SPANDIDOS PUBLICATIONS

Isolation of novel cell-penetrating peptides from a random peptide library using *in vitro* virus and their modifications

KAEKO KAMIDE[1,2], HIROSHI NAKAKUBO[1], SHUSEI UNO[1] and AKIYOSHI FUKAMIZU[2]

[1]Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033; [2]Graduate School of Life and Environmental Science, University of Tsukuba, Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8572, Japan

Received August 3, 2009; Accepted September 22, 2009

DOI: 10.3892/ijmm_00000311

Abstract. A number of cell-penetrating peptides (CPPs) have been reported, but their transduction efficiencies are too low to be used as intracellular carriers for therapeutic purposes. We conducted a comprehensive search to find novel CPPs using an *in vitro* virus (IVV) library, which presented random peptides consisting of 15 amino acids (diversity of the library was $>10^{12}$). We found 9 kinds of novel CPPs with an intracellular translocation efficiency higher than that of the TAT peptide (YGRKKKRRQRRR). Interestingly, one of the novel CPPs, No. 14 (KLWMRWYSPTTRRYG), showed a dramatic improvement in translocation activity relative to the TAT peptide in CHO cells (>10-fold efficiency in 50 μM). As the intracellular translocation efficiency of No. 14 was increased by substitution Arg for Lys^1 (14-1), we carried out alanine scanning on the basis of 14-1 to determine important amino acids for the intracellular translocation. The Ala substitution analysis showed that both Arg and Trp residues were important for the cell-penetrating activity and that their contribution was in the order Trp^3<Arg^{12}<Arg^1<Arg^5, Arg^{13}<Trp^6. Moreover, it was possible to substitute two Trp with other bulky amino acids such as Ile or Tyr. In this study, we showed that novel CPPs could be acquired by screening random peptides and modifying some amino acids could increase their cell-penetrating activity.

Introduction

Cell-penetrating peptides (CPPs), also known as protein transduction domains (PTDs), are of interest because of their ability to translocate across cellular membranes. CPPs consist of 30 or fewer amino acids classified as either cationic or amphipathic in nature. TAT (basic domain of HIV-1 TAT protein) (1-3) is representative of the cationic CPPs, which contain clusters of arginine and lysine residues. Penetratin (4,5) is representative of the secondary amphipathic CPPs and penetratin-fusion proteins work well for proteins with <100 amino acid residues; however, toxicity is always a concern with these peptides (6). In addition, polylysine (7,8), polyarginine (9), HSV VP22 (10-12), MPG (HIV gp41/SV40 Tag NLS) (13,14), pep-1 (15), MAPs (16), HIV-1 Rev (17), pIsl (18), and WR peptide (19) have been reported as CPPs. Most of the currently recognized CPPs are derived from viral, insect, or mammalian proteins endowed with membrane translocation properties.

The large number of different cargo molecules that have been efficiently delivered by CPPs ranges from small molecules to proteins and even liposomes and magnetic particles (20). However, the exact mechanisms underlying the internalization of CPPs across the cellular membrane are not fully understood. Recently, increasing lines of evidence for endocytosis as a major route of entry, especially, non-classical, clathrin-independent pathways such as macropinocytosis and endocytosis via lipid rafts have been identified (21,22). Although there seems to be no consensus regarding one specific mechanism as the preferred route of CPPs uptake, it is reasonable to think that the specific route(s) of internalization depends on a number of factors, with the main one being the characteristics of CPPs and the characteristics of the cargos.

The mechanism of internalization is unclear, but the CPP approach is currently a major tool in delivery systems that mediate the import of problematic cargos into cells. Therefore, it is important to find novel CPPs having high abilities to translocate across cellular membranes. Trials using a phage-display library have been reported in the search for comprehensive CPPs (23-26). However, in the screening of the phage-display library, bias might occur in the sequence that is expressed, and the diversity is inadequate for screening CPPs. We therefore tried to obtain novel CPPs from a random peptide library using IVV for screening on a larger scale.

Materials and methods

Preparation of IVV molecules. The design and preparation of IVV was consigned to Zoegene Corporation. DNA template

Correspondence to: Dr Kaeko Kamide, Advanced Medical Research Laboratory, Research Division, Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan
E-mail: kamide.kaeko@me.mt-pharma.co.jp

Key words: TAT, penetratin, protein transduction domain, cell-penetrating peptide, *in vitro* virus, alanine scanning

