3 times above the controls but preferably 5-10 times above the control values), it was apparent that regardless of the conditions and targets used that there was an over-representation of Tryptophan (amino acid symbol = W) and Proline (amino acid symbol = P) in the peptide binders relative to their predicted frequency for a random library, see Tables 1A and 1B.

-19-

Table 1A: Predicted frequencies of amino acids (aa) due to the stability of the genetic code. Frequencies of occurrences of aa, if codons are randomly used (in % of total aa content). For example, Alanine (A) is present at a frequency of 6 %, Tryptophan (W) at a frequency of 1.58%, Proline (P) at a frequency of 5.35%, and so on.

AA	Th.Rd Fre.
A	0.0599
V	0.0608
L	0.0794
I	0.0521
C	0.0265
M	0.0184
F	0.0224
Y	0.0308
W	0.0158
G	0.0721
S	0.087
T	0.0696
K	0.0547
R	0.1071
H	0.0304
D	0.0365
E	0.0471
N	0.0366
Q	0.0392
P	0.0535

25

Table 1B: Predicted vs. measured for all libraries and all conditions. A total of 538 binder sequences (16,181 AA's) were analyzed. Ratios of expected vs. measured AA frequencies for all sequences of the best binders. Only W, P and I were condition- and target-independent. All other deviations were condition and/or target-specific.

30

-20-

AA	Th.Rd Freq	All - all cond	ratio (M/Th)	
A	0.0599	0.053087	0.88626	
V	0.0608	0.039862	0.655618	
L	0.0794	0.087016	1.095915	
I	0.0521	0.010259	0.196909	I
C	0.0265	0.005315	0.200561	C
M	0.0184	0.012978	0.705336	
F	0.0224	0.029355	1.31051	
Y	0.0308	0.039058	1.268122	
W	0.0158	0.123107	7.791604	W
G	0.0721	0.058896	0.816869	
S	0.087	0.069464	0.798439	
T	0.0696	0.058031	0.833779	
K	0.0547	0.015636	0.285843	K
R	0.1071	0.051048	0.476634	
H	0.0304	0.060318	1.984133	H
D	0.0365	0.03776	1.03453	
E	0.0471	0.016377	0.347712	E
N	0.0366	0.022619	0.618009	
Q	0.0392	0.026513	0.676341	
P	0.0535	0.185155	3.460849	P
Cumulative freq =		1.001854		
Total # seq =		538		

From the variations in amino acid composition of the selected binder sequences, only Tryp (W) and Pro (P) were augmented by factors of nearly 4-fold and 8-fold regardless of the conditions used and regardless of the nature of the target. The decreased I (Isoleucine) density seemed to play only a minor role.

As can be seen in Table 2, adding the frequencies for both P and W leads to a constant ratio between the expected/predicted frequencies and the measured ones regardless of the conditions and the length. It is thus the cumulated density of both W and P that represents the optimal density (W+P) for the obtained binders which are further designated as a novel class of constrained linear peptides. The WP rule – by opposition to the densities of other amino-acids -was condition independent.

Table 2: frequency of W and P in the isolated peptides

	Th.Rd Freq		All - all cond		All - RT, pH6		All - RT, pH7.5		All - 37°pH7.5		All - 37°, pH6	
				ratio		ratio		ratio		ratio		ratio
W	0,0158		0,123107	7,791604	0,105933	6,704620028	0,124255	7,864236	0,098853	6,256494	0,145674	9,219848
P	0,0535		0,185155	3,460849	0,141119	2,637742456	0,229253	4,285096	0,223742	4,182099	0,19324	3,611972
SUM	0,0693		0,308263	4,448236	0,247052	3,564967069	0,353508	5,101119	0,322595	4,655049	0,338914	4,890535

Thus, the optimal density values of W+P were deduced from the densities yielding the highest ELISA values. See figures 1, 2, and 3 showing the ELISA's as a function of the density of W, of P and of W+P. Only W+P demonstrates an optimum. The boundaries of said optimum can be set at W+P densities between 18 and 50% or more preferably between 15 and 55% of total sequence length. Identical results are shown for binders with an affinity for HSP and HSC in Figure 4.

The W+P algorithm allows for the definition of a novel class of linearly constrained short peptides having a W+P density comprised between 15 and 50 % of the total number of AA in the sequence. The precise positions of those key structural residues and the length of the peptides are not relevant to the efficacy of binding the target. This was confirmed for sequence lengths ranging between 6 and 60 AA's.

The relevance of the W+P rule in terms of binding efficacy can also be illustrated by near identical sequences differing in terms of cumulative W+P frequency. See Table3.

A series of sequences exemplified, according to the invention, are depicted in Tables 5 and 6.

Table 3: ELISA values of closely related binders complying or failing to comply with the invention-defined cumulative W+P densities (= the W + P rule)

5	Sequence 1: FGTWAYYLSVATAYS (15 AA) - W+P rule not applied (W+P at 6.7%), G at 1/15 (or 7% which is the predicted random frequency). Poor binder, ELISA values (0.029 Test vs. 0.010 Ctrl.).
10	Sequence 2: FGTWATTYLVHSPTL (15 AA) - W+P rule applied (W+P at 20%), G at 1/15 Excellent binder, ELISA values (0.257 Test vs. 0.019 Ctrl.).
	Sequence 3: FGTWATTYLGHRLLIL (15 AA) – W+P rule not applied (W+P at 6.7%), G at 2/15 15 (twice random frequency) Poor binder, ELISA values (0.052 Test vs. 0.042 Ctrl)
	Sequence 4: PSAYLVPHAPWQGSV (15 AA) – W+P rule applied (W+P at 27%), G at 1/15 (random 20 frequency) Excellent binder, ELISA values (2,146 Test vs. 0.160 Ctrl)
	Sequence 5: PSAYLVTHAQWQGSV (15 AA) – W+P rule not applied (W+P at 13%), G idem 25 Poor binder, ELISA values (0.080 Test vs. 0.040 Ctrl)

2. Affinity of the linear constrained peptides in a cell free system

- 30 We observed that the higher ELISA values obtained with the binders translated into high affinity and highly effective binders.
- Figure 5 shows the apparent K_d in planar bilayer experiments in the femtomolar range.
- Figure 6 shows the surface plasmon resonance (SPR) data of 4 different experiments wherein one high affinity linear peptide against HVDAC1 proteoliposomes, immobilized into a SPR
35 device, are measured. The maximal binding capacity that could be expected during those experiments was 36.8 RU assuming a one-to-one binding (see figure legend for details). There is a very good agreement between the chip immobilized VADC protein (at 30 femtoMol) and the peptide that becomes bound (1.10^{-14} M). Hence the experimental result closely matches the theoretical predictions and confirms the ultra low K_d in the femtomolar range in

cell-free systems.

Peptide binding against VADC1 reconstituted in proteoliposomes. Starting at 1.10^{-16} M, one starts to see selective binding onto the target. 1.10^{-22} and 1.10^{-17} M display no measurable signals just on-off artifacts from the SPR device. Selective binding was obtained by subtracting from the total sensorgram signal the matched signals due to i. plain liposomes devoid of any VADC1 protein (negative control) and ii. the buffer bulk effects. Every measurement was always performed with those 2 matching controls during the experimental runs.

Those binding studies were performed on a chip with a low VDAC-immobilization level around 1000 response units (RU), as recommended for kinetic measurements where the binding rate is determined by the interaction kinetics and not limited by mass transport of the analyte to the

$$B_{\text{cap}} (\text{in RU}) = n \times \text{captured ligand (in RU)} \times \frac{MW_{\text{anal}}}{MW_{\text{ligand}}}$$

surface (Biacore handbook, ⁵⁴). In those conditions, a theoretical maximal binding capacity (B_{cap}) can be calculated using the following formula:

In these experiments, VDAC is the captured ligand (MW 30,000 da) and the flowing analyte is the K7H peptide (MW 1104 da). Considering a one to one binding stoichiometry (n), the maximum peptide binding response at 1000 RU would be 36.8 RU. Having 1000 RU worth of VDAC as capture agent on the chip corresponds to 1 ng of protein or to 30 femtoM. This corresponds remarkably well with the amounts of bound peptide (at 1.10^{-14} M).

3. Effects and affinity of the linear constrained peptides in a cellular system

In the next experiment we convincingly showed that the linear constrained peptides with an affinity for HVDAC1 can block their target inside the cells, thereby inducing the effect of: switching aerobic metabolism off, hence bringing the cell in the cell death pathways. The effective intracellular concentrations (as measured by labeled peptides are in the low nanomolar affinity range (5 of the most active peptides are effective below 3 nanoM).

We showed that HVDAC1 specific binding peptides are able to decrease the ATP levels in living cells, by switching off 80% of the mitochondrial aerobic metabolism (see Figure 9).

Depending on the cellular condition –i.e. the cellular balance between energy production and energy needs - this ATP block may have different outcomes.

There are 3 possible scenarios:

- A. If cells are very actively dividing such as cancer cells: the anti-HVADC1 peptides kill up to 80 % of the cells that are treated with the peptides. See below 2A. Such peptide sequences when delivered intracellularly have applications in cancer therapy and are lead molecules for all potential clinical applications hereby linked. Any growth process

like benign tumors, malignant tumors, tissue overgrowths, tissue neogenesis, angiogenesis are potential clinical applications for those HVDAC1 blockers in this case.

5 B. If cells are quiescent, the peptides have no effect on the cellular death rates. See below 2B. Hence misdirected peptides should have lower rates or no rates of side-effects, except on highly dividing cells.

10 C. If cells are already undergoing apoptosis, blocking HVADC1 and ATP production blocks the apoptotic process. The same peptide sequences are thus useful for all neurological pathologies involving an increased apoptotic process. Any process where apoptosis or necroptosis is present can be a potential clinical application for the HVDAC1 blockers in this case.

If cells are very actively dividing like in cancer cell lines, blocking ATP forces the cells to enter cellular death. This cell death is split between apoptosis and necrosis depending on the intracellular levels of ATP at the moment of the challenge with the peptides.

15 In the presence of high intracellular ATP concentrations, apoptotic cell death is favored, whereas in lowered intracellular ATP concentrations, necrosis is favored. Total cellular death remains the same. ATP levels in the cells were increased (measured and calibrated) according to ^{55,56}. In particular cases the different binding peptides have different balancing effects towards apoptosis and necrosis. The cell death induced by HVDAC1 block depends on the metabolic needs of the cells. In quiescent cells, or in primary cultures (DC immatures), the peptides have no effect. The basal metabolism that is left during the block is sufficient. In activated cells (see activated mature DC cells), apoptotic first then late apoptosis and necrosis can be measured killing 80% of the population 2 days before the die off of the primary culture.

