Chemically synthesized peptide libraries as a new source of BBB shuttles. Use of mass spectrometry for peptide identification.

B. Guixer[a], X. Arroyo[a,b], I. Belda[b], E. Sabidó[c,d], M. Teixidó[*a], E. Giralt[*a,e].

a) Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Baldiri Reixac 10, 08028 Barcelona, Spain.

b) Intelligent Pharma, Baldiri Reixac 4, 08028 Barcelona, Spain.

c) Proteomics Unit, Centre de Regulació Genòmica (CRG), Dr. Aiguader 88, 08003 Barcelona, Spain.

d) Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, 08003 Barcelona, Spain.

e) Department of Organic Chemistry, University of Barcelona, Martí Franquès 1-11, 08028 Barcelona, Spain.

Corresponding authors (*):

Dr. Meritxell Teixidó and Prof. Ernest Giralt, Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Baldiri Reixac 10, 08028 Barcelona, Spain. Tel.: +34 93 4037125, Fax: +34 93 4037126.

E-mail: ernest.giralt@irbbarcelona.org, meritxell.teixido@irbbarcelona.org

ABSTRACT:

The blood-brain barrier (BBB) is a biological barrier that protects the brain from neurotoxic agents and regulates the influx and efflux of molecules required for its correct function. This stringent regulation hampers the passage of brain parenchyma-targeting drugs across the BBB. BBB shuttles have been proposed as a way to overcome this hurdle since these peptides can not only cross the BBB but also carry molecules which would otherwise be unable to cross the barrier unaided.

Here we developed a new high-throughput screening methodology to identify new peptide BBB shuttles in a broadly unexplored chemical space. By introducing D-amino acids, this approach screens only protease-resistant peptides. This methodology combines combinatorial chemistry for peptide library synthesis, *in vitro* models mimicking the BBB for library evaluation, and state-of-the-art mass spectrometry techniques to identify those peptides able to cross the *in vitro* assays. BBB shuttle synthesis was performed by the mix-and-split technique to generate a library based on the following: Ac-D-Arg-XXXXX-NH_2, where X were: D-Ala (a), D-Arg (r), D-Ile (i), D-Glu (e), D-Ser (s), D-Trp (w), or D-Pro (p). The assays used comprised the *in vitro* cell-based BBB assay (mimicking both active and passive transport) and the PAMPA (mimicking only passive diffusion). The identification of candidates was determined using a two-step mass spectrometry approach combining LTQ-Orbitrap and Q-trap mass spectrometers.

Identified sequences were postulated to cross the BBB models. We hypothesized that some sequences cross the BBB through passive diffusion mechanisms and others through other mechanisms, including paracellular flux and active transport.

These results provide a new set of BBB shuttle peptide families. Furthermore, the methodology described is proposed as a consistent approach to search for protease-resistant therapeutic peptides.

KEYWORDS:

Blood-brain barrier (BBB), mass spectrometry (MS), LTQ-Orbitrap, Q-trap, single reaction monitoring (SRM), *in vitro* cell-based BBB model, parallel artificial membrane permeability assay (PAMPA), mix-and-split, solid-phase peptide synthesis (SPPS).

Introduction

An estimated 1.5 billion people suffer from CNS diseases worldwide. To tackle such diseases, there is an urgent need not only for new therapeutics but also for systems that efficiently deliver therapeutics to their site of action. The blood-brain barrier (BBB) is located in the brain capillaries, and it is formed by endothelial cells, highly influenced by astrocytes, perycites and other glial cells. Features such as tight junctions , few endocytotic vesicles and high metabolic activity allows the BBB to control the passage of molecules from the blood stream to the brain parenchyma, thus protecting the brain from potential neurotoxic substances circulating in the blood.

Two groups of transport mechanisms take place in the BBB: passive transport mechanisms (i.e. paracellular flux and passive diffusion) and active transport mechanisms (including receptor-mediated transcytosis, carrier-mediated transcytosis and adsorptive-mediated transcytosis [1]).

The effectiveness of drugs for CNS diseases depends not only on their therapeutic effect but also on efficient delivery to their site of action. To this purpose, invasive [2,3] and pseudo-invasive [4] strategies are highly discouraged due to sever secondary effects and sequelae. A promising approach to deliver drugs to the brain is the use of BBB shuttles, also called the Trojan horse strategy [5,6], through intravenous injection. This strategy involves linking a drug unable to cross the BBB unaided to a BBB shuttle—a molecule that can cross this barrier. Ideally, this construct (BBB shuttle–drug) should cross the BBB together, thereby facilitating the uptake of the drug into the brain parenchyma. Ideally, a universal BBB shuttle should assist the passage of any molecule attached to it. Since passive diffusion relies on physicochemical properties, linkage of a drug (henceforth cargo) could dramatically affect shuttle efficiency. Although successful and promising passive shuttles have been described [7-11], BBB shuttles using active transport mechanisms are the molecules of choice to launch this shuttle strategy. Targeting highly expressed BBB receptors such as transferrin, lactoferrin, insulin, leptin receptors, and the LDL-receptor related protein 1 and 2 with the aim to achieve drug transport across the BBB by means of receptor-mediated transcytosis might be the most effective strategy.

Historically, peptides have been overlooked as effective therapeutics mainly because of their peptidase lability, which basically ruled out their use as efficient therapeutics or part of therapeutic constructs. However, in recent years it has been demonstrated that peptides are a feasible alternative, being especially suited to mimic and/or adapt to active sites and thus effectively match the pockets of receptors and transporters. Specifically focusing on BBB shuttles, peptides have emerged as a consistent alternative [12-15]. A variety of peptides have been described to be efficient BBB shuttles. Angiopep2 has reached Phase II in clinical trials [12]. Other examples of molecules with potential to become BBB shuttles are SynB1, SynB3, penetratin and TAT peptides, as well as THRre-peptide and apamin derivatives [13,14,16-18].

A variety of strategies can be envisaged to discover and/or design novel BBB shuttles. From natural products or chemically modified molecules to high-

throughput screening of molecular libraries, including chemically synthesized strategies or display technologies [19]. To the best of our knowledge, to date, there are no examples of peptide BBB shuttles designed through a structure-based approach.

Here we propose to search for water-soluble and protease-resistant BBB-shuttle peptides using a high-throughput screening methodology that combines peptide combinatorial chemistry, *in vitro* cellular BBB models, and state-of-the-art mass spectrometry.

Results and discussion

Peptide library design and synthesis

A mix-and-split method was used to synthesize the peptide library [20]. This approach enables the synthesis of peptide mixtures comprising permutations of amino acids. The number of amino acid residues (*m*) in which variability is introduced in each peptide within the library and amino acid variability (*r*) on each peptide residue determine the total number of peptides in the library, $R = r^m$. To assess library size, combinations of *r* and *m* values were evaluated. Eventually, *m = 5* and *r = 7 values* leading to a library comprising 16.807 peptides were selected to synthesize our peptide library (see Section S1 in Supporting Information). On the basis of physicochemical and statistical parameters, as well as on synthetic performance in order to minimize coupling side reactions, a set of seven amino acids comprising D-tryptophan, D-arginine, D-alanine, D-glutamic acid, D-serine, D-isoleucine and D-proline was selected (**Table 1**).

Tryptophan is a fluorescent amino acid that plays a key role in many potential hot spots for protein-surface molecular recognition. Arginine is also described to have many potential interactions (electrostatic, H-bond) and is relevant for molecular recognition. Moreover, with a pI = 10.76, arginine has a positive charge in physiological pH. It is also described as a crucial residue for peptides able to cross membranes. Alanine is an abundant residue in natural systems. Since this amino acid has a small side chain, it can act as a spacer for distinct interaction sites on peptides. Glutamic acid contributes to negative charges and shows better synthetic performance than aspartic acid as it does not present side reactions leading to aspartimide formation. Serine is a polar non-charged amino acid that shows better performance on SPPS than β-branched threonine. Isoleucine is an aliphatic β-branched amino acid with a bulky side chain. It has a high tendency to form a β-sheet structure and has a higher presence than valine or leucine in active sites. Finally, proline is the only natural *N*-alkyl amino acid. It facilitates β-turns, and poly-Pro leads to particular structures. It has been described to play a crucial role in hydrophobic-hydrophobic interactions.

The library design included the acetylation of the *N*-terminus and amidation of the *C*-terminus. Since these *N*- and *C*-terminus positions are regarded as potentially suitable sites to link eventual therapeutic cargos, acetyl and amide groups were used as the simplest cargo moieties. This approach allows to high-throughput screen in physicochemical conditions that closely resemble those of

the final shuttle-moiety construct. In addition, we introduced a D-Arg moiety at the *N*-terminus of the 5 variable amino acid positions in the peptides to complement the library design, thus leading to a hexapeptide library. On the one hand, this approach ensures at least one position for charge stabilization for MS ionization. On the other hand, this singular Ac-D-Arg- fingerprint facilitates the identification of peptides within complex mixtures comprising endogenous peptides from *in vitro* cellular assays and *in vivo* assays. Library synthesis was performed by the mix-and-split methodology in SPPS. The synthetic workflow and a summary of the library design are provided in **Figure 1**,

Library synthesis was validated through two methods. First, before whole library cleavage, single peptide analysis was performed by cleavage of single beads. Mix-and-split synthesis led to one-bead-one-compound (OBOC) libraries– meaning that each bead contained several copies of the same peptide sequence. By cleaving peptides carried by single beads and the subsequent analysis by MALDI-TOF, we found that the peptides had the expected pattern (m = 5 + D-Arg + CH_3COO- and composed only by the 7 amino acids used in the synthesis; homogeneous, single sequences were found, not presenting deletions). Second, once the library was cleaved from the resin, high-resolution mass spectrometry analysis was performed by direct infusion of the library on the FT-ICR mass spectrometer [23]. Quality control of the library synthesis was evaluated by comparison between the theoretically predicted library mass spectrum and the experimentally measured library mass spectrum as shown in **Figure 2**. Deviations such as peptides with a molecular weight lower or larger than that expected from the library design are useful to evaluate library quality. Both techniques are complementary tools for the quality control of combinatorial mix-and-split peptide libraries.

Library assay using *in vitro* BBB models

Two *in vitro* assays were used to test the Ac-D-Arg-library, namely the parallel artificial membrane permeability assay (PAMPA) and an *in vitro* cell-based BBB assay. PAMPA is an *in vitro* transport assay based on artificial phospholipid membranes that mimics only passive diffusion properties of the BBB (**Figure 3A**).

The PAMPA was run with a saturated solution of Ac-D-Arg-library in MilliQ water (neither commercial pION PAMPA buffer nor co-solvents were used). Samples were sonicated to ensure concentrations of peptides as high as possible, close to saturation, and pH was adjusted to 7.0. To eliminate non-water soluble peptides, samples were centrifuged and precipitate was removed. This step serves as a first screening stage to avoid selecting peptides that will have solubility problems on next stages of shuttle development process. The PAMPA was run for 4 h in triplicate. Acceptor (upper) and donor (lower) compartments and t_0 aliquots were kept for analysis. A blank sample was run in parallel.

In vitro cell-based BBB models are still far from fully mimicking the *in vivo* BBB, which is influenced by many other stimuli than those present *in vitro*. Several models based on cell co-cultures have been proposed with a variety of combinations of endothelial cells and astrocytes from distinct animals. These models have shown a significant correlation between *in vitro* and *in vivo* BBB permeability, although do not fully mimic the BBB. We selected the model

based on a co-culture of bovine brain endothelial cells and rat astrocytes, as shown in **Figure 3B**, as it is one of the most accepted and validated BBB models described to date [24,25]. This model mimics both passive diffusion properties and active transport mechanisms.

The in vitro cell-based BBB assay was run with a , 1:10 dilution of the saturated solution of Ac-D-Arg-library in Ringer/HEPES buffer (buffer required for cell-based BBB model). Dilution was applied to avoid high peptide concentrations to minimize cell stress and disruption of the *in vitro* BBB. The pH was adjusted to 7.0. To eliminate non-water soluble peptides, samples were centrifuged and precipitate was removed. Lucifer yellow was added as an internal standard to each well. The *in vitro* cell model was run for 2 h at 37ºC in triplicate. Donor (apical) and acceptor (baso-lateral) compartments and t_0 aliquot were kept for analysis. TEER measures (prior to the assay) of co-cultured cell layers of the *in vitro* cell-based assay were 125 ± 18 $\Omega \cdot cm^2$. A blank sample was run in parallel.

Peptide identification by mass spectrometry

The greatest challenge of this approach to discover novel BBB shuttles is the identification of peptides able to cross the *in vitro* BBB assays. We were confronted with a mixture of extremely similar peptides, especially those sequences with the same composition and that share similar physicochemical properties. Such similarity therefore hinders their differentiation. This complex peptide mixture can be analyzed on the basis of proteomic analysis, although a combination of mass spectrometry techniques should be considered. Mass spectrometry techniques were preferred over others such as pool sequencing method by Edman degradation since allow higher versatility on the screened libraries.

While a stringent assay such as the *in vitro* cell-based BBB assay enormously reduces the number of peptides to be analyzed in the acceptor (baso-lateral) compartment, coupled chromatographic and mass spectrometric techniques are still required to fractionate the complexity of the acceptor mixture.

However, from such similar mixtures, high retention time differences to separate peptides by chromatographic techniques are not expected to be sufficient, even working with isocratic gradients. In many cases, peptides will be simultaneously ionized and analyzed in the mass spectrometer. Mass spectrometry techniques provide *m/z* information about the molecules ionized, thus identifying all those with characteristic *m/z* signals in the peptide ion spectra, also called MS1. The high resolution and precision (10 ppm) of this advanced technique allows the unequivocal assignation of signals to molecular compositions.

Molecular masses represent mostly a single amino acid composition, but in some cases combinations of pairs, trios or larger groups of amino acids may have the same overall mass and thus not be distinguishable with 10 ppm precision. In our library, the following combinations of amino acids led to indistinguishable compositions (10 ppm precision): (D-Ala)2-(D-Glu)1 and (D-Ser)2-(D-Pro)1, 0.04 ppm; and (D-Glu)2-(D-Ile)2 and (D-Ala)2-(D-Arg)1-(D-Trp)1, 2.75 ppm. Therefore it is critical not to assume that the same molecular mass denotes a single amino acid composition (from now on we will use the term peptide family to denote a single amino acid composition).

Indeed, each peptide family contains a diverse number of specific peptide sequences (family members) depending on the amino acid distribution within the sequence: from a single member where all amino acids are the same to 120 members when each member differs (**Table 2**).

Given that the members of a family share the same m/z and specific family members cannot be identified using solely MS1 spectrometric techniques, we designed an analytical MS strategy comprising two mass spectrometric levels.

The first level of identification consisted of injecting the sample into an LTQ-Orbitrap (coupled to LC) mass spectrometer. This first step focuses only on MS1 spectrum data and provides a clear picture of all precursor ions (in this case they can also be called peptide ions or molecular ions because no later fragmentation is performed) present in the sample.

However, MS1 spectrometric techniques alone are unable to identify specific peptide sequences (family members). In this regard, MS1 spectra information allows the identification of specific molecular compositions (peptide families) in the sample. To this purpose, we developed a computer program called *Bibliopepfinder.* We used this tool to sort LTQ-Orbitrap output data (MS1 spectra peaks) that match peptide families of the synthesized library (see Section 2 in Supporting Information).

This first identification level resulted in a LTQ-Orbitrap MS1 precursor ion list, which contained information only on the composition of peptides present in the sample, as shown in **Figure 4**. At this point, we focused our efforts on targeting a few selected families since the second identification level cannot be performed with diverse families at the same time.

In the second identification level, we used single reaction monitoring (SRM) technology, now extensively used for targeted proteomic analysis [26]. For this purpose, triple quadrupole (QqQ) equipment was used, taking advantage of its high sensitivity. The combination of m/z settings on both the first and third quadrupoles of QqQ is referred to as transition. To filter the desired peptide parent ion, the m/z value of the first quadrupole is determined by the mass and the most abundant charge state of the targeted peptide. In the third quadrupole, one fragment ion (or daughter ion) is selected and thus a transition is defined.

In contrast to a classical proteomic SRM approach—which monitors only a reduced set of transitions from a selected number of peptides—our analysis aimed to explore all the sequences from selected compositions (peptide families). For this purpose, we monitored the complete transition set of each sequence of the selected peptide composition.

Transitions corresponding to all *b* and *y* ions from singly and doubly charged precursor ions were monitored. When enough transitions that define a sequence are found in the resulting spectrum, we can assure that this specific peptide sequence is present within the *m/z* LTQ-Orbitrap MS1 signal, and thus that it crosses the BBB *in vitro* assay.

The QqQ equipment allows the monitoring of around 120 transitions per injection (considering a dwell time of 20 ms) [27]. Thereafter, few families can

be screened because there is not enough sample for recurrent injections (approximately 4-5 injections were performed per sample). For this reason, it is of paramount importance to have a rational selection criterion to choose the most promising families from the MS1 LTQ-Orbitrap list.

To perform this selection, we formulated four requirements to be fulfilled by all candidate families in order to pass through to the second identification level (**Table 3**). First, the family has to be present in all *in vitro* assay triplicates. Second, their molecular peak must fulfil typical mass spectrometry fingerprint profile of peptide species in the MS1 LTQ-Orbitrap spectra: molecular and isotopic peaks must have a consistent and peptide-like profile (considering peptide masses between 500 a.m.u. and 1200 a.m.u.) [28] and, furthermore, it must be confirmed that the molecular peak is indeed a molecular peak and not an isotopic peak from another signal (i.e. peptide specie). Third, consistency is primed among triplicates, so as to prioritize peaks that are consistently more intense. This makes the choice much more robust and avoids possible false positive results. And fourth, families with more than 30 members are directly discarded.

Thanks to the combined use of the PAMPA and *in vitro* cell-based BBB assay, an extra selection criterion was considered. Since the PAMPA mimics only passive diffusion properties of the BBB, all families detected in PAMPA samples were hypothesized to cross the BBB through this mechanism. On the other hand, the *in vitro* cell-based BBB model mimics both active transport and passive diffusion properties of the BBB. It can therefore be hypothesized that families found in both assays cross the BBB by passive diffusion, while those present only in the *in vitro* cell-based BBB model cross by another mechanism (either active transport related or paracellular flux). Thus, two groups of peptides were generated on the basis of hypothetical passive diffusion permeability or other mechanisms of transport.

After assaying the library and previous injection into the mass spectrometer, a sample cleaning protocol was performed. Acceptor well (baso-lateral) and t_0 samples from *in vitro* cell-based BBB assay were desalted with mini-C_{18} columns following a standard desalting protocol. The peptides present in the samples were retained within mini-C_{18} columns and then washed with water to remove salts. The eluted peptides were then used for HPLC-MS analysis. Similarly, acceptor well and t_0 PAMPA samples were cleaned following the same protocol.

Once desalted, the library was injected into the LTQ-Orbitrap equipment. After injection, a list of all peaks and their basic information (*m/z*, intensity, relative intensity and charge) was extracted from equipment software (**Figure 4**).

All data were evaluated with *Bibliopepfinder* to filter peaks that might correspond to peptides present in the Ac-D-Arg-library. Data were analyzed considering an instrumental error of 10 ppm. *Bibliopepfinder* provides output information for all peaks potentially corresponding to peptides belonging to the Ac-D-Arg-library: theoretical mass, experimental mass, amino acid composition (family), matching *m/z* signals, a general family number code and specific family code (e.g. (D-Glu)3-(D-Ala)1-(D-Pro)1-(D-Arg)1 family would appear as e3-a1-

p1-r1 and general family number code would be 3111 referring to the number of distinct amino acids present in the family).

The families selected from PAMPA found in all triplicates were pooled and all replicates were removed. Signals matching library families found in the blank sample (3) were subtracted from the list (**Figure 5A**). Finally, *Bibliopepfinder* assigned 187 families to 169 m/z signals in the LTQ-Orbitrap spectrum. 151 m/z signals matched a single family, while 18 m/z signals were assigned to two families.

Triplicate samples from the *in vitro* cell-based assay analyzed by LTQ-Orbitrap were pooled and blank signals (10) matching possible families from the library were deleted (**Figure 5B**). *Bibliopepfinder* assigned 367 families to 335 m/z signals in the LTQ-Orbitrap spectrum. 304 m/z signals matched a single family, while 30 m/z signals were assigned to two families and in one case to three distinct families. Of the families assigned by *Bibliopepfinder*, 78% were present in all triplicates, while 15% were present in two triplicates and 7% in only one triplicate (**Figure 5C**).