20 In cells where apoptosis is triggered at the same time than the peptide is given, inhibiting ATP production leads to the suppression of this apoptotic cell death. This evidences that in cells where HVDAC1 is closed, apoptosis, which is an active process and hence needs ATP, is markedly decreased

30 In Figure 14, the upper bar represents the addition of the effects of both the separate DMSO treatment and peptide treatment; this is thus an expected value if the combined effects were to be strictly additive. The second bar in figure 14 represents the result of such cumulative experiment when both treatments are given together. The total cell death is decreased (difference between 1st and second line) and the apoptosis linked to the DMSO treatment is replaced by necrosis (grey replaced by white in the 2nd line).

35 In Figure 15, P815 cells are applied, a cell line that is highly sensitive to apoptosis induced by extremely low doses of ethanol, the simultaneous addition of a HVDAC1 blocker at 10 nM and ethanol at 0.01% could decrease the ethanol-induced cell death by 40%.

4. Competition between linear constrained peptides and antibodies for their target inside the cells.

When equimolar concentrations of antibodies and various selected binding peptides against the target HVDAC1 are applied intracellularly it was witnessed that these peptides are able to

5 displace the antibodies. Those results were obtained both with monoclonal and polyclonal antibodies, and were specific for the channel isoforms. Only HVDAC1 antibodies (nor HVDAC2, nor HVDAC3) markedly showed competition with the peptides. Depending on their sequences, the peptides (see figure 8) could displace between 20 and 80% of the already bound antibodies.

- 10 Figure 8 illustrates the competition between 14 short peptides and a polyclonal HVDAC1 antibody (commercially obtained and tested) inside permeated and non-fixed living cells followed by FACS fluorescence.

5. Production of Linearly Constrained Random Epitopes (LiCoRE) libraries

- 15 Based on the obtained data on the binders in Example 1 it is possible to generate peptide libraries comprising the linearly constrained peptides of the invention (herein designated as LiCoRE libraries).

- Generally linear (non-constrained) random peptide libraries contain too many sequences that are not useful. This leads to a waste of time during the construction of such libraries, and add
- 20 loss of time during screening by yielding too many poor quality binders even when their apparent K_d seems highly desirable.

- In the present example the LiCoRE libraries are constructed around the structurally-relevant randomized peptides, enabling to focus both library-making and screening processes to the
- 25 structural fractions of the libraries. This allows a better use of the power of randomization as the libraries will include more of the relevant constrained sequence spaces. The library is made up of various sub-libraries that are then pooled together into the final product to reach the desirable variation numbers. Each of those sub-libraries are also built by having the needed proportion of constraining residues (according to the W+P algorithm defined in
- 30 Example 1).

5.1 A random peptide library with 6 positions

- One starts with a set of 2 fixed positions where W and P are introduced [i.e. $W+P = 33\%$] and randomize all the other remaining positions to obtain a sub-library. The 4 randomized positions
- 35 are obtained via the NNS or another appropriate technique. Then in a next iteration 2 other possible fixed positions are chosen and the randomization process for the remaining 4 positions is repeated. This requires a total of 30 different randomized sub-libraries (SB) (6 times 5 fixed positions), all of which can be built in parallel prior their merging into a single shape library. This covers all W-P arrangements, and has to be repeated for the case of the

-26-

W-W and P-P combinations, i.e. an additional 60 randomized sub-libraries.

The total variation number of such library is then 90.20^4 or $1.48 \cdot 10^6$. This value contains essentially the whole of the possible sequence space with a 99% confidence interval. As the most of the sequence space in linear random epitope libraries is not constrained, it is not structurally-relevant, - here in the invention - by decreasing the variation in the random sequence space we eliminate most of those unconstrained sequences, and we keep in the library the shape space, which is more relevant in the case of binders. Table 4 depicts the building scheme.

10 Table 4: The "open spaces" can be any amino acid except W or P:

Positions		1	2	3	4	5	6
Iteration 1	SB1	W	P				
	SB2	W		P			
	SB3	W			P		
	...	W				P	
	...	W					P
Iteration 2	...	P	W				
	...		W	P			
	...		W		P		
	SB9		W			P	
	...		W				P
	...		W				
Iteration 3	...			W			
	...	P		W			
	...		P	W			
	...			W	P		
	...			W		P	
	SB17			W			P
Iteration 4	...	P			W		
	...		P		W		
	...			P	W		
	...				W	P	
	...				W		P
Iteration 5	...	P				W	
	...		P			W	
	...			P		W	

-27-

Positions		1	2	3	4	5	6
	...				P	W	
	...					W	P
Iteration 6	...	P					W
	...		P				W
	...			P			W
	...				P		W
	...					P	W
	SB30					P	W

5.2 A random peptide library with 6 amino acids and a total of 50% W or P

In the case for the construction of a random peptide library with 6 amino acids and a total of 50% W or P, if X is a W or a P, there are: 6 times 5 times 4, that are 120 sub-libraries containing 3 times X. This has to be multiplied by 6 to account for all cases: XXX being WWW, WWP, WPP, PPP, PPW, PWW. This gives a total of 720 sets to be pooled. Once again the total variation number ($5.76 \cdot 10^6$) is well within the possible content of a library, and corresponds to 3 fully randomized positions.

10

5.3 A random peptide library with 7 or more amino acids and a total of 50% W or P

A 50% strategy applied to a longer peptide library (like a 7-mer or an 8-mer) requires 1260 or 2016 independent sub-libraries to be merged and enables the exploration of a larger shape space with variation numbers at $2 \cdot 10^8$ (6 times $7 \times 6 \times 5 \times 20^4$) and $3.23 \cdot 10^8$ (6 times $8 \times 7 \times 6 \times 20^4$). This keeps the fully randomized positions at 4 and 5 respectively.

15

5.4 A random peptide library with 7 amino acids with 2 fixed positions being W or P (28 % of W or P residues)

In the case of a random library with 7 amino acids with 2 fixed positions (W or P), a total of 144 sub-libraries is synthesized. 36 libraries with 2 fixed positions multiplied by 4 (WW, WP, PW and PP as possible combinations). Here with 5 fully randomized positions, the total variation numbers is $4.8 \cdot 10^8$. Those iterations represent the combinatorial arrangements of 7-mers with 3 fixed positions.

20

It is apparent that still other examples for libraries made of longer peptides, up to 12-13 AA's can be conceived following the same principles as in examples 5.1, 5.2 or 5.3.

25

It is worth noting that the number of sub-libraries that are needed does not increase markedly because the ratios of fixed positions relative to total length of the sequences remain analogous or the same. Those will always be between 18% and 55% of W+P, and the free randomized positions will always be 3, 4, 5, or 6, 8, 10 when doubling the examples mentioned above.

30

Table 5: examples of HVDAC1 binders. The peptides R7H, S7H and D7H (highlighted in red) do not comply with the W + P rule (as defined in the claims) and gave poor binding results in the ELISA experiments.

	K7H		W6H		W7H		V6H		Reference
X	WSGK PAT		WRWPT		WRIWPT WRIWPT		WTGSV W		
#	7		6		12		6		
W	1	14.28%	2	33.3%	4	33.3%	2	33.3%	6.7-33%
P	1	14.28%	1	16.67%	2	16.6%	0	-	10-25%
G	1		0	-	0	-	1	16.67%	0-7% / 1G
I	0	-	1	16.67%	2	16.6%	0	-	5.2-6.6%
W		28.6%		50%		50%		50%	
	Y7H		L7H		P7H		R7H		Reference
S	(PSAY LVPH APWQ GSV) x2		FGTWAT T YLWHSP TL		(KPPHK T)x4		RSRLA R		
#	1X (15) and 2X(30)		15		24		6		
W	2	6.67%	2	13.3%	0	-	0	-	6.7-33%
P	6	20%	1	6.67%	8	33.33%	0	-	10-25%
G	1	3.3%	0	-	0	-	0	-	0-7% / 1G
I	-	-	0	-	0	-	0	-	5.2-6.6%
W		27%		20%		33.3%		0%	
	M7H		S7H		F6H		N7H		Reference
S	LPGA MWIH PSPW TSS		SGSLAG	0.078	PPAPVL P SGHRT WFI		NYLAG AM		
#	15		6		15		7		

-29-

W	2	13.33%	0	-	1	6.67%	0	-	6.7-33%
P	3	20%	0	-	4	26.67%	0	-	10-25%
G	1	6.67%	2	33.3%	1	6.67%	1	14.28%	0-7% / 1G
I	1	6.67%	0	-	1	6.67%	0	-	5.2-6.6%
W		33%		0%		33.3%		0%	
	F7H		F8H		D7H				Reference
S	(PPAP VLP SGHR TWFI) x2		(PPAPVL P SGHRTW FI)x3		(LDQER P)x2	0.22			
#	30		45		12				
W	2	6.67%	3	6.67%	0	0			6.7-33%
P	8	26.67%	12	26.67%	2	16.6			10-25%
G	2	6.67%	3	6.67%	0				0-7% / 1G
I	2	6.67%	3	6.67%	0				5.2-6.6%
W		33.3%		33.3%		16.6%			

-30-

Table 6: examples of linear constrained peptides according to the invention binding to HSP70 and HSC70 binders.

	1. (RLWWRP)2x	HSC-37°C / HSC-RT / HSC-42°C HSP-37°C / HSP-RT / HSP-42°C
5	2. (RWW RTP) 2x	HSC-RT / HSC-37°C / HSP-RT / HSP-37°C
	3. (WYPWRT) 2x	HSP-RT / HSC-RT
	4. (RYHWWH) 2x	HSC-37°C / HSC-42°C / HSP-37°C
	5. (RWWPLR) 2x	HSC-37°C / HSP-37°C / HSP-42°C
	6. (IWRWPG) 2x	HSC-37°C
10	7. (HWLHRW) 2x	HSC-37°C
	8. (WWHRPT) 2x	HSC-42°C
	9. (WWSRWD) 2x	HSC-42°C
	10. (IPWWRC) 2x	HSC-RT / HSP-42°C
	11. (WRWPTL) 2x	HSC-RT / HSP-42°C
15	12. (WRPTRW) 2x	HSC-RT / HSP-RT / HSP-42°C
	13. (RSWFYP) 2x	HSC-RT
	14. (LRHRWM) 2x	HSP-37°C
	15. (HTWWRP) 2x	HSP-RT
	16. (HFLRPW) 2x	HSP-RT
20	17. (WWSRWN) 2x	HSP-42°C
	18. (YYVMVP) 2x	HSP-42°C
	19. (IWFWPR) 2x	HSP-42°C
	20. (HVRPWW) 2x	HSC-37°C
	21. (LSPWWWAPANRGSLT) 2x	HSC-37°C
25	22. RRPFGWWRLRPS	HSC-37°C
	23. (QHAWWR) 2x	HSC-RT
	24. VVPFYAAQPAMA	HSC-42°C / HSP-37°C
	25. (PLHRAWRPWWPSNSP) 2x	HSC-37°C