The data from the *in vitro* cell-based BBB model can be interpreted in two ways. First, it shows high reproducibility from well to well in the cell based assay plate. Second, 286 out of 462 possible families in the Ac-D-Arg-library (62%) probably show a certain percentage of paracellular flux because the endothelial cell layer forming the *in vitro* cell-based BBB assay is too loose. This is an intrinsic general concern regarding the *in vitro* cell-based BBB models described in the bibliography. The improvement of this feature is currently one of the hottest topics in the field.

At this point, we applied an extra selection criterion, thus generating two lists of candidate peptides: those postulated to cross the BBB by passive diffusion and those postulated to cross by other mechanisms (including active transport related mechanisms and paracellular flux) (**Figure 5D**).

Of note, 172 (92%) out of 187 families found in PAMPA were also found in the *in vitro* cell-based BBB assay (not necessarily in all triplicates). These data demonstrate the satisfactory performance of this combined assay approach, which shows a high degree of agreement when predicting molecular passive diffusion properties.

The list for “other mechanisms of transport” included signals found in all triplicates of *in vitro* cell-based BBB assay but not in PAMPA. *Bibliopepfinder* assigned 131 families to 119 m/z signals in the LTQ-Orbitrap spectrum. 108 m/z signals matched a single family, while 10 m/z signals were assigned to two families, and in one case to three distinct families.

Selection of candidates to be further studied and validated

We then selected some families from each list following the rational selection criteria stated above. The statements shown in **Table 3** were applied. These criteria led to the selection of a set of peptides to be further studied at the second level of identification: 4 candidates corresponding to the “other mechanisms of transport” list (i3-a1-p1-r1, e2-p2-a1-r1, p2-w2-i1-r1 and i4-p1-

r1), and 3 candidates corresponding to the "passive diffusion transport" list (i3-p2-r1, p3-s2-r1 and a3-r2-p1).

Skyline proteomics software was used to generate all desired transitions from the sequences selected. This approach allows Q-trap to monitor each transition in order to confirm the presence of each monitored sequence. Transitions from both singly and doubly charged precursor ions were monitored. On one hand, for singly charged precursor ions, 10 transitions were explored: b_1^+, b_2^+, b_3^+, b_4^+, b_5^+ and y_5^+, y_4^+, y_3^+, y_2^+, and y_1^+. On the other hand, for doubly charged precursor ions, 20 transitions were followed: b_1^+, b_2^+, b_3^+, b_4^+, b_5^+, y_5^+, y_4^+, y_3^+, y_2^+, y_1^+ and b_1^{++}, b_2^{++}, b_3^{++}, b_4^{++}, b_5^{++}, y_5^{++}, y_4^{++}, y_3^{++}, y_2^{++}, and y_1^{++}.

To validate any sequence, transitions corresponding to the fragmentation of each amide bond must be found. Therefore, b_1 or y_5, b_2 or y_4, b_3 or y_3, b_4 or y_2, and b_5 or y_1 must be found to unequivocally confirm the presence of this sequence in the sample. These signals must not appear in the blank sample injected in QqQ.

In some cases, where there are two or more equal amino acids next to each other, not all transitions need to be found. For example, "riiipa" could be assigned by finding b_1 or y_5, b_4 or y_2, and b_5 or y_1. However, assignments finding fragmentation ions that validate all transitions show greater consistency and robustness.

Q-trap spectra were manually and individually assigned using Skyline Proteomics software. Co-elution of enough transitions (unequivocally describing the screened sequence) in the chromatogram ensures the presence of that particular sequence. **Figure 6** shows, as an example, the co-elution of all transitions defining "raiipi" assignment to the peak at t_r = 16.7 min.

Figure 7 summarizes all the data analyzed from SRM experiments performed in Q-trap for all sequences of the peptide compositions selected. Information is filtered and rated taking into account whether identification evidence is consistent among triplicates, retention time, and precursor ion.

The candidates monitored on the passive diffusion transport list suggest significant differences between sequences from each composition/family. From the i3-p2-r1 family, ripipi, rpiiip and rpipii were found in high accordance and consistency. Highly consistent results were found for rsppsp, rspspp and rssppp from the p3-s2-r1 family, and for raaarp, rraaap and rrapaa from the a3-r2-p1 family.

On the other hand, individual sequence evaluation from the "other mechanisms of transport" list showed a higher presence of sequences representing a specific monitored family/composition. Importantly, some of them showed high consistency over triplicates. In fact, i4-p1-r1 was represented by all five peptides sequences (riiiip, riiipi, riipii, ripiii, rpiiii), while i3-a1-p1-r1 was represented by rapiii, riiaip, riiiap, riiipa, riipai, riipia, ripiai, rpaiii and rpiiai. No significant information was found for e2-p2-a1-r1 or p2-w2-i1-r1.

In forthcoming stages identified peptides will be tested individually in the *in vitro* cell-based BBB assay for further validation and study of the transport mechanism involved.

Conclusions

Here we set up a novel high-throughput screening identification method based on mass spectrometry to identify the peptides within a library that showed greatest capacity to cross an *in vitro* BBB model (Figure 8). A workflow based on two MS-identification levels was established. The first MS-step consisted of choosing the peptides species corresponding to specific compositions (peptides families) by means of LTQ-Orbitrap MS, operating only on MS1 level. Subsequently, targeted mass spectrometry (SRM) was applied as the second MS-identification level by monitoring transitions defining specific amino acid sequences. Hence, peptide sequences (family members) within the selected composition were unequivocally validated.

On the basis of data obtained from the PAMPA and the *in vitro* cell-based BBB model, we propose that some peptide families cross the BBB by passive diffusion transport while others use other mechanisms of transport (including active transport and paracellular flux). Most sequences in the i3-a1-p1-r1 and i4-p1-r1 families corresponding to other mechanisms of transport were validated in SRM experiments. Similarly, most sequences in the i3-p2-r1, p3-s2-r1 and a3-r2-p1 families corresponding to passive diffusion transport were also validated using the same experiments.

Here we demonstrate the suitability of this novel high-throughput screening method for a variety of systems in which peptide libraries might be a valuable source of active molecules. Using specialized but widely accessible, mass spectrometry equipment, this approach allows the selection of the most promising peptide sequences interacting with or crossing an assay that mimics a biological target.

Materials and methods

Chemicals, consumables and equipment

Protected amino acids and resins were supplied by Neosystem (Strasbourg, France), Calbiochem-Novabiochem AG (Laüfelfingen, Switzerland), Bachem AG (Bubendorf, Switzerland) and Iris Biotech, (Marktredwitz, Germany). Oxyma Pure (ethyl 2-cyano-2-(hydroxyimino)acetate) was provided by Luxembourg Industries (Tel-Aviv, Israel) and COMU by Aldrich (Milwaukee, WI USA). DIEA and ninhydrin were from Fluka Chemika (Buchs, Switzerland). Solvents for peptide synthesis and RP-HPLC were from Scharlau or SDS (Barcelona, Spain). Trifluoroacetic acid was purchased from KaliChemie (Bad Wimpfen, Germany). The other chemicals used were from Aldrich (Milwaukee, WI USA) and were of the highest purity commercially available. PAMPA plates and PAMPA system solution were from pION (Woburn, MA USA). Porcine polar brain lipid extract (PBLEP) was supplied by Avantis Polar Lipids (Alabaster, AL USA). HPLC

chromatograms were recorded on a Waters model Alliance 2695 with photodiode array detector 996 (Waters, Milford, USA) using a symmetry C_{18} column (150 × 4.6 mm × 5 µm, 100 Å, Waters), solvents: H_2O (0.045% TFA) and MeCN (0.036% TFA); flow rate of 1 mL/min; and software Millenium version 4.0. HPLC-MS [Waters model Alliance 2796, quaternary pump, Waters 2487 with UV/Vis dual absorbance detector, ESI-MS model Micromass ZQ and Masslynx version 4.0 software (Waters)] was done using a Symmetry 300 C_{18} column (150 × 3.9 mm × 5 µm, 300 Å, Waters); solvents: H_2O (0.1% formic acid) and MeCN (0.07% formic acid); and flow rate of 1 mL/min. All the reagents used for the set-up of the *in vitro* cell-based BBB model were supplied by Biological industries, Calbiochem, Gibco, Lonza, Panreac, Sigma-Aldrich and Worthington. Bovine brain microvascular endothelial cells were purchased from *Cell Applications* and Wistar rats from *Charles River*.

Peptide library synthesis

Library synthesis was performed using the mix-and-split [20] methodology and solid-phase peptide synthesis (SPPS) using the 9-fluorenylmethoxycarbonyl/*tert*-butyl (Fmoc/*t*Bu) strategy. In all cases peptide elongation and other manipulations were done manually in polypropylene syringes provided with porous polyethylene disks. Solvents and soluble reagents were removed by suction. Washings between synthetic steps were carried out with DMF (5 x 30 s) and DCM (5 x 30 s) using approximately 5 mL of solvent/g of resin. During coupling and deprotection reactions, the mixture was allowed to react with intermittent manual stirring (Teflon stick). Seven small reactors, consisting on 20-mL polypropylene syringes, each provided with a polyethylene porous disk, were used for coupling reactions and Fmoc group removal. Each 20-mL reactor was devoted to the coupling of a single amino acid (D-Ala, D-Arg, D-Ile, D-Glu, D-Ser, D-Trp and D-Pro). A 60-mL polypropylene syringe provided with a polyethylene porous disk was used as a container for mixing all resin for further splitting. Coupling reactions were performed with Fmoc-D-amino acids (4 eq.), oxima (4 eq.), COMU (4 eq.) and DIEA (8 eq.) for 90 min. Recoupling was performed in all cases. Fmoc group deprotection was carried out with a solution of piperidine (20% in DMF) with one treatment of 1 min and 2 treatments of 10 min. An additional treatment with a solution of 5% DBU, 20% piperidine, and 5% of toluene in DMF was applied on the D-Pro reactor for 1 min and two treatments of 5 min. The sixth amino acid (D-Arg), which is common to all peptides in the library, was coupled in the 60-mL reactors under the same conditions: Fmoc-D-Arg(Pbf)-OH (4 eq.), oxyma (4 eq.), COMU (4 eq.) and DIEA (8 eq.) with two treatments of 90 min. On the same reactor, acetylation was performed with acetic acid (3 eq.) and DIPCDI (6 eq.) in DCM with two treatments of 15 min. Prior to addition of the resin, anhydride formation was performed in a separate vessel. Finally, peptides were cleaved from the resin using a cleavage cocktail composed by TFA, TIS, and H_20 (95:2.5:2.5). Three treatments of 1 h each were performed for cleavage. Cleaved peptides were precipitated through addition of cold *tert*-butyl methyl ether and centrifuged at 4ºC and 4000 rpm for 10 min. The ether fraction was discarded and precipitation was repeated up to 3 times to remove all scavengers and by-products from the cleavage reaction. Finally, cleaved peptides were dried with a N_2 flow, resuspended in H_2O/MeCN (1:1), separated

from the resin through filtration, and finally lyophilized. Lyophilised peptides were redissolved with 1M HCl H_2O/MeCN (1:1) and lyophilized three times to replace trifluoroacetate counterion for a chloride ion.

Peptide library characterization

Peptide library characterization was performed using two methods.

a) Cleavage of single beads

Proper synthesis should lead to a one-bead-one-compound library. Prior to the cleavage step, some resin beads were picked up and cleaved in individual vials with a cleavage cocktail composed by TFA, TIS, H_20 (95:2.5:2.5) for 1 h. Cleavage cocktail was dried under a N_2 flux and dissolved in H_2O/MeCN (1:1). 1 µL of peptide solution was mixed with 1 µL of ACH matrix in a MALDI-plate and MALDI-TOF spectra were recorded. Each peptide was analyzed by MALDI-TOF MS in tandem mode. Analysis was performed on 15 beads, and single peptide sequences were characterized on each bead.

b) Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR)

500 µg of library sample was reconstituted in 50 µL of H_2O/MeCN (1:1) with 0.1% formic acid. Direct infusion (automated nanoelectrospray) of sample was performed on an LTQ-FT Ultra apparatus (Thermo Scientific). The NanoMate (Advion BioSciences, Ithaca, NY, USA) aspirated the samples from a 384-well plate (protein Lobind) by means of disposable, conductive pipette tips and infused the sample through the nanoESI Chip (which consists of 400 nozzles in a 20x20 array) towards the mass spectrometer set to positive ionization. Spray voltage was 1.90 kV and delivery pressure was 0.8 psi. The capillary temperature was 200ºC, capillary voltage 44 V and tube lens 120 V. The m/z range screened comprised between 200 and 1500 a.m.u. Data were acquired with Xcalibur software, vs.2.0SR2 (ThermoScientific).

Primary cell lines: General considerations

Coating of culture flask surfaces is required for proper cell seeding: poly-D-lysine coating for astrocytes and collagen and fibronectin coatings for bovine brain endothelial cells.

Poly-D-lysine coating: cell culture flasks were treated with 6 mL of poly-D-lysine (10 µg/mL in H_2O) for 2 h at 37ºC. Poly-D-lysine solution was removed by suction, and flasks were left on the hood until completely dry (30 min).

Collagen and fibronectin coating: cell culture flasks were coated by adding 6 mL of collagen (10 µg/mL in H_2O, 0.1% v/v acetic acid) for 2 h at room temperature. Collagen coating solution was removed by suction and flasks were washed three times with PBS. Subsequently, 5 mL of fibronectin coating solution was added (10 µg/mL in H_2O) for 30 min. Fibronectin coating solution was aspirated, and flasks were immediately used for cell culture purposes.

Three distinctly enriched DMEM (4500 mg/L glucose) standard media was used. DMEM (-) for rat dissection: 2 mM glutamine, 50 U/mL of penicillin, and 0.05 g/mL streptomycin. DMEM (+) for astrocyte conditioning: 10% heat inactivated calf serum, 2 mM glutamine, 50 U/mL of penicillin, and 0.05 g/mL streptomycin. DMEM (+/+): 10% heat inactivated calf serum, 2 mM glutamine, 25 mM HEPES, 5 mL of MEM non-essential amino acids, 50 U/mL of penicillin, and 0.05 g/mL streptomycin. DMEM (+/+) was used as the standard media for cell maintenance, both rat astrocytes and bovine brain endothelial cells. Culture medium was changed every other day.

Trypsinization of cells was performed with 0.25% trypsin-EDTA for 1 min at 37ºC. Cells were washed with DMEM (+/+) and centrifuged at 1000 rpm for 1 min to remove all trypsin. If cells were not detached with trypsin treatment, a scraper was used. Cell pellet was resuspended in DMEM (+/+) and cell density was determined in a Neubauer counting plate.

Astrocyte isolation

The whole process was performed with strictly sterilized instruments. Solutions 1 to 5 were freshly prepared the same day of dissection. Solution 1: 50 mL Krebs-Ringer buffer 1X, 0.15 g BSA, and 0.4 mL $MgSO_4$ at 3.8% (w/v). Solution 2: 10 mL solution 1, and 2.5 mg trypsin. Solution 3: 10 mL solution 1, 0.8 mg DNAse, 5.2 mg STBI, and 0.1 mL $MgSO_4$ at 3.8% (w/v). Solution 4: 8.4 mL solution 1 and 1.6 mL of solution 3. Solution 5: 5 mL solution 1, 40 µL $MgSO_4$ at 3.8% (w/v), and 6 µL $CaCl_2$ at 1.2% (w/v). Six newborn Wistar rats (2 pups/flask) at the age of 2 to 3 days were decapitated with scissors and each head was dipped in a 70% ethanol solution. The brains were extracted from the skulls and placed on a petri dish containing ice-cold DMEM (-). The cortex of each brain was carefully isolated and meninges were peeled out. Cortexes were pulled, chopped with a blade, and transferred with a pipette to a 50-mL centrifuge tube, together with solution 1. The mixture was centrifuged at 4ºC and 1000 rpm for 5 sec. The supernatant was discarded and solution 2 was added. The mixture was carefully shaken at room temperature for 5 min, after which solution 4 was added, and the mixture was centrifuged at 4ºC and 1000 rpm for 5 sec. The supernatant was discarded, and two Pasteur pipette volumes of solution 3 were added. The pellet was resuspended and thoroughly homogenized using a Pasteur pipette with a narrowed end. Solution 5 was added to the homogenized mixture. The mixture was centrifuged at 4ºC and 1000 rpm for 5 min. The supernatant was discarded and the pellet was carefully resuspended in DMEM (+) and seeded in three p75 cell culture flasks. Cell cultures were maintained at 37ºC and 5% CO_2 for 3 days (it is highly recommended not to manipulate flasks during the first three days), when culture media had to be replaced. At day 6, culture media was renewed, and cell cultures were shaken at 90 rpm under anaerobic conditions overnight. Culture media were removed, and cells still attached to the flasks were carefully rinsed with PBS and cultured again with DMEM (+)at 37ºC and 5% CO_2 until day 9. Astrocytes were subcultured in p75 culture flasks previously coated with poly-D-lysine.

All experiments using rats were approved by the appropriate institutional review committee and performed in strict compliance with the European Community Guide for the Care and Use of Laboratory Animals.

Construction of *in vitro* cell-based BBB model

The established model consists of a co-culture of bovine brain endothelial cells and rat astrocytes. The model was built on 24-well plates containing inserts or transwells containing a polycarbonate membrane with a surface area of 0.33 cm^2 and pore-size of 0.4 µm. The upper surface of transwells was coated with collagen type IV and fibronectin. First, filter membranes were incubated with 100 µL of collagen type IV solution (10 µg/mL) in H_2O (0.1% v/v acetic acid) for 2 h at room temperature. Transwells were rinsed three times with PBS to remove traces of acetic acid, and 100 µL of fibronectin solution (100 µg/mL in PBS) was added for transwell membrane coating for 30 min at room temperature, after which the solution was removed by suction. Transwells were placed upside down in plastic boxes containing 10 mL of PBS in order to maintain high relative humidity inside the box. Immediately, 90-µL droplets containing approximately 45000 astrocytes were seeded on the bottom part of each filter. Astrocytes were incubated upside down for 75 min. Fresh DMEM was added to each filter every 15 min. Inserts were carefully transferred to the 24-well plate containing 800 µL of DMEM media. The upper well was filled with 200 µL of DMEM. The transwell system was incubated at 37ºC and 5% CO_2 for 3 days. It is highly recommended not to manipulate transwell system during the first three days. Medium was changed by enriched differentiation medium (clonetics EMVB SingleQuots, Lonza) with heparin (125 µg/mL in PBS, freshly prepared), HEPES (25 mM in PBS) and MEM (10 µL/mL) 2 h prior to seeding the bovine brain endothelial cells. 45000 cells contained in droplets of 20 µL were seeded in the upper part of the filter. Cells were homogeneously distributed across the membrane by carefully pipetting up and down the upper transwell suspension. The transwell system was incubated at 37ºC and 5% CO_2 for three days. At day 3 of co-culture, medium was replaced by differentiation medium (clonetics EMVB SingleQuots, Lonza) supplemented with 8-(4-chlorophenylthio)-cAMP (154 mg/L in H_2O) and 4-(3-butoxy-4-methoxybenzyl)-2-imidazolidinone (4.9 mg/L in DMSO) and incubated at 37ºC and 5% of CO_2. In general, at day 8 of co-culture, the system was ready to perform transport studies. To assess the maturity of the cell layer transendothelial electrical resistance measurements (TEER) were performed.

Transport assay: *in vitro* blood-brain barrier model

Inserts were washed with Ringer/HEPES solution. Acceptor (lower or baso-lateral) compartments were filled with 0.8 mL of Ringer/HEPES solution, while the donor compartments (upper or apical) were filled with 0.2 mL of peptides dissolved in Ringer/HEPES. The assay was performed for 2 h, incubating the plate at 37ºC and 5% CO_2. Assayed peptides were co-incubated with 20 µM of Lucifer Yellow (LY). After the experiment, LY permeability was determined by transferring an aliquot of the samples to a fluorescence plate reader (excitation at 428 nm and emission at 536 nm). A LY permeability over $1.7 \cdot 10^{-5}$ cm/s indicates cell membrane disruption or breakage during the assay, and such wells discarded.

P_{app} was determined by:

$$P_{app} = \frac{\partial Q}{\partial t} \cdot \frac{1}{A \cdot C_0}$$

where (dQ/dt) is the amount of sample present in the acceptor compartment in function of time (nmol/s), A is the area of the insert (cm^2), and C_0 is the initial concentration of sample applied to the donor compartment (nmol/mL).

The percentage of transport is calculated using the following equation:

$$\%T = \frac{Q_A(t)}{Q_D(t_0)} \cdot 100$$

where $Q_A(t)$ is the compound concentration in the acceptor well at time t and $Q_D(t_0)$ is the compound concentration in the donor well at $t = 0$.

All samples were tested in triplicate. The Ac-D-Arg-library was tested at 10% of saturation concentration.