30

35

Reference List

- 5 1. Sijbren, O., Furlan, R. L. E. & Sanders, J. K. M. Selection and amplification of hosts
from dynamic combinatorial libraries of macrocyclic disulfides. *Science* **297**, 590-593
(2002).
2. Smith, G. P. Filamentous fusion phage: Novel expression vectors that display cloned
antigen on the virion surface. *Science* **228**, 1315-1317 (2012).
- 10 3. Smith, G. P. & Scott, J. K. Libraries of peptides and proteins displayed on filamentous
phages. *Methods Enzymol.* **217**, 228-257 (1993).
4. Clackson, T. & Wells, J. A. *In vitro* selection from protein and peptide libraries. *TIB Tech*
12:173-184. **12**, 173-184 (1994).
5. Roman, I., Figys, J., Steurs, G. & Zizi, M. Hunting interactomes of a membrane
protein: obtaining the largest set of VDAC-interacting protein epitopes. *Molecular &*
15 *Cellular Proteomics* (2006).
6. Devlin, J. J., Panganiban, F. C. & Devlin, P. E. Random libraries: a source of specific
protein binding molecules. *Science* **249**, 404-406 (1990).
7. Jespers, L., Lasters, I. & Winter, G. Concatenated nucleic acid sequences.
WO00XCCCC. 2002. Ref Type: Patent
- 20 8. Bogarad, L. D. & Deem, M. W. *Proc. Natl. Acad. Sci. USA*, 96:2591-2595. **96**, 2591-
2595 (1999).
9. Wolynes, P. G. *Proc. Natl. Acad. Sci. USA* **96**, 9568-9573 (1996).
10. Coles, M. *et al. Curr. Biol.* , 9:1158-1168. **9**, 1158-1168 (1999).
11. Lang, D., Thomas, R., Henn-Sax, M., Sterner, R. & Wilmanns, M. *Science* **289**, 1546-
25 1550 (2000).
12. Alberts *Molecular Biology of the Cell*. Garland Science, New York, NY, USA (2002).
13. Yokota, T., Milenic, D. E., Whitlow, M. & Schlom, J. Rapid tumor penetration of a
single-chain Fv and comparison with other immunoglobulin forms. *Cancer Res.* **52**,
3402-3408 (1992).
- 30 14. Blumenthal, R. D., Kashi, R., Sharkey, R. M. & Goldenberg, D. M. Quantitative and
qualitative effects of experimental radioimmunotherapy on tumor vascular
permeability. *Int. J. Cancer* **61**, 557-566 (1995).
15. Ladner, R. C., Sato, A. K., Gorzelany, J. & de Souza, M. Phage display-derived
peptides as therapeutic alternatives to antibodies. *Drug Discovery Today* **9**, 525-529
35 (2004).
16. Watt, P. M. Screening for peptide drugs from the natural repertoire of biodiverse
protein folds. *Nat. Biotechnol.* **24**, 177-183 (2006).
17. Nilsson, F., Tarli, L., Viti, F. & Neri, D. The use of phage display for the development of
tumour targeting agents. *Adv. Drug Deliv. Rev.* **43**, 165-196 (2000).

18. Willick, E. G. *Constrained Peptides*. Bentham Science Publishers, (2004).
19. Pak, V. V., Koo, M., Yun, L. & Kwon, D. Y. Recognized sequence and conformation in design of linear peptides as a competitive inhibitor for HMG-CoA reductase. *J. Mol. Recognit.* **20**, 197-203 (2007).
- 5 20. Beckers, T. *et al.* Structure-function studies of linear and cyclized peptide antagonists of the GnRH receptor. *Biochem. Biophys. Res. Commun.* **289**, 653-663 (2001).
21. Herman, R. E. *et al.* Phage Display Screening of Epithelial Cell Monolayers Treated with EGTA: Identification of Peptide FDFWITP that Modulates Tight Junction Activity. *J. Biomol. Screen.* **12**, 1092-1101 (2007).
- 10 22. Hruby, V. J. Conformational restrictions of biologically active peptides via amino acid side chain groups. *Life Sci.* **31**, 189-199 (1982).
23. Hruby, V. J. Designing molecules: specific peptides for specific receptors. *Epilepsia* **30**, S64-S68 (1989).
24. Sieber, S. A. & Marahiel, M. A. Learning from nature's drug factories: nonribosomal synthesis of macrocyclic peptides. *J. Bacteriol.* **185**, 7036-7043 (2003).
- 15 25. Perdih, A. & Kikelj, D. The application of Freidinger lactams and their analogs in the design of conformationally constrained peptidomimetics. *Curr. Med. Chem.* **13**, 1525-1556 (2006).
26. Fletcher, J. M. & Hughes, R. A. Novel monocyclic and bicyclic loop mimetics of brain-derived neurotrophic factor. *J. Pept. Sci.* **12**, 515-524 (2006).
- 20 27. Tamamura, H. *et al.* Structure-activity relationship studies on CXCR4 antagonists having cyclic pentapeptide scaffolds. *Org. Biomol. Chem.* **3**, 4392-4394 (2005).
28. Tyndall, J. D. *et al.* Synthesis, stability, antiviral activity, and protease-bound structures of substrate-mimicking constrained macrocyclic inhibitors of HIV-1 protease. *J. Med. Chem.* **43**, 3495-3504 (2000).
- 25 29. Tan, G. H., Yusoff, K., Seow, H. F. & Tan, W. S. A phage-displayed single chain variable fragment that interacts with hepatitis B core antigen: library construction, selection and diagnosis. *J. Clin. Virol.* **38**, 49-56 (2007).
- 30 30. Tan, W. S., Tan, G. H., Yusoff, K. & Seow, H. F. A phage-displayed cyclic peptide that interacts tightly with the immunodominant region of hepatitis B surface antigen. *J. Clin. Virol.* **34**, 35-41 (2005).
31. Laudet, B. *et al.* Structure-based design of small peptide inhibitors of protein kinase CK2 subunit interaction. *Biochem. J.* **408**, 363-373 (2007).
32. Nam, N. H., Ye, G., Sun, G. & Parang, K. Conformationally constrained peptide analogues of pTyr-Glu-Glu-Ile as inhibitors of the Src SH2 domain binding. *J. Med. Chem.* **47**, 3131-3141 (2004).
- 35 33. Khananshvili, D. *et al.* Inhibition of the cardiac sarcolemma Na⁺/Ca²⁺ exchanger by conformationally constrained small cyclic peptides. *Mol. Pharmacol.* **51**, 126-131 (1997).

34. Karle, I. L., Ranganathan, D. & Lakshmi, C. Demonstration of a cystine unit as a promising turn scaffold for the design of a parallel U-shaped two-helix bundle motif: crystal structure of the homodimer Cys(Aib(n))(2) (n = 3, 4). *Biopolymers* **59**, 301-304 (2001).
- 5 35. Ranganathan, D., Haridas, V., Nagaraj, R. & Karle, I. L. Double-helical cyclic peptides: design, synthesis, and crystal structure of figure-eight mirror-image conformers of adamantane-constrained cystine-containing cyclic peptide cyclo (Adm-Cyst)(3). *J. Org. Chem.* **65**, 4415-4422 (2000).
- 10 36. Ueda, S. *et al.* Structure-activity relationships of cyclic peptide-based chemokine receptor CXCR4 antagonists: disclosing the importance of side-chain and backbone functionalities. *J. Med. Chem.* **50**, 192-198 (2007).
37. Sia, S. K., Carr, P. A., Cochran, A. G., Malashkevich, V. N. & Kim, P. S. Short constrained peptides that inhibit HIV-1 entry. *Proc. Natl. Acad. Sci. U. S. A* **99**, 14664-14669 (2002).
- 15 38. Helicomimetics and stabilized LXXLL peptidomimetics. PCT 60/274,846(WO/2002/072597). 2002. Ref Type: Patent
39. Kostidis, S. *et al.* The relative orientation of the Arg and asp side chains defined by a pseudodihedral angle as a key criterion for evaluating the structure-activity relationship of RGD peptides. *J. Pept. Sci.* **10**, 494-509 (2004).
- 20 40. Andronati, S. A., Karaseva, T. L. & Krysko A.A. Peptidomimetics - antagonists of the fibrinogen receptors: molecular design, structures, properties and therapeutic applications. *Curr Med Chem.* 2004 May;11(9):1183-211. **May 11**, 1183-1211 (2004).
41. Zhu, J. & Marchant, R. E. Solid-Phase Synthesis of Tailed Cyclic RGD Peptides Using Glutamic Acid: Unexpected Glutarimide Formation. *J Pept Sci.* **14**, 690-696 (2008).
- 25 42. Reichelt, A. & Martin, S. F. Synthesis and properties of cyclopropane-derived peptidomimetics. *Acc. Chem. Res.* **39**, 433-442 (2006).
43. Shepherd, N. S. *et al.* Modular alpha-Helical Mimetics with Antiviral Activity against Respiratory Syncytial Virus. *J. Am. Chem. Soc.* **128**, 13284-13289 (2006).
44. Herman, R. E. *et al.* The Trp Cage Motif as a Scaffold for the Display of a Randomized Peptide Library on Bacteriophage T7. *Journal of Biological Chemistry* **282**, 9813-9824 (2007).
- 30 45. Cochran, A. G. *et al.* A minimal peptide scaffold for beta-turn display: optimizing a strand position in disulfide-cyclized beta-hairpins. *J Am Chem Soc.* **123**, 625-632 (2001).
- 35 46. Souriau, C., Chiche, L., Irving, R. & Hudson, P. New binding specificities derived from Min-23, a small cystine-stabilized peptidic scaffold. *Biochemistry* **44**, 7143-7155 (2005).
47. Cheng, R. P., Suich, D. J., Cheng, H., Roder, H. & DeGrado, W. F. Template-constrained somatostatin analogues: a biphenyl linker induces a type-V' turn. *J. Am.*

- Chem. Soc.* **123**, 12710-12711 (2001).
48. Suich, D. J. *et al.* Template-constrained cyclic peptide analogues of somatostatin: subtype-selective binding to somatostatin receptors and antiangiogenic activity. *Bioorg. Med. Chem.* **8**, 2229-2241 (2000).
- 5 49. Real, E. *et al.* Antiviral drug discovery strategy using combinatorial libraries of structurally constrained peptides. *J. Virol.* **78**, 7410-7417 (2004).
50. Binz, H., Amstutz, P. & Pluckthun, A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **23**, 1257-1268 (2005).
- 10 51. Ladner, R. C. & Ley, A. C. Novel frameworks as a source of high-affinity ligands. *Curr. Opin. Biotechnol.* **12**, 406-410 (2001).
52. Quay, S. C. *et al.* Phage displayed Trp cage ligands. 10976942 filed on 10/29/2004(7329725). 12-2-2008. Ref Type: Patent
53. Neidigh, J. W., Fesinmeyer, R. M. & Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Biol.* **9**, 425-430 (2002).
- 15 54. Myszka, D. G. Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr. Opin. Biotechnol.* **8**, 50-57 (1997).
55. da Silva, C. G., Jarzyna, R., Specht, A. & Kaczmarek, E. Extracellular nucleotides and adenosine independently activate AMP-activated protein kinase in endothelial cells: involvement of P2 receptors and adenosine transporters
- 20 1. *Circ. Res.* **98**, e39-e47 (2006).
56. Duflot, S. *et al.* ATP-sensitive K(+) channels regulate the concentrative adenosine transporter CNT2 following activation by A(1) adenosine receptors. *Mol. Cell Biol.* **24**, 2710-2719 (2004).
- 25

CLAIMS

1. A non-naturally occurring polypeptide consisting of between 6 and 15 contiguous amino acids which is characterized as having:
 - 5 - between 1 and 4 P,
 - between 1 and 5 W,
 - which contains optionally one G.
2. A non-naturally occurring polypeptide according to claim 1, wherein:
 - 10 - the frequency of P is between 10% and 25%, and
 - the frequency of W is between 6.7% and 33%.
3. A non-naturally occurring polypeptide according to claim 1, wherein the total W+P frequency of the total sequence length is between 15% and 55%.
4. A non-naturally occurring polypeptide according to claim 1 which contains no I in peptides smaller than 14 amino acids.
- 15 5. A non-naturally occurring polypeptide according to any one of claims 1 to 4 wherein the non-naturally occurring polypeptide comprises one or more detectable labels.
6. A non-naturally occurring polypeptide according to any one of preceding claims, wherein the non-naturally occurring polypeptide is fused to a moiety.
7. A random peptide library comprising non-naturally occurring polypeptides as defined in
20 any one of claims 1 to 4.
8. A method to produce a random peptide library according to claim 7 wherein said method comprises the following steps: i) designing the specific amino acid sequence for the peptides applying the rules for the frequency of P, W, G and I as defined in claims 2 or 3, ii) producing the designed peptides by chemical synthesis and iii) obtaining said random
25 peptide library.
9. A non-naturally occurring polypeptide as defined in any one of claims 1 to 6 for use as a medicament.
10. A non-naturally occurring polypeptide as defined in any one of claims 1 to 6 but wherein the length of the polypeptides is between 6 and 60 amino acids for use as a medicament.
- 30 11. A non-naturally occurring polypeptide as defined in any one of claims 1 to 6 for use in a method of diagnosis.

-36-

12. A non-naturally occurring polypeptide as defined in any one of claims 1 to 6 but wherein the length of the polypeptides is between 6 and 60 amino acids for use in a method of diagnosis.
13. Use of a random peptide library according to claim 7 for the selection of a peptide
5 specifically binding to an agent.
14. Use of a random peptide library according to claim 7 for the selection of a peptide specifically binding to an agent but wherein the length of the polypeptides is between 6 and 60 amino acids.