Transport assay: Parallel artificial membrane permeability assay

The parallel artificial membrane permeability assay (PAMPA) was used to assess the capacity of the peptides to cross the BBB by passive diffusion by means of effective permeability (P_e). The initial concentration of compounds was set at 200 µM in a buffer solution prepared by dilution from a commercial buffer solution by pION, following the manufacturer's instructions. The PAMPA sandwich consists of two 200-µM wells separated by a polycarbonate filter membrane, which is coated with 4 µL of phospholipids in dodecane (20 mg/mL). Peptide sample (195 µL) was placed in the donor compartment (lower) containing a magnetic seed, while the acceptor compartment (upper) was filled with 200 µL of plain buffer. The plate was covered and incubated at room temperature in a saturated humidity atmosphere for 4 h in Gut-Box™ chamber at 25 µm of unstirred water layer (UWL). The content of acceptor and donor compartments was evaluated by HPLC-UV or by MALDI-TOF, and P_e, percentage of transport ($\%T$) and membrane retention (% membrane retention) were calculated.

$$P_e = \frac{-218.3}{t} \cdot log\left[1 - \frac{2 \cdot Q_A(t)}{Q_D(t_0)}\right] 10^{-6}\ cm/s$$

$$\%T = \frac{Q_A(t)}{Q_D(t_0)} \cdot 100$$

$$\%\ membrane\ retention = [Q_D(t_0) - Q_A(t) - Q_D(t)] \cdot 100$$

Off-line reverse phase purification protocol

Selected columns (Ultra MicroSpin C_{18}, 300Å silica (5-60 µg) commercial column) were conditioned with 200 µL of "conditioning buffer" (MeOH) two times, and spun in a centrifuge (1000 x g for 1 min) to push the flow through. The column was equilibrated with 200 µL "equilibration buffer" three times with

spinning. Peptide samples were dissolved in H_2O (with 5% formic acid) and loaded onto the column. Washing buffer was applied to the column three times (200 µL), and finally peptides were eluted with three treatments of 200 µL of "elution buffer". Solvent from eluted peptides was evaporated using a SpeedVac system. Samples were dissolved in 40 µL of 0.1% formic acid for injection. Conditioning buffer: 100% methanol. Equilibration/washing buffer: 5% formic acid in H_2O. Elution buffer: 5% formic acid in H_2O /MeCN (1:1, v/v).

Sample treatment before HPLC-MS injection

Acceptor samples and t_0 samples from the PAMPA assay and *in vitro* cell-based BBB assay were evaporated using a SpeedVac system. Samples were redissolved in 200 µL of H_2O (with 5% formic acid) and desalted with off-line reverse phase purification.

LTQ-Orbitrap-XL

Library peptides were analyzed using a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an Agilent Technologies 1200 Series (CA, USA). Peptides were loaded onto C_{18} Zorbax precolumn (Agilent Technologies, cat #5065-9913) and separated by reverse-phase chromatography using a 12-cm column with an inner diameter of 75 µm, packed with 5 µm C_{18} particles (Nikkyo Technos Co., Ltd. Japan). Chromatographic gradients started at 97% buffer A and 3% buffer B with a flow rate of 300 nL/min, and gradually increased to 90% buffer A and 10% buffer B in 1 min, and to 65% buffer A / 35% buffer B in 30 min. After each analysis, the precolumn and column were washed for 10 min with 10% buffer A / 90% buffer B. Buffer A: 0.1% formic acid in H_2O. Buffer B: 0.1% formic acid in MeCN.

The mass spectrometer was operated in positive ionization mode with nanospray voltage set at 2.5 kV and source temperature at 200 °C. Ultramark 1621 for the FT mass analyzer was used for external calibration prior to the analyses. Moreover, an internal calibration was also performed using the background polysiloxane ion signal at m/z 445.1200. Full MS scans with 1 micro scans at resolution of 60000 were used over a mass range of m/z 350-2000 with detection in the Orbitrap to obtain the exact mass of peptide precursors. Auto gain control (AGC) was set to $1 \cdot 10^6$, dynamic exclusion (60 seconds). All data were acquired with Xcalibur software v2.2.

Q-trap

The peptide mixtures were analyzed using a 5500 Q-trap mass spectrometer (AB Sciex, Framingham, MA, USA) coupled to a nanoLC Ultra-1DPlus (AB Sciex (Eksigent)). Peptides were loaded onto C_{18} Acclaim PepMap precolumn (Thermo Scientific, cat # 164564) and separated by reverse-phase chromatography using a 12-cm column with an inner diameter of 75 µm, packed with 5-µm C_{18} particles (Nikkyo Technos Co., Ltd. Japan). Chromatographic gradients started at 98% buffer A and 2% buffer B with a flow rate of 300 nL/min for 5 min and gradually increased to 60% buffer A and 40% buffer B in 35 min. After each analysis, the precolumn and column were washed for 10 min with

2% buffer A / 98% buffer B. Buffer A: 0.1% formic acid in H_2O. Buffer B: 0.1% formic acid in MeCN. SRM data were processed using the Skyline software (v 2.1).

Acknowledgements

We thank the IRB Barcelona Mass Spectrometry Core Facility and the Proteomic Platform at the Barcelona Science Park. This work was supported by MINECO-FEDER (BIO2013-40716-R and CTQ2013-49462-EXP), *Generalitat de Catalunya* (XRB and 2014-SGR-521) and Gate2Brain-RecerCaixa 2014.

Captions:

Figure 1. A) General structure of the synthesized mix-and-split peptide library. B) Set of selected amino acids used for the synthesis of the mix-and-split peptide library. This library design leads to 16.807 peptides ($R = r^m$; where r is the number of amino acids used to introduce variability, and m is the number of peptide residue positions where variability is introduced). C) Mix and split SPPS Ac-D-Arg-library synthetic workflow. Library synthesis was performed by the mix-and-split methodology in SPPS, using Rink-amide AM resin (200 mesh). Synthetic workflow comprises 5 rounds of mix-and-split. Each round involved dividing the resin in a large vessel into 7 smaller reactors, coupling amino acid building blocks, Fmoc deprotection, and mixing the resin again in the large vessel. Once five residues were coupled, a final coupling step was performed in the larger vessel to introduce Fmoc-D-Arg(Pbf)-OH as the *N*-terminus amino acid in all peptides of the library. Finally, the whole library was *N*-terminus-acetylated and cleaved from the resin.

Figure 2. Comparison of relative abundance of the library peptides in theoretical and experimental spectra recorded by FT-ICR mass spectrometrer, up and upside down respectively.

Figure 3. A) Scheme of PAMPA. An artificial phospholipid membrane mimics only passive diffusion properties of the BBB. The thickness of the phospholipidic membrane has been described to be equivalent to 300 lipid bilayers. B) Scheme of *in vitro* cell-based BBB assay. A co-culture of bovine brain endothelial cells and rat astrocytes provides an assay mimicking both passive diffusion and active related transports mechanisms described in the BBB.

Figure 4. General workflow followed to obtain peptide composition lists (peptide families). Molecular ions from all MS1 spectra recorded throughout the total ion chromatogram (TIC) chromatogram matching peptides masses within Ac-D-Arg-library are selected by *Bibliopepfinder*.

Figure 5. Venn diagrams showing relations between sets of experimental data: A) PAMPA library sample vs. PAMPA blank sample. B) *In vitro* cell-based BBB assay library sample vs. *in vitro* cell-based BBB assay blank sample. C) Triplicate samples from *in vitro* cell-based BBB assay. D) PAMPA library sample vs. *in vitro* cell-based BBB assay sample.

Figure 6. Case example (raiipi), the process followed for each single targeted sequence to search for complete sets of co-eluting transitions. A, B and C correspond to the SRM spectra and zoom in images of the same spectra. A also shows the blank spectra monitoring the same sequence.

Figure 7. Summarized information extracted from SRM experiments performed in Q-trap. Information shown in the table refers to co-eluting transitions found in triplicate samples and the retention time at which they were found. No significant information was found for e2-p2-a1-r1 or p2-w2-i1-r1.

Figure 8: General scheme of proposed novel high-throughput screening methodology based on mass spectrometry. Mix-and-split methodology was applied on SPPS for library synthesis. The library was subsequently assayed in an *in vitro* cell-based BBB model consisting of a co-culture of bovine brain endothelial cells and rat astrocytes. Concerning peptide identification, mass spectrometry techniques were used. A workflow based on two MS-identification levels was set up. 1) First MS-identification step involving the selection of the most promising peptide species corresponding to specific compositions (peptide families) with LTQ-Orbitrap MS, only operating on MS1 level. 2) Subsequently, targeted mass spectrometry (SRM) was applied as the second MS-identification level by monitoring transitions defining specific amino acidic sequences. Hence, peptide sequences (family members) comprised within the selected composition can be unequivocally validated. RAIIPI is shown as an example.

Table 1. Summary of parameters considered for the selection of a reduced amino acid set. Second column in the table shows the propensity of each amino acid to form a secondary structure: α-helix and β-sheet. Symbol ≈ indicate a similar value for α-helix and β-sheet structures; ≈β a medium propensity to form a β-sheet; and α and β high propensity to form α-helix and β-sheet, respectively. The third column shows the hydrophobicity of selected amino acids in the Kyte-Doolittle scale. Selected amino acids include hydrophobic, hydrophilic and neutral molecules. The fourth column shows the percentage of amino acid presence in active sites with respect to the total number of selected amino acids in the sequence. Alanine value is not known since alanine scanning was the methodology to calculate this parameter. This amino acid set includes 3 top amino acids on this parameter: Trp, Arg and Ile. Furthermore, proline was included in these selections since it is the only proteinogenic *N*-alkyl amino acid.

Table 2. Family column stands for a general molecular formula code expressing the variety of amino acids and their repetition, which determine the number of possible family member columns for all families comprised within the library.

Table 3. Selection criteria for peptide composition/family selection for further analysis on the *second level of identification*.

Figure 1

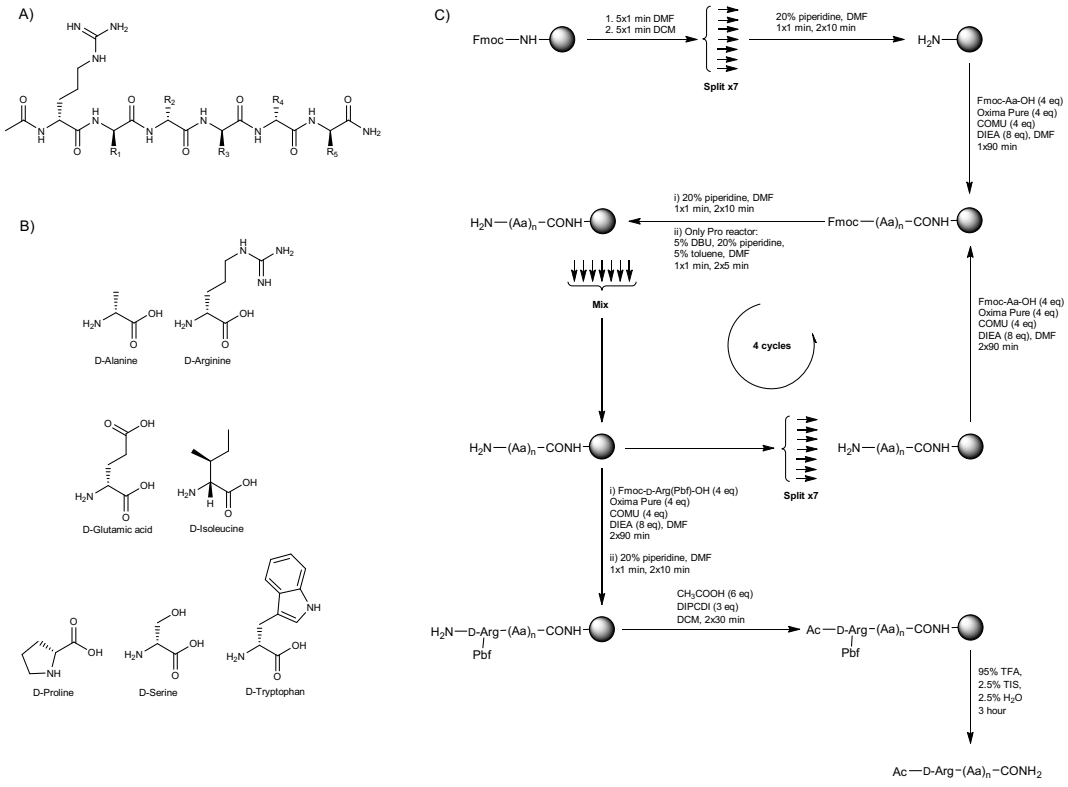

A)
B)
D-Alanine
D-Arginine
D-Glutamic acid
D-Isoleucine
D-Proline
D-Serine
D-Tryptophan
C)
Fmoc—NH
1. 5x1 min DMF
2. 5x1 min DCM
Split x7
20% piperidine, DMF
1x1 min, 2x10 min
H2N
Fmoc-Aa-OH (4 eq)
Oxima Pure (4 eq)
COMU (4 eq)
DIEA (8 eq), DMF
1x90 min
i) 20% piperidine, DMF
1x1 min, 2x10 min
ii) Only Pro reactor:
5% DBU, 20% piperidine,
5% toluene, DMF
1x1 min, 2x5 min
Mix
4 cycles
Fmoc-Aa-OH (4 eq)
Oxima Pure (4 eq)
COMU (4 eq)
DIEA (8 eq), DMF
2x90 min
Split x7
i) Fmoc-D-Arg(Pbf)-OH (4 eq)
Oxima Pure (4 eq)
COMU (4 eq)
DIEA (8 eq), DMF
2x90 min
ii) 20% piperidine, DMF
1x1 min, 2x10 min
CH3COOH (6 eq)
DIPCDI (3 eq)
DCM, 2x30 min
Pbf
95% TFA,
2.5% TIS,
2.5% H2O
3 hour
Ac—D-Arg-(Aa)n-CONH2

Table 1

Amino acid	Secondary structure propensity	Hydropathy index [21]	Preferences in hot spots (%) [22]
D-Trp	≈β	-0.9	21.05
D-Arg	≈	-4.5	13.30
D-Ala	α	1.8	-
D-Glu	α	-3.5	3.64
D-Ser	≈	-0.8	1.12
D-Ile	β	4.5	9.62
D-Pro	≈	-1.6	6.74

Table 2

Type of family	Number of family members
5	1
4-1	5
3-2	10
3-1-1	20
2-2-1	30
2-1-1-1	60
1-1-1-1-1	120

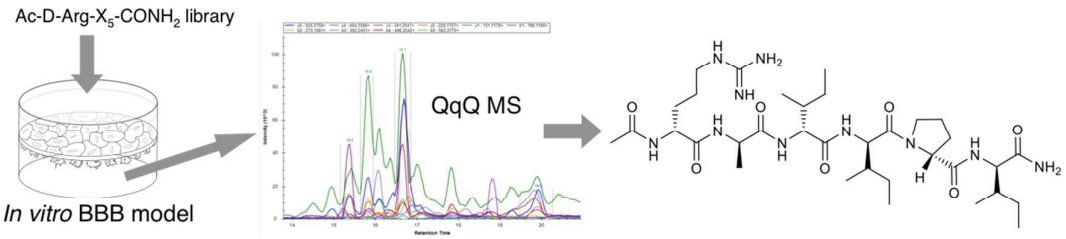
Table 3

1 - Must be in all triplicates.
2 - Its molecular peak must have a typical mass ratio profile in mass spectrometry of peptide species in the MS1 LTQ-Orbitrap spectra.
3 - Apply a score criterion based on weighted intensity ranking on all triplicates.
4 - Families with more than 30 members are automatically excluded.

Graphical table of contents

Chemically synthesized peptide libraries as a new source of BBB shuttles. Use of mass spectrometry for peptide identification.

B. Guixer, X. Arroyo, I. Belda, E. Sabidó, M. Teixidó*, E. Giralt*



Description of a new high-throughput screening methodology to search for new protease-resistant BBB shuttle peptides. This methodology combines combinatorial chemistry for peptide library synthesis, *in vitro* models mimicking the BBB for library evaluation, and state-of-the-art mass spectrometry techniques to identify those peptides able to cross the BBB assays.

References:

1. Abbott NJ, Ronnback L, Hansson E. Astrocyte-endothelial interactions at the blood-brain barrier. *Nat. Rev. Neurosci.* 2006; **7**: 41-53.
2. Krewson CE, Klarman ML, Saltzman WM. Distribution of nerve growth factor following direct delivery to brain interstitium. *Brain Res.* 1995; **680**: 196-206.
3. Ai Y, Markesbery W, Zhang Z, Grondin R, Elseberry D, Gerhardt GA, Gash DM. Intraputamenal infusion of GDNF in aged rhesus monkeys: distribution and dopaminergic effects. *J. Comp. Neurol.* 2003; **461**: 250-261.
4. Zunkeler B, Carson RE, Olson J, Blasberg RG, DeVroom H, Lutz RJ, Saris SC, Wright DC, Kammerer W, Patronas NJ, Dedrick RL, Herscovitch P, Oldfield EH. Quantification and pharmacokinetics of blood-brain barrier disruption in humans. *J. Neurosurg.* 1996; **85**: 1056-1065.
5. Pardridge WM. Molecular Trojan horses for blood-brain barrier drug delivery. *Curr. Opin. Pharmacol.* 2006; **6**: 494-500.
6. Deeken JF, Loscher W. The blood-brain barrier and cancer: transporters, treatment, and Trojan horses. *Clin. Cancer Res.* 2007; **13**: 1663-1674.
7. Teixido M, Zurita E, Malakoutikhah M, Tarrago T, Giralt E. Diketopiperazines as a tool for the study of transport across the blood-brain barrier (BBB) and their potential use as BBB-shuttles. *J. Am. Chem. Soc.* 2007; **129**: 11802-11813.
8. Malakoutikhah M, Teixido M, Giralt E. Toward an optimal blood-brain barrier shuttle by synthesis and evaluation of peptide libraries. *J. Med. Chem.* 2008; **51**: 4881-4889.
9. Malakoutikhah M, Prades R, Teixido M, Giralt E. N-methyl phenylalanine-rich peptides as highly versatile blood-brain barrier shuttles. *J. Med. Chem.* 2010; **53**: 2354-2363.
10. Malakoutikhah M, Guixer B, Arranz-Gibert P, Teixido M, Giralt E. 'À la carte' peptide shuttles: tools to increase their passage across the blood-brain barrier. *ChemMedChem*. 2014; **9**: 1594-1601.
11. Arranz-Gibert P, Guixer B, Malakoutikhah M, Muttenthaler M, Guzman F, Teixidó M, Giralt E. Lipid bilayer crossing – the gate of symmetry. Water-soluble phenylproline-based blood-brain barrier-shuttles. *J. Am. Chem. Soc.* 2015; DOI: 10.1021/jacs.5b02050.
12. Demeule M, Regina A, Che C, Poirier J, Nguyen T, Gabathuler R, Castaigne JP, Beliveau R. Identification and design of peptides as a new drug delivery system for the brain. *J. Pharmacol. Exp. Ther.* 2008; **324**: 1064-1072.
13. Prades R, Oller-Salvia B, Schwarzmaier SM, Selva J, Moros M, Balbi M, Grazu V, de La Fuente JM, Egea G, Plesnila N, Teixido M, Giralt E. Applying the retro-enantio approach to obtain a peptide capable of overcoming the blood-brain barrier. *Angew. Chem. Int. Ed.* 2015; **54**: 3967-3972.

14. Oller-Salvia B, Teixido M, Giralt E. From venoms to BBB shuttles: Synthesis and blood-brain barrier transport assessment of apamin and a nontoxic analog. *Biopolymers* 2013; **100**: 675-686.
15. Oller-Salvia B, Sánchez-Navarro M, Giralt E, Teixidó M. Blood-brain barrier shuttle peptides: an emerging paradigm for brain delivery. *Chem. Soc. Rev*. 2016; *in press*.
16. Rousselle C, Smirnova M, Clair P, Lefauconnier JM, Chavanieu A, Calas B, Scherrmann JM, Temsamani J. Enhanced delivery of doxorubicin into the brain via a peptide-vector-mediated strategy: saturation kinetics and specificity. *J. Pharmacol. Exp. Ther.* 2001; **296**: 124-131.
17. Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science* 1999; **285**: 1569-1572.
18. Prades R, Guerrero S, Araya E, Molina C, Salas E, Zurita E, Selva J, Egea G, Lopez-Iglesias C, Teixido M, Kogan MJ, Giralt E. Delivery of gold nanoparticles to the brain by conjugation with a peptide that recognizes the transferrin receptor. *Biomaterials* 2012; **33**: 7194-7205.
19. Sergeeva A, Kolonin MG, Molldrem JJ, Pasqualini R, Arap W. Display technologies: application for the discovery of drug and gene delivery agents. *Adv. Drug Deliv. Rev.* 2006; **58**: 1622-1654.
20. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 1991; **354**: 82-84.
21. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982; **157**: 105-132.
22. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 1998; **280**: 1-9.
23. Palmblad M, Drijfhout JW, Deelder AM. High resolution mass spectrometry for rapid characterization of combinatorial peptide libraries. *J. Comb. Chem.* 2010; **12**: 65-68.
24. Cecchelli R, Dehouck B, Descamps L, Fenart L, Buee-Scherrer VV, Duhem C, Lundquist S, Rentfel M, Torpier G, Dehouck MP. In vitro model for evaluating drug transport across the blood-brain barrier. *Adv. Drug Deliv. Rev.* 1999; **36**: 165-178.
25. Gaillard PJ, de Boer AG. 2B-Trans technology: targeted drug delivery across the blood-brain barrier. *Methods Mol. Biol.* 2008; **437**: 161-175.
26. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat. Methods* 2012; **9**: 555-566.
27. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* 2008; **4**: 222.
28. Smith RM. *Understanding Mass Spectra: A Basic Approach*. John Wiley & Sons Inc., New Jersey, **2004**.