DRAWINGS

Fig. 1

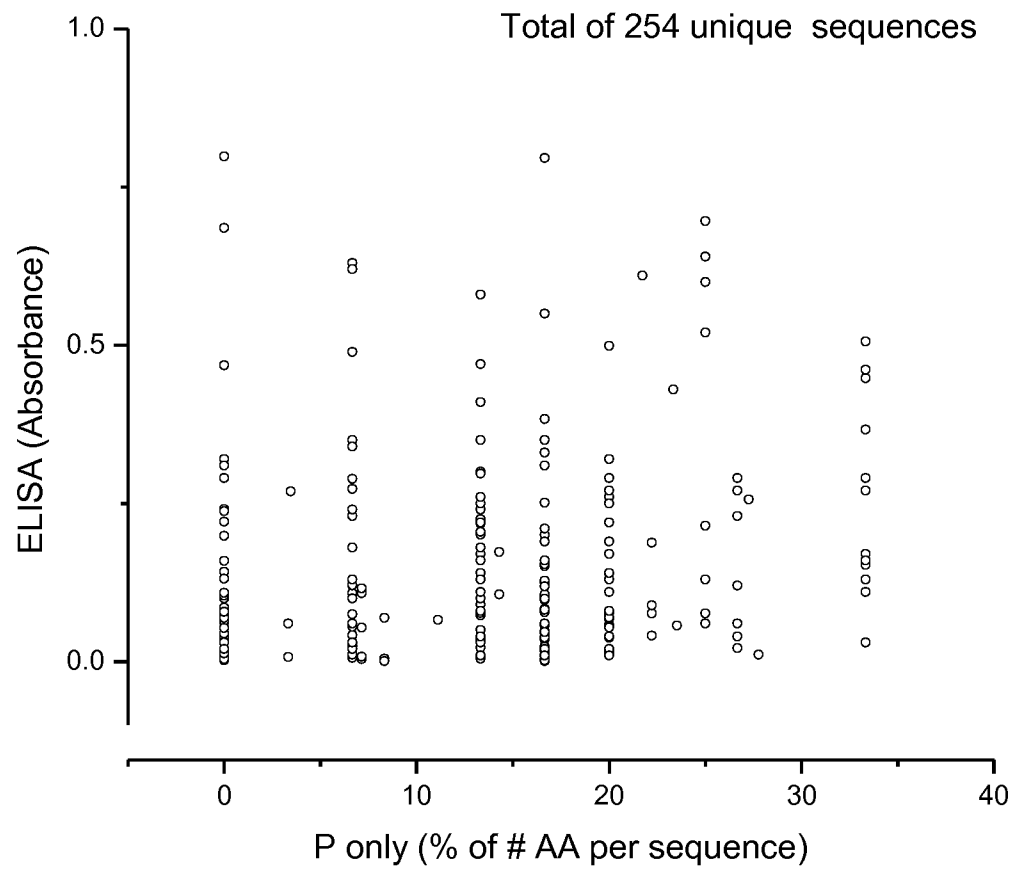


Fig. 2

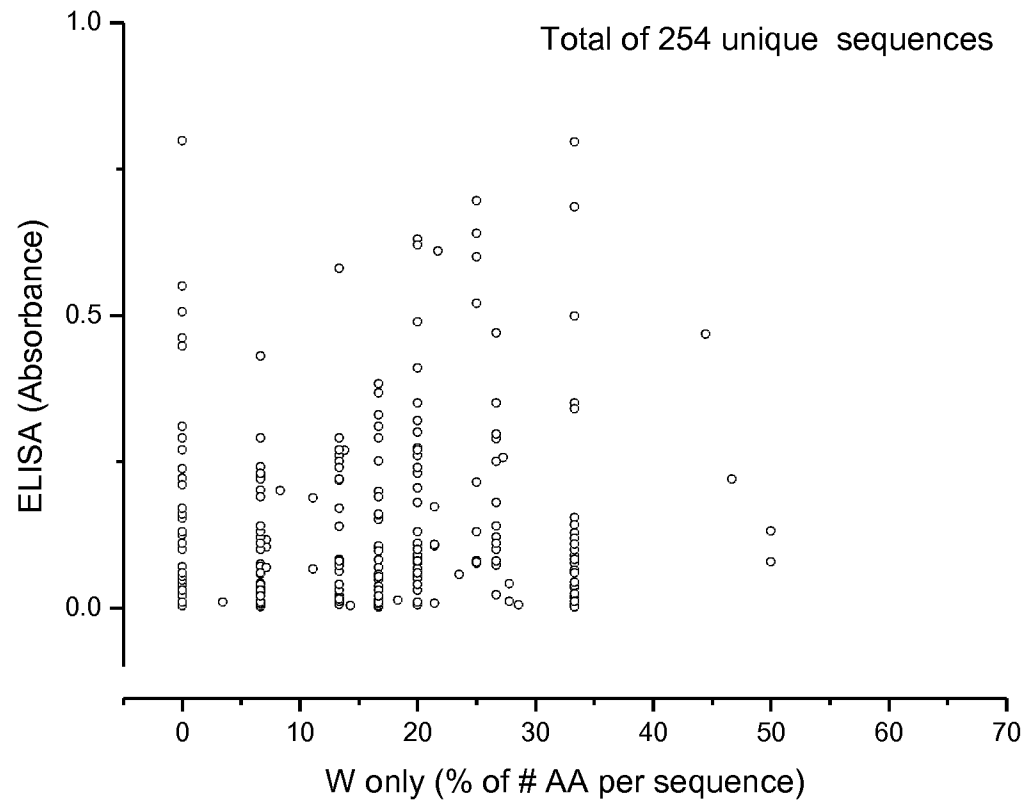


Fig. 3

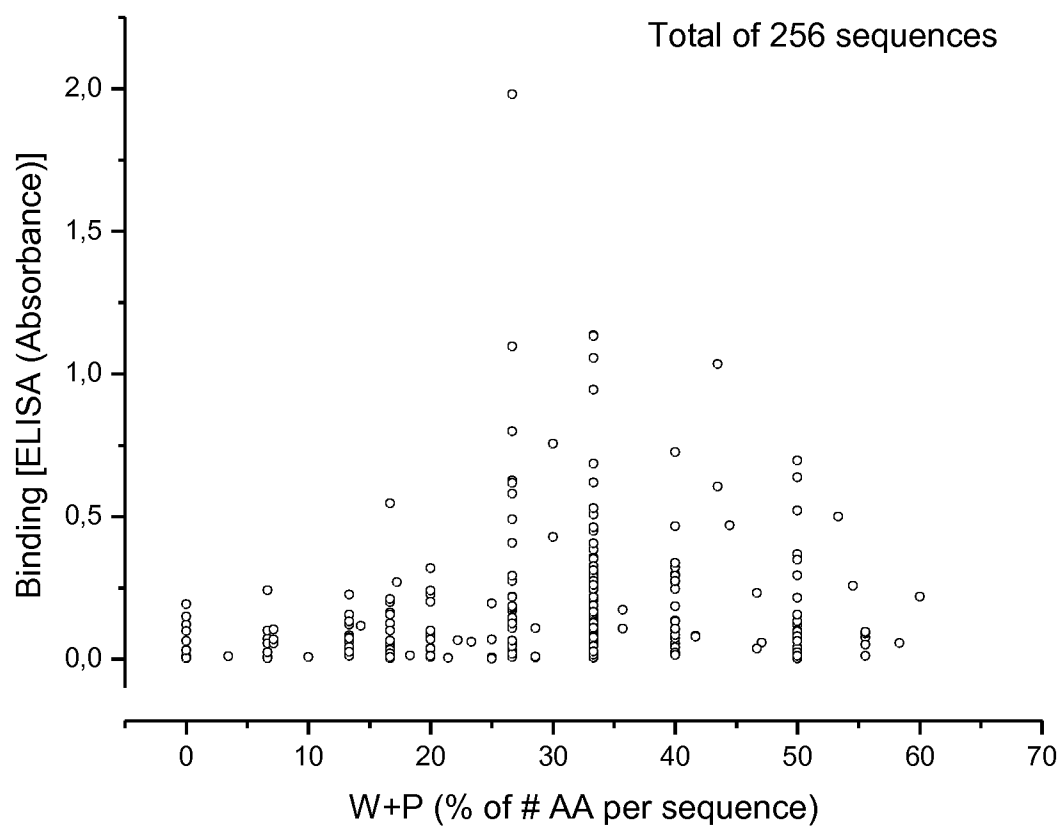


Fig. 4A

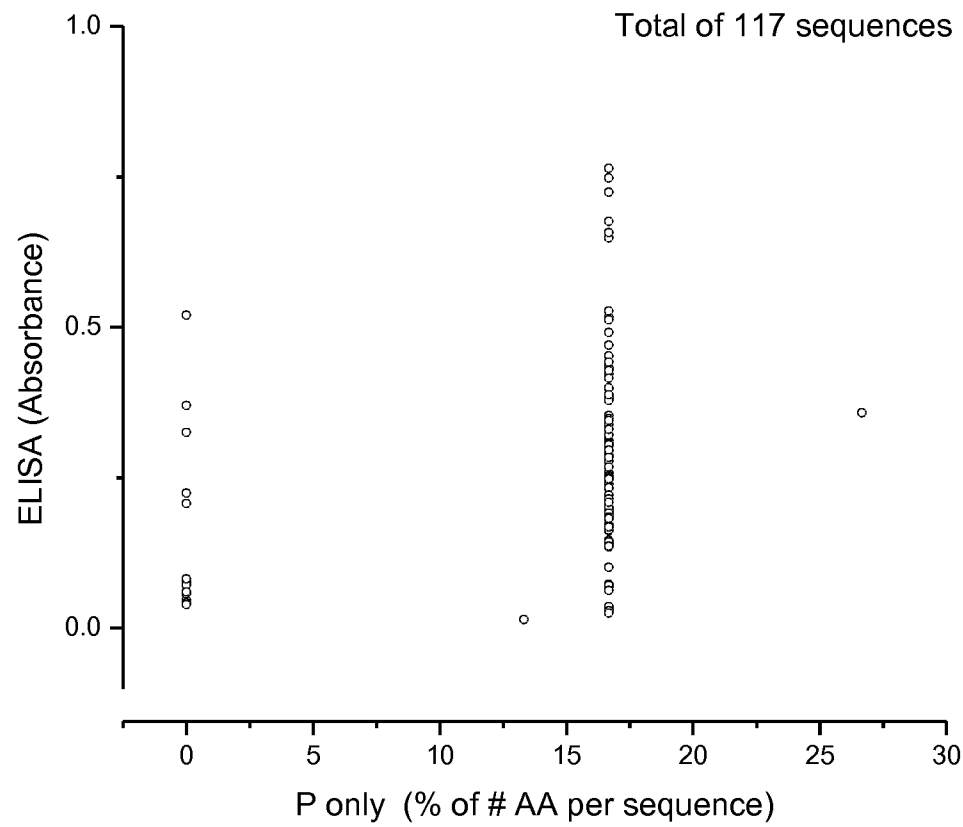


Fig. 4B

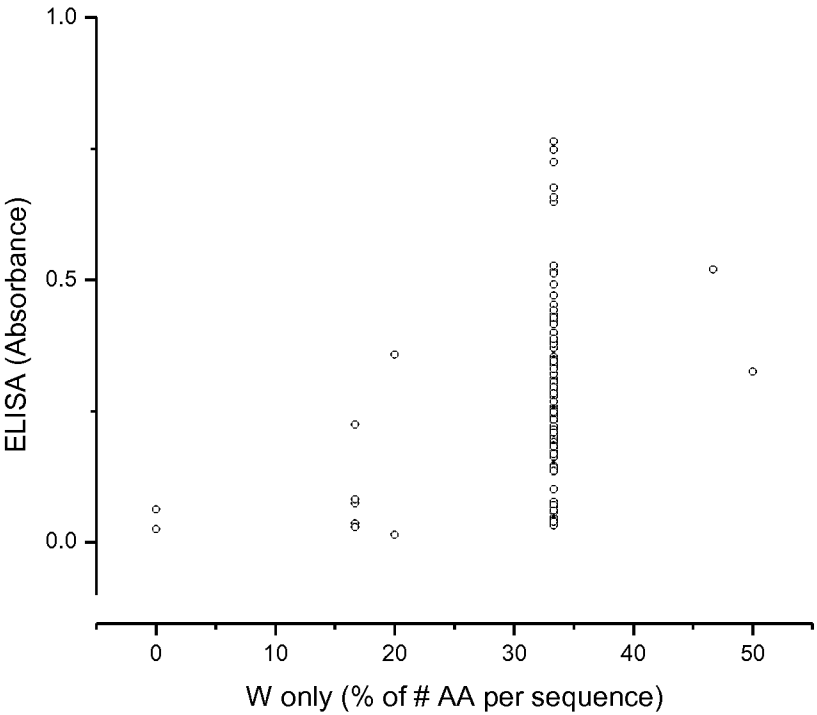


Fig. 4C

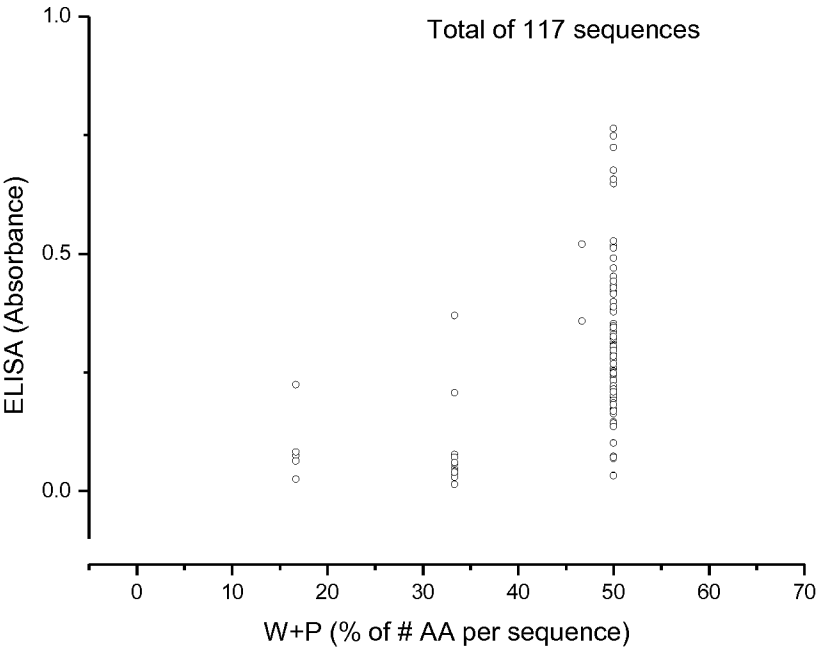


Fig. 5