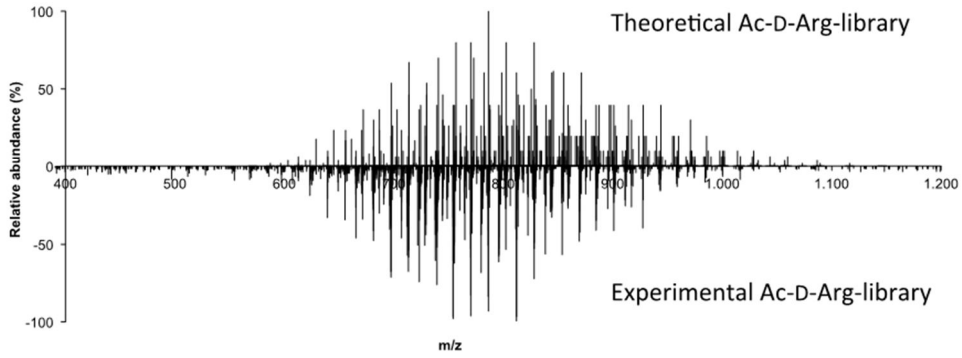


Figure 2. Comparison of relative abundance of the library peptides in theoretical and experimental spectra recorded by FT-ICR mass spectrometrer, up and upside down respectively.
239x100mm (96 x 96 DPI)

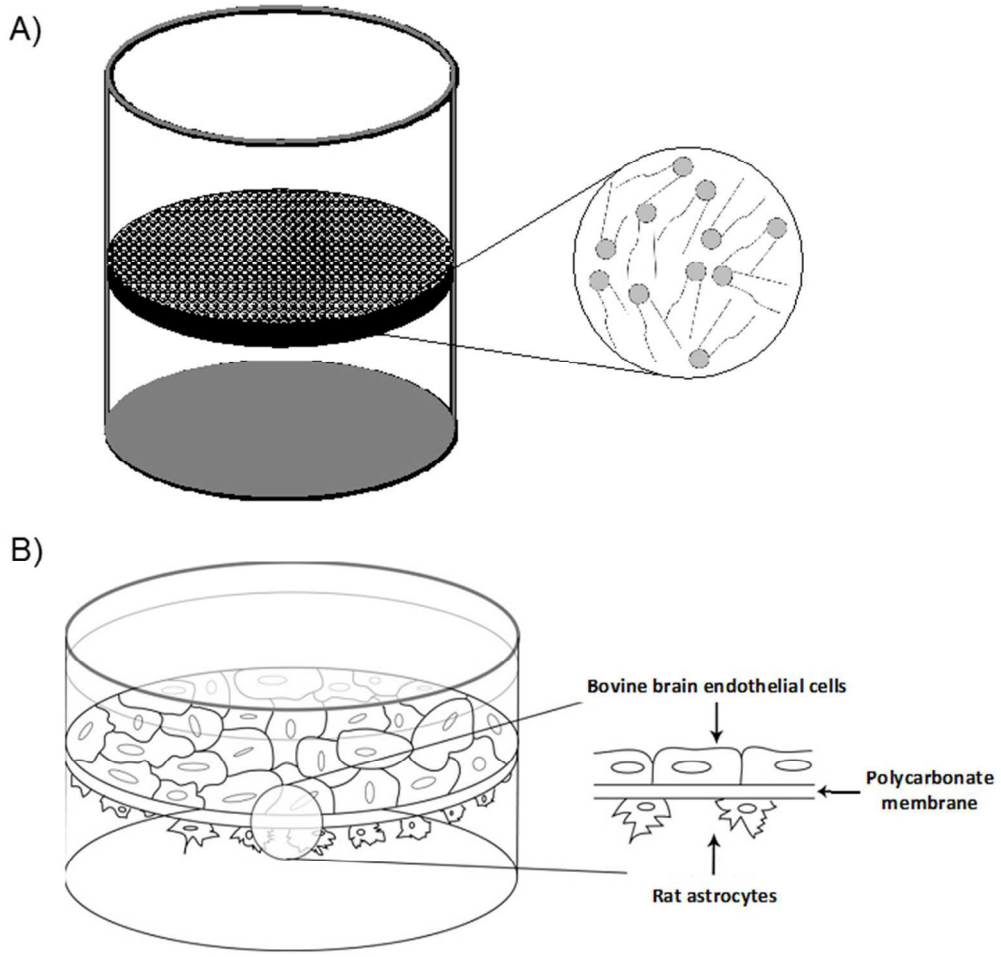


Figure 3. A) Scheme of PAMPA. An artificial phospholipid membrane mimics only passive diffusion properties of the BBB. The thickness of the phospholipidic membrane has been described to be equivalent to 300 lipid bilayers. B) Scheme of in vitro cell-based BBB assay. A co-culture of bovine brain endothelial cells and rat astrocytes provides an assay mimicking both passive diffusion and active related transports mechanisms described in the BBB.
191x190mm (96 x 96 DPI)

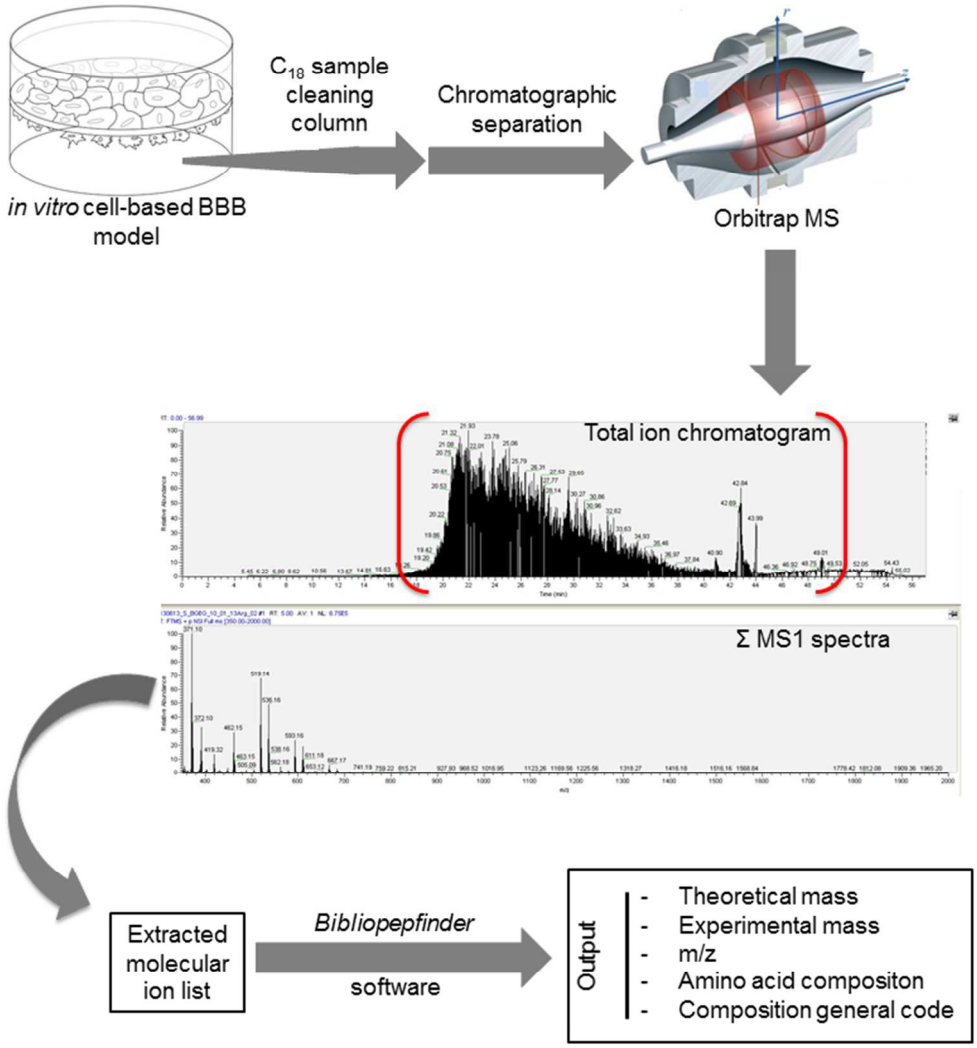


Figure 4. General workflow followed to obtain peptide composition lists (peptide families). Molecular ions from all MS1 spectra recorded throughout the total ion chromatogram (TIC) chromatogram matching peptides masses within Ac-D-Arg-library are selected by Bibliopepfinder.
210x230mm (96 x 96 DPI)

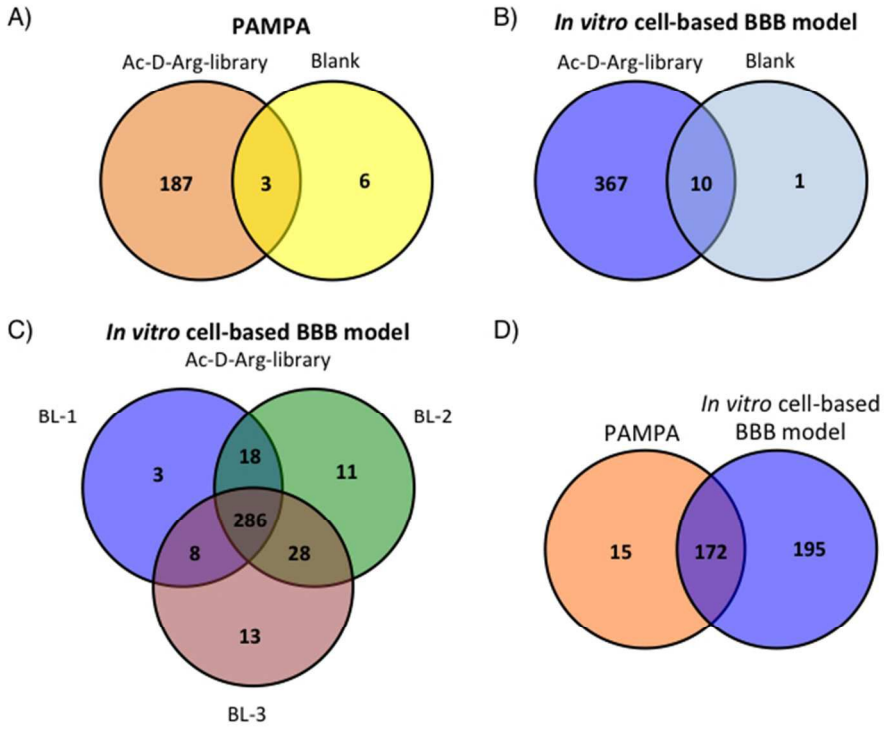


Figure 5. Venn diagrams showing relations between sets of experimental data: A) PAMPA library sample vs. PAMPA blank sample. B) In vitro cell-based BBB assay library sample vs. in vitro cell-based BBB assay blank sample. C) Triplicate samples from in vitro cell-based BBB assay. D) PAMPA library sample vs. in vitro cell-based BBB assay sample.
254x190mm (72 x 72 DPI)

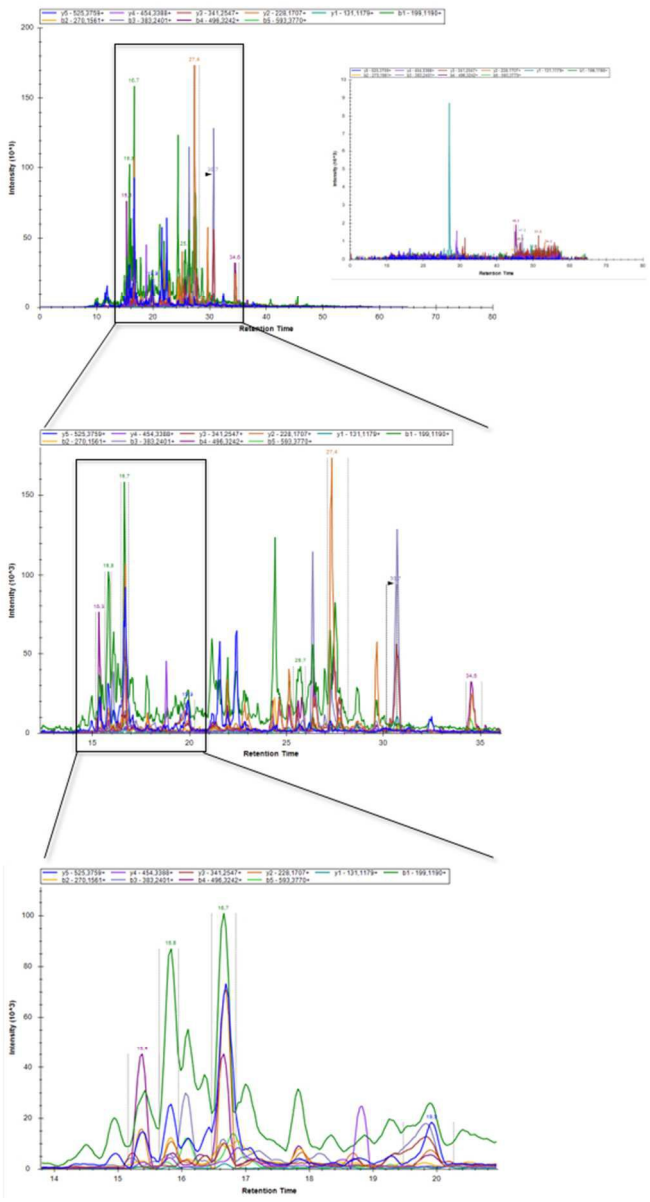


Figure 6. Case example (raiipi), the process followed for each single targeted sequence to search for complete sets of co-eluting transitions. A, B and C correspond to the SRM spectra and zoom in images of the same spectra. A also shows the blank spectra monitoring the same sequence.
155x280mm (96 x 96 DPI)

Other mechanisms of transport list		*Passive diffusion transport list*		
i3-a1-p1-r1	i4-p1-r1	i3-p2-r1	p3-s2-r1	a3-r2-p1
raiiip	riiiip	riiipp	rpppss	raaapr
raiipi	riiipi	riipip	rppsps	raaarp
raipii	riipii	riippi	rppssp	raapar
rapiii	ripiii	ripiip	rpspps	raapra
riaiip	rpiiii	ripipi	rpspsp	raarap
riaipi		rippii	rpsspp	raarpa
riapii		rpiiip	rsppps	rapaar
riiaip		rpiipi	rsppsp	rapara
riiapi		rpipii	rspspp	rapraa
riiiap		rppiii	rssppp	raraap
riiipa				rarapa
riipai				rarpaa
riipia				rpaaar
ripaii				rpaara
ripiai				rparaa
ripiia				rpraaa
rpaiii				rraaap
rpiaii				rraapa
rpiiai				rrapaa
rpiiia				rrpaaa

Legend
Co-eluting transitions not found in any triplicate
Co-eluting transitions found in some triplicates, but not at the same retention time
Co-eluting transitions found in some triplicates, at the same retention time
Co-eluting transitions found in all triplicates, but not at the same retention time
Co-eluting transitions found in all triplicates, at the same retention time, but not on the same precursor ion

Figure 7. Summarized information extracted from SRM experiments performed in Q-trap. Information shown in the table refers to co-eluting transitions found in triplicate samples and the retention time at which they were found. No significant information was found for e2-p2-a1-r1 or p2-w2-i1-r1.
190x275mm (96 x 96 DPI)

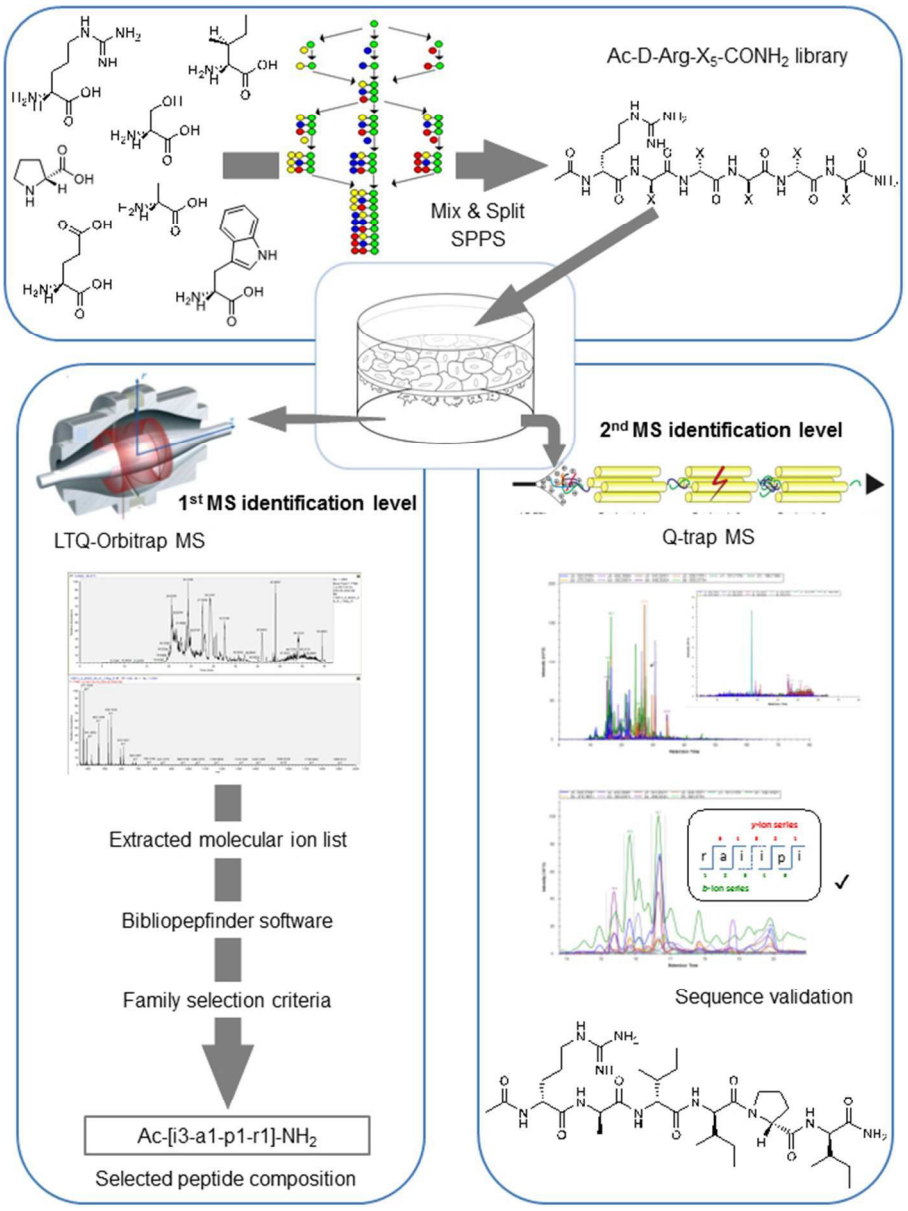


Figure 8: General scheme of proposed novel high-throughput screening methodology based on mass spectrometry. Mix-and-split methodology was applied on SPPS for library synthesis. The library was subsequently assayed in an in vitro cell-based BBB model consisting of a co-culture of bovine brain endothelial cells and rat astrocytes. Concerning peptide identification, mass spectrometry techniques were used. A workflow based on two MS-identification levels was set up. 1) First MS-identification step involving the selection of the most promising peptide species corresponding to specific compositions (peptide families) with LTQ-Orbitrap MS, only operating on MS1 level. 2) Subsequently, targeted mass spectrometry (SRM) was applied as the second MS-identification level by monitoring transitions defining specific amino acidic sequences. Hence, peptide sequences (family members) comprised within the selected composition can be unequivocally validated. RAIIPI is shown as an example.
190x254mm (96 x 96 DPI)