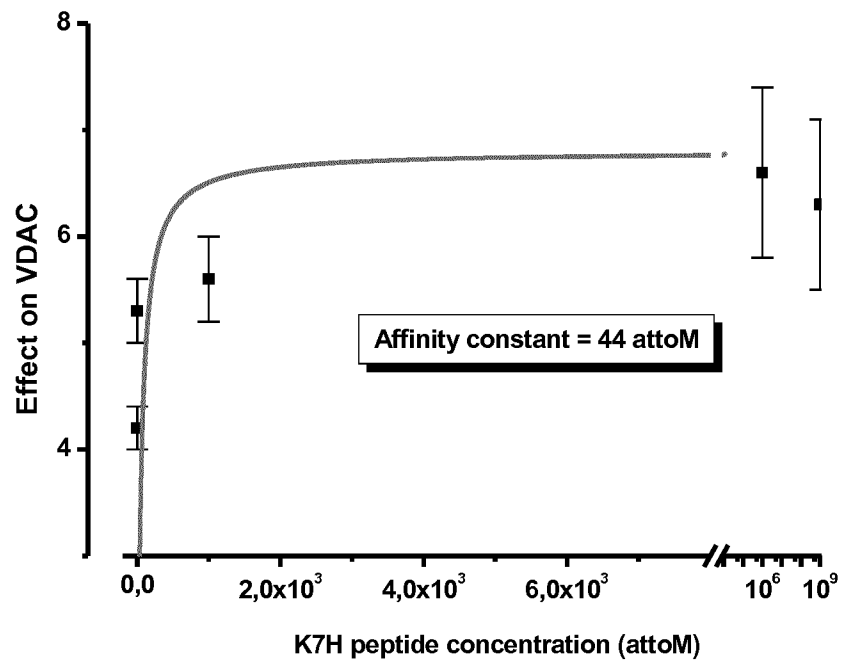


Fig. 6

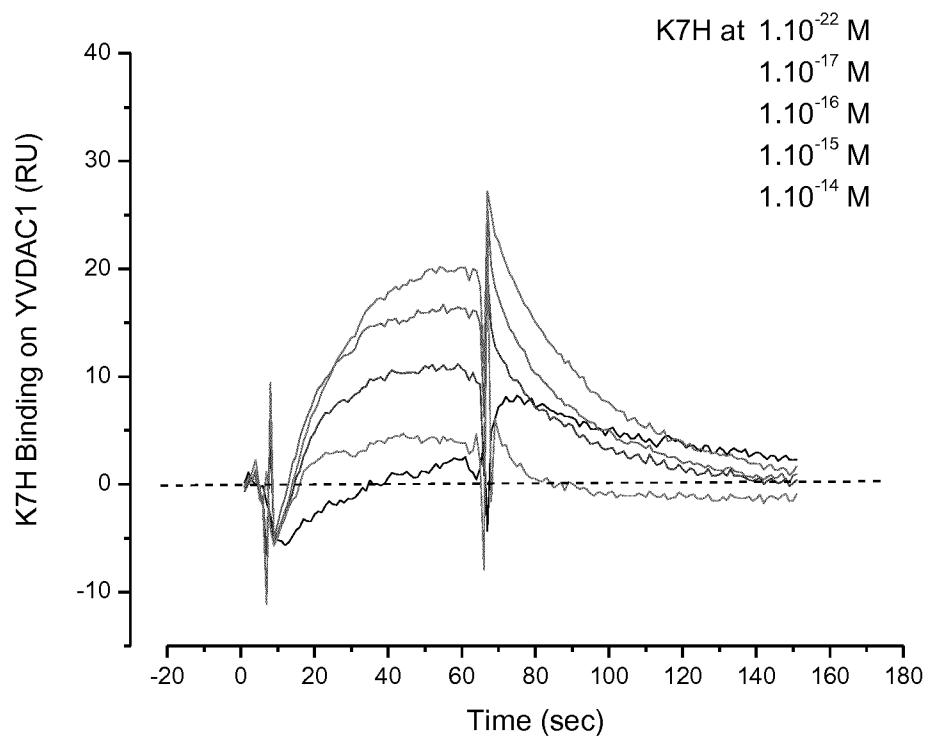


Fig. 7A

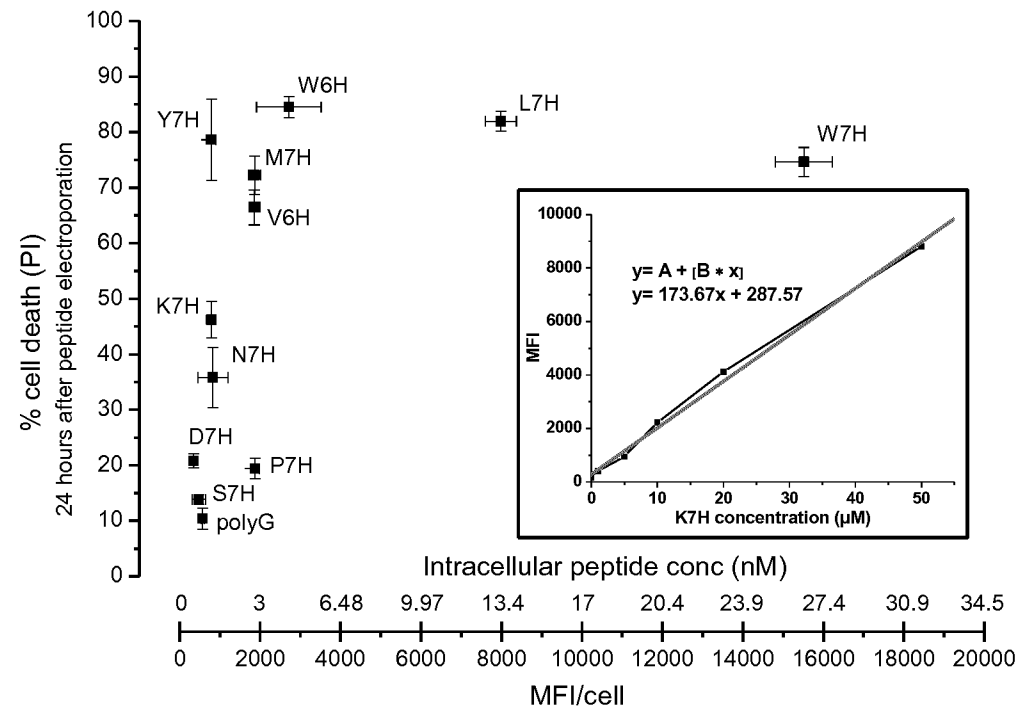


Fig. 7B

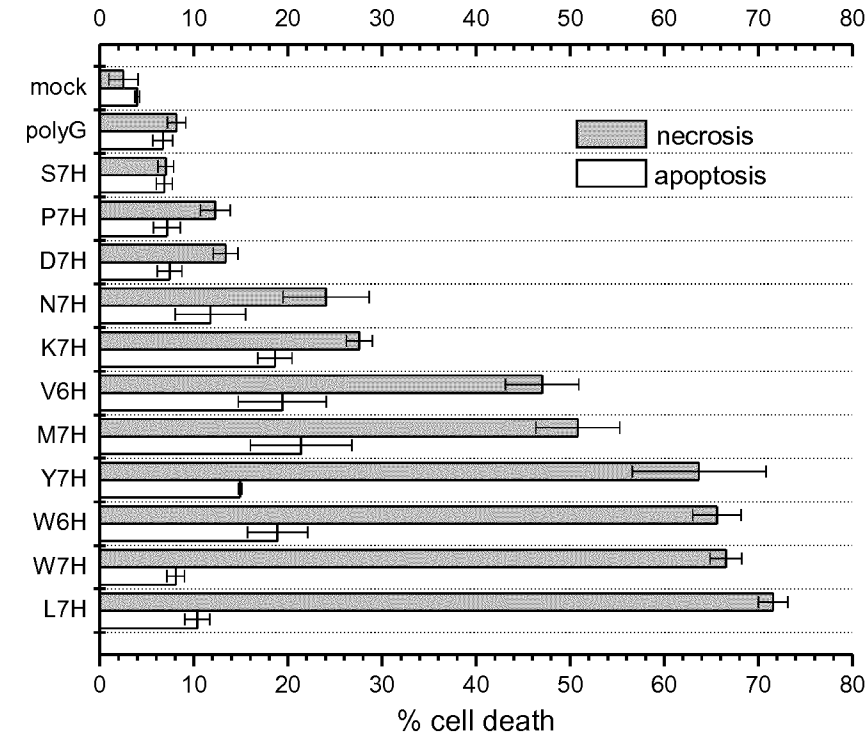
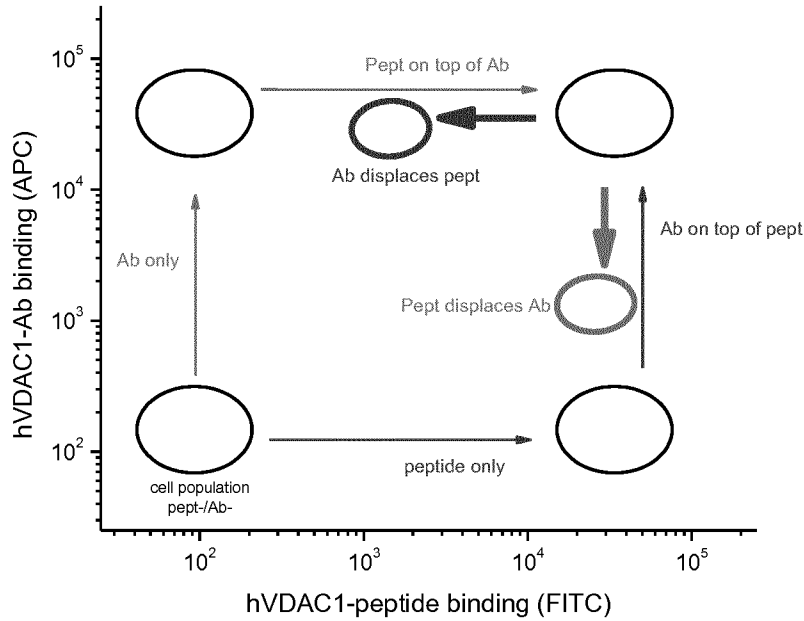


Fig. 8

A



B

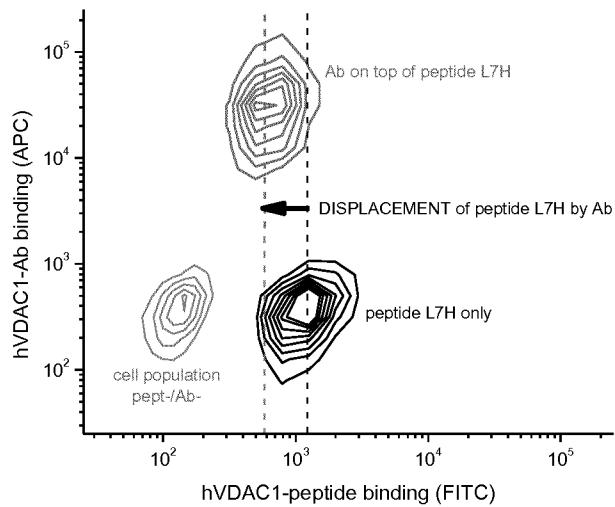
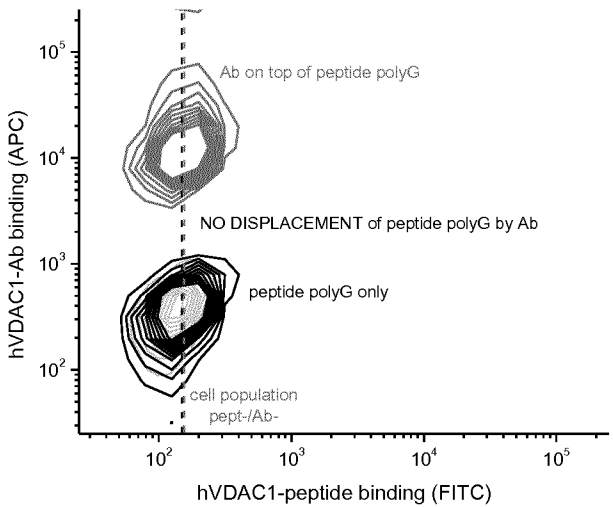


Fig. 8 continued

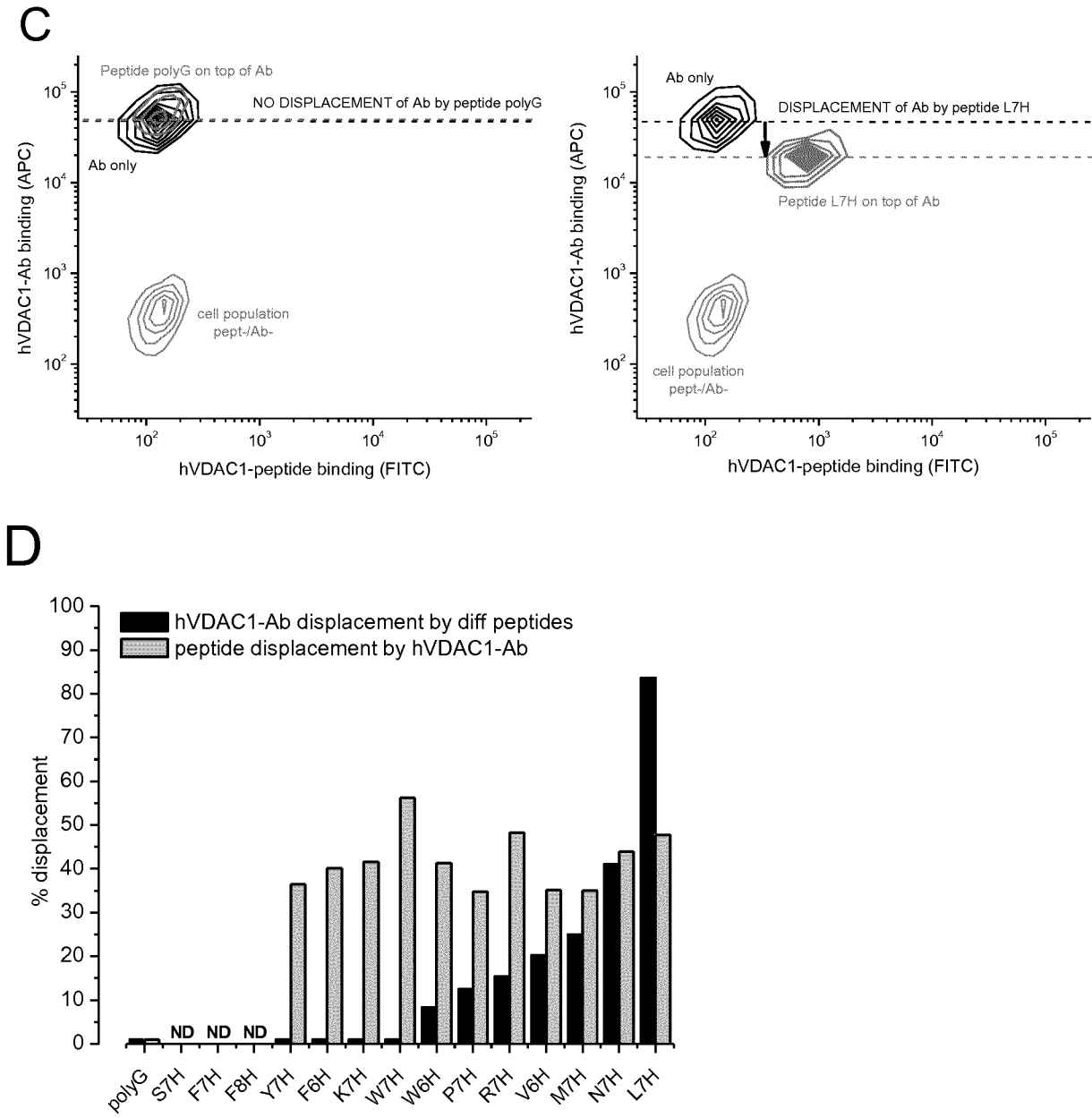


Fig. 9

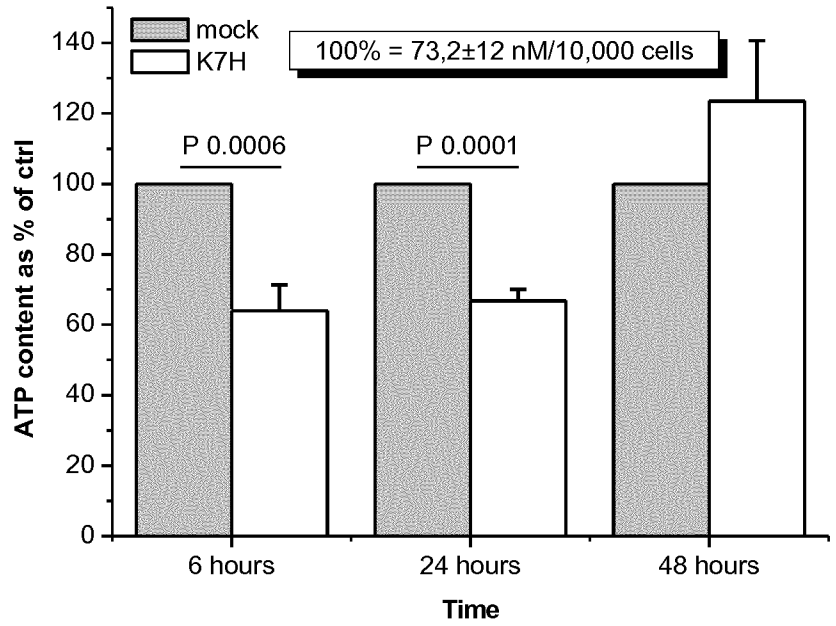


Fig. 10

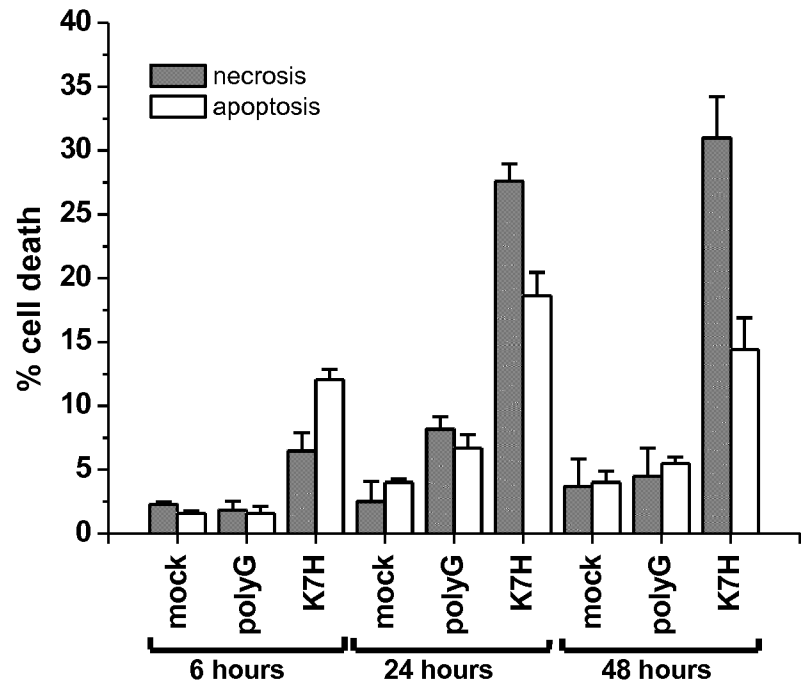


Fig. 11

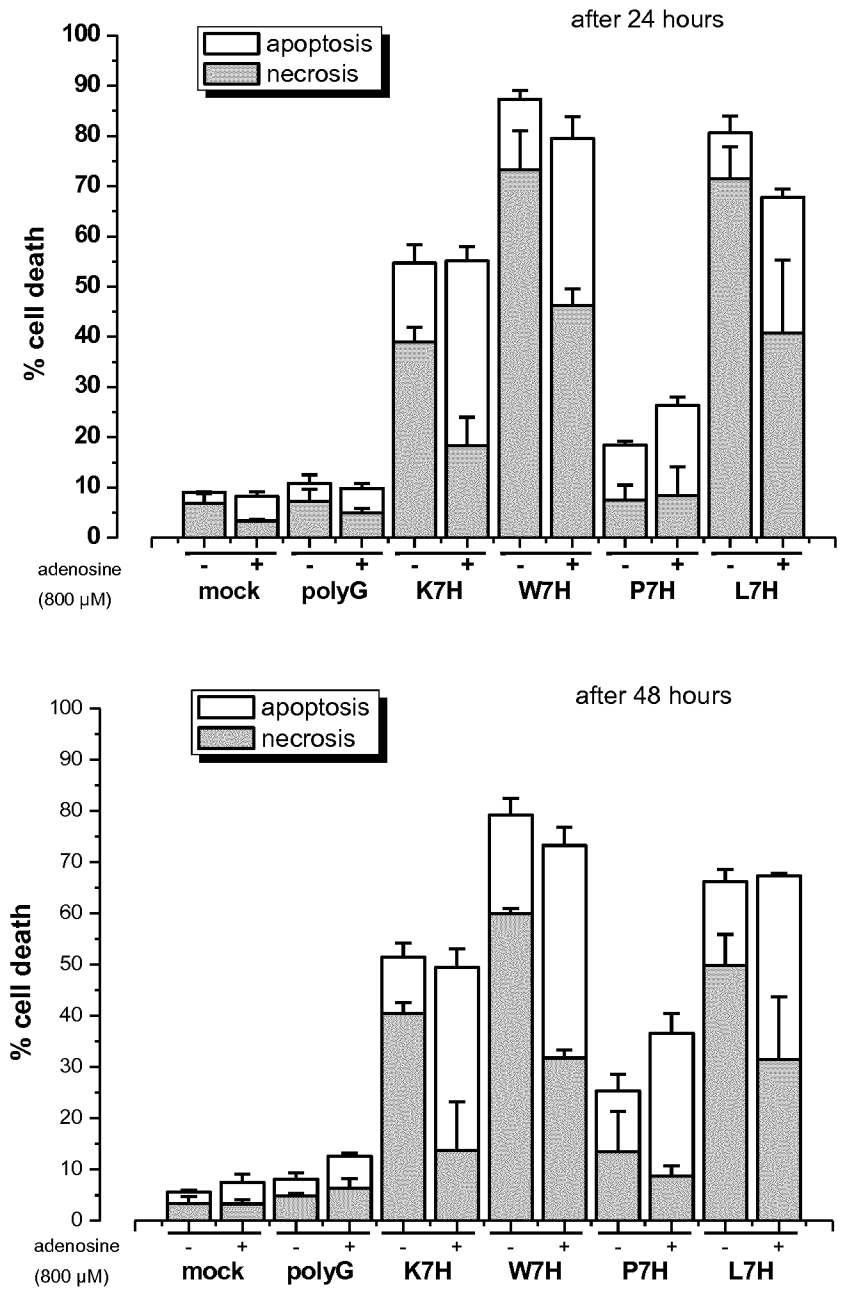
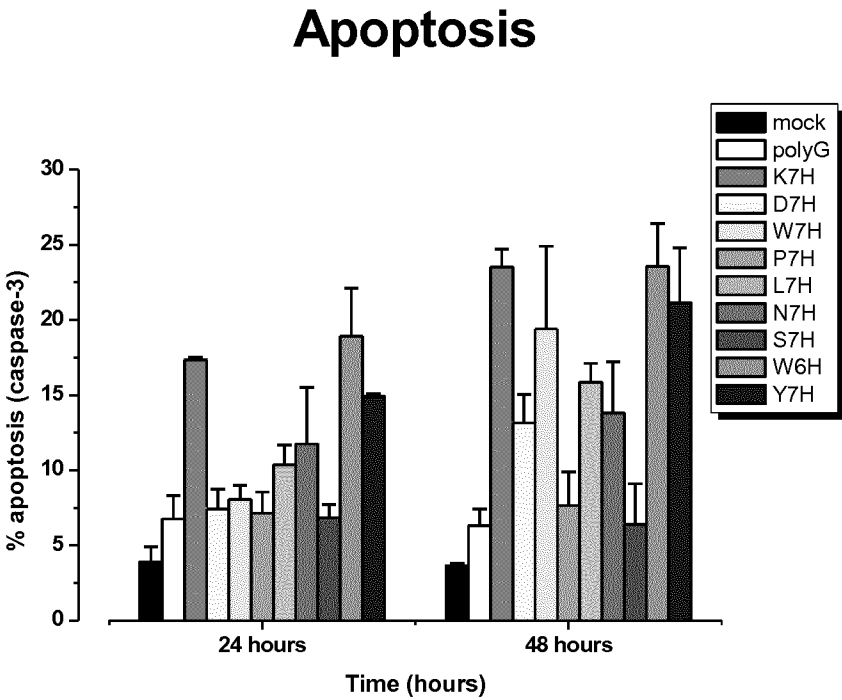
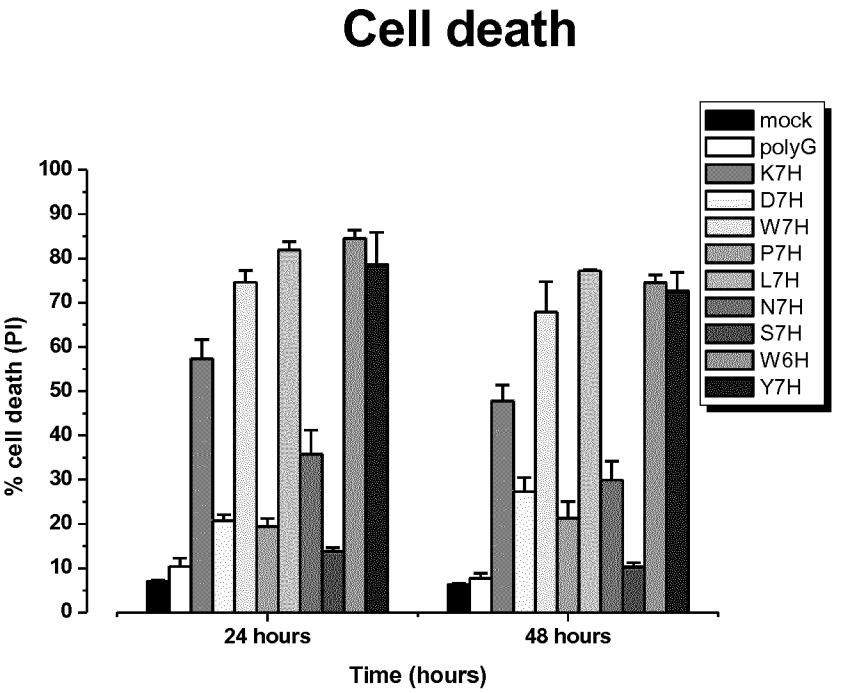


Fig. 12



polyG: GGGGGG
K7H: WSGKPAT
D7H: LDQERP
W7H: WRIWPT (x2)
P7H: KPPHKT (x4)

L7H: FGTWATTYLWHSPTL
N7H: NYLAGAM
S7H: SGSLAG
W6H: WRIWPT
Y7H: PSAYLVPHAPWQGSV (x2)

Fig. 13A

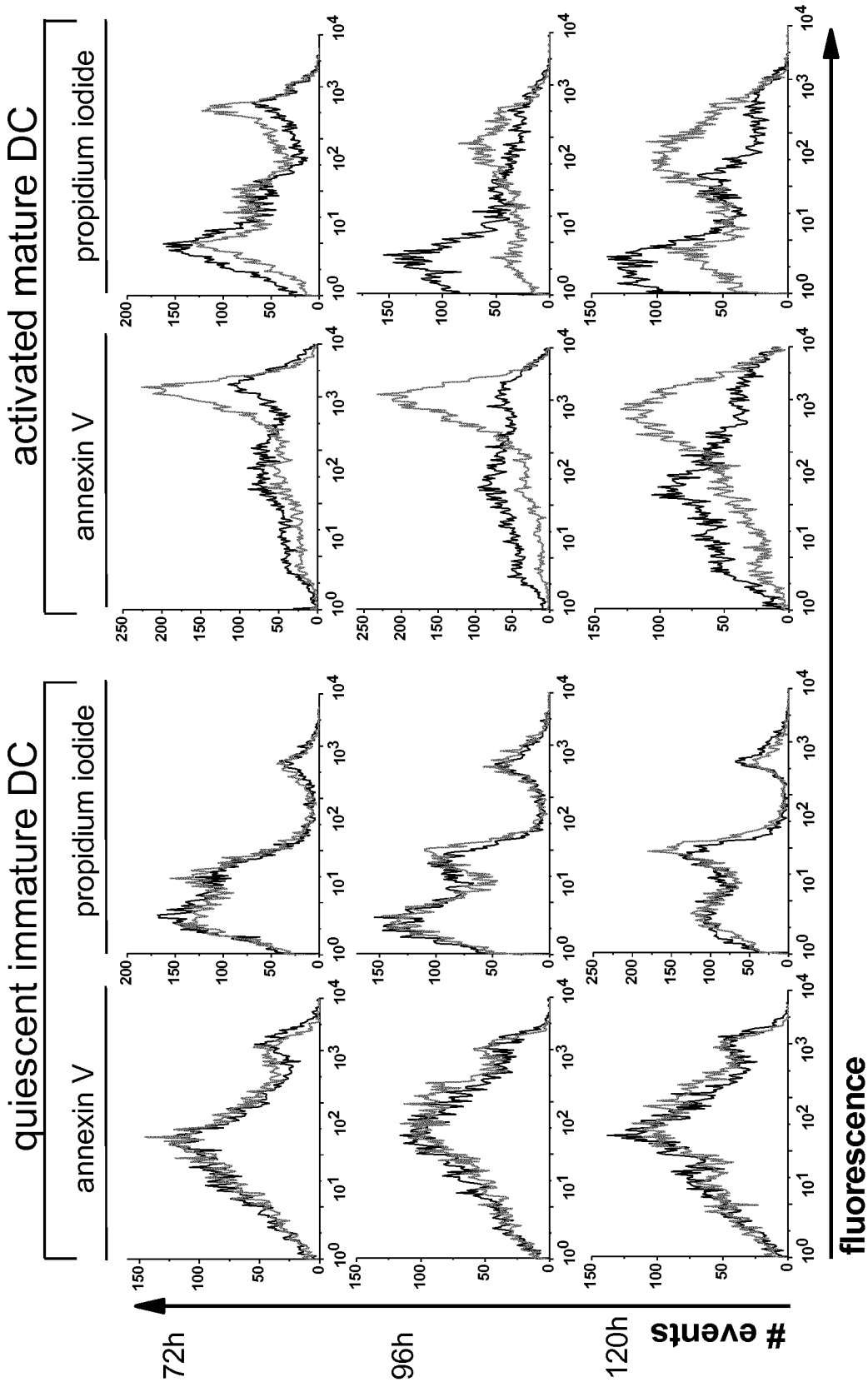


Fig. 13B

1. Erythropoietin

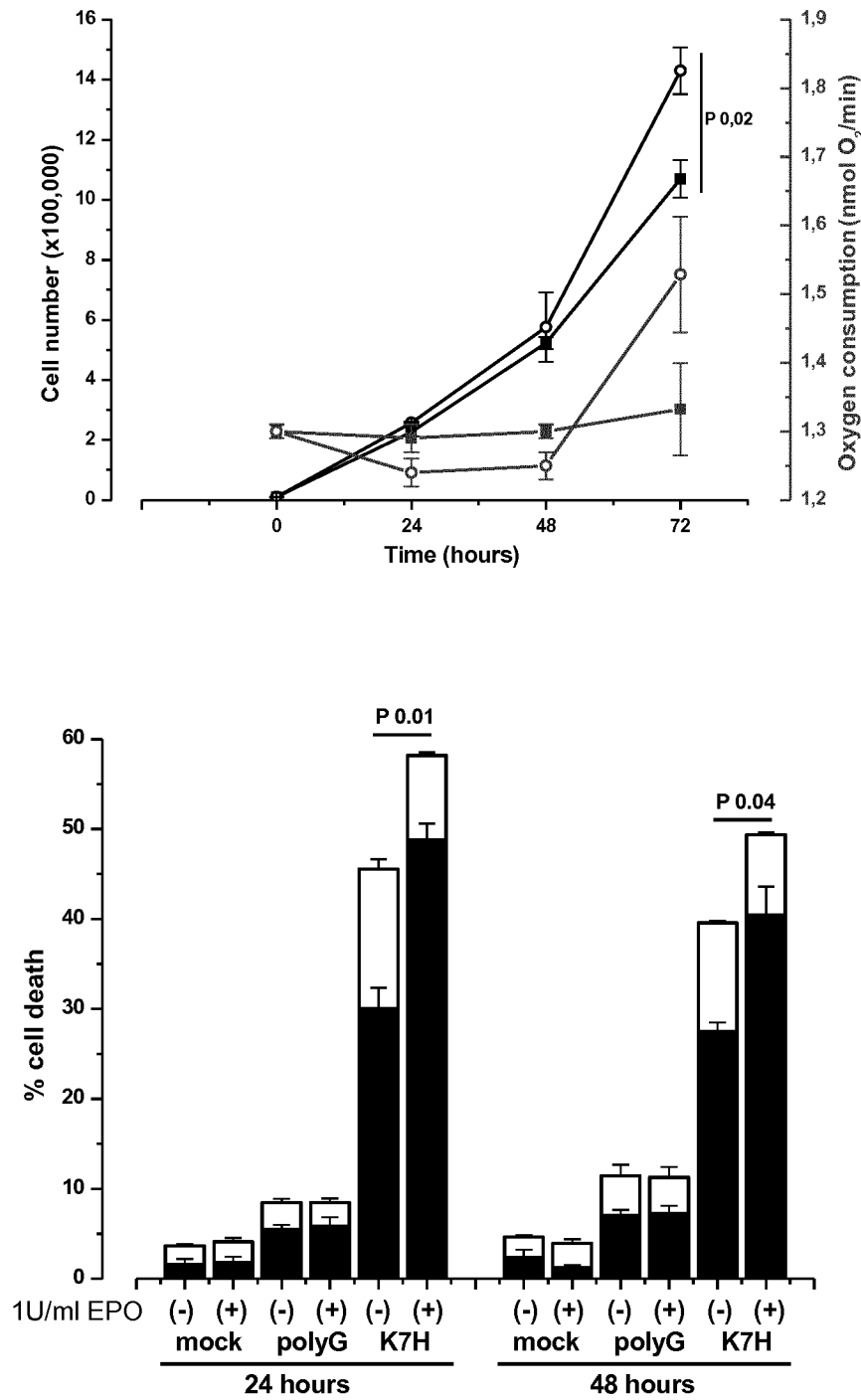


Fig. 14

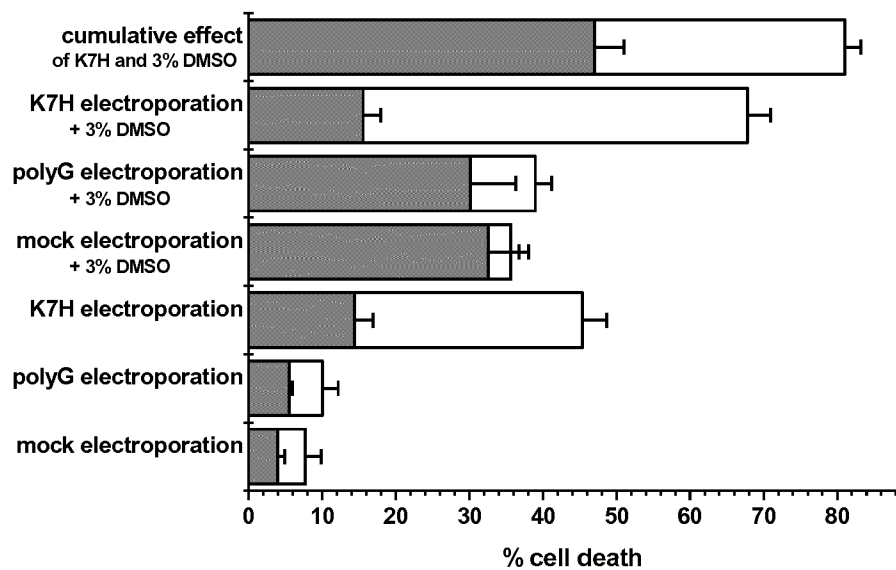


Fig. 15