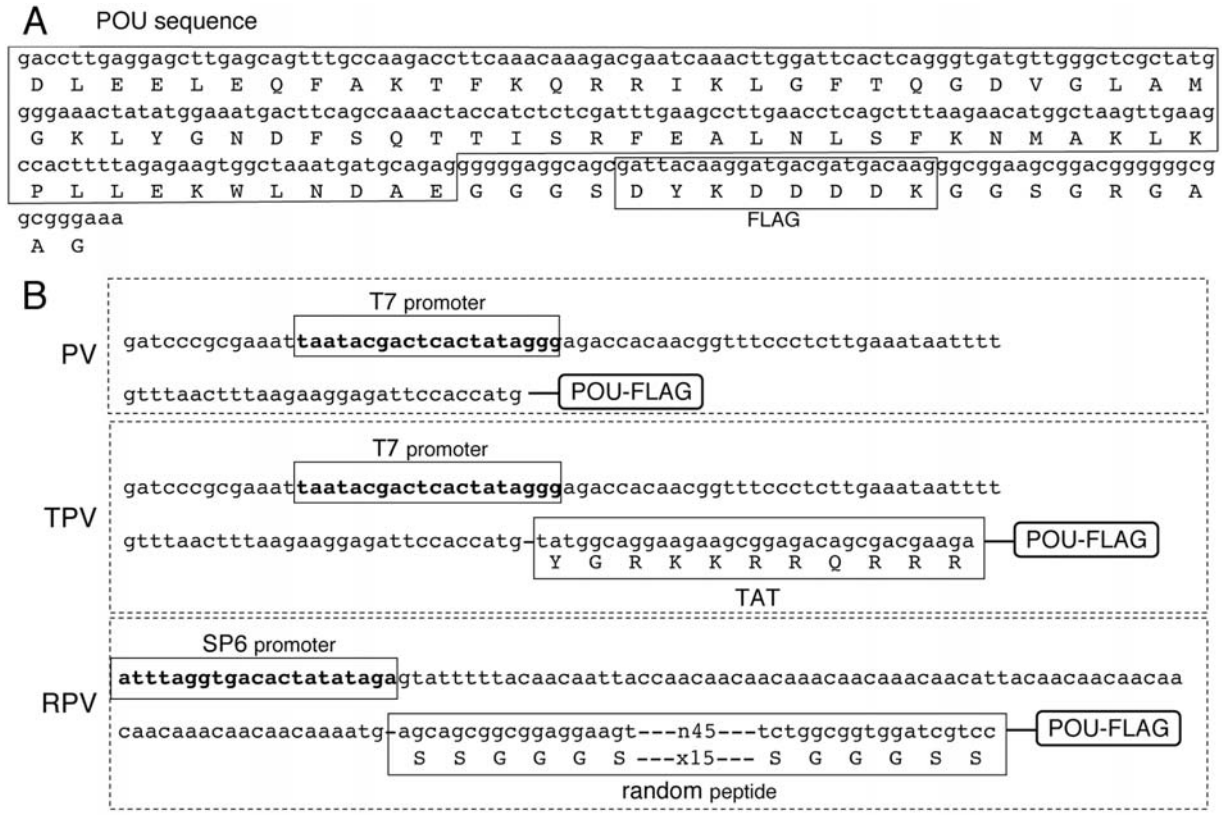


Figure 1. Nucleic acid and amino acid sequences of IVVs. (A) Nucleic acid and amino acid sequences of POU with FLAG-tag. (B) Nucleic acid and amino acid sequences of upstream of POU. PV, the upstream of POU region consisted of 96 nucleotides containing T7 promoter. TPV, TAT sequence was inserted between ATG (translation initiating codon) and the POU sequence of PV. RPV, the SP6 promoter was used instead of the T7 promoter, and random peptide sequences were inserted between ATG and the POU sequence of PV.

for the mRNA of TAT-presenting IVV (TAT-POU-IVV, TPV) was prepared using the TAT sequence of HIV-1 (27) as a known CPP and a POU sequence (Fig. 1A), which was DNA-binding domain of Oct-1, as a support protein for presenting peptides. The constructed DNA template encoded the TAT sequence, the POU sequence and a Flag-tag downstream of the T7 promoter (Fig. 1B). POU-IVV (PV) that did not contain the TAT sequence was also prepared, as a negative control (Fig. 1B). IVVs (TPV and PV) were prepared according to methods of Biyani *et al* (28). They were purified using anti-FLAG antibody immobilized on agarose beads (Sigma, St. Louis, MO, USA) and eluted with 3X FLAG peptide (Sigma).

Construction of the random peptide library. The nucleotide sequences of the random peptide were designed to include 20 amino acids equally using the MLSDS method (29). Then DNA template of random peptide for IVV library was constructed using synthetic-designed oligonucleotides. The constructed DNA templates encoded the random peptide sequence of 15 amino acids, the POU sequence and the Flag-tag downstream of the SP6 promoter (Fig. 1B). IVVs presenting random peptide (Random peptide-POU-IVV, RPV) were prepared according to methods of Biyani *et al* (28) using these DNA templates.

Recovery of the IVV genome that was translocated into cells. The prepared RPV was added to Jurkat cells ($1x10^5$ cells) and incubated at 37˚C for 1 h. After then, the cells were trypsinized and washed with PBS (phosphate-buffered saline). The washed cells were lysed with Lyse-N-Go reagent (Pierce, Rockford, IL, USA) according to the manufacturer's protocol. After the IVV genome RNA was digested with RNase H (Toyobo, Osaka, Japan), we performed ethanol precipitation of IVV genome DNA.

The pellet of the IVV genetic material was dissolved in distilled water and it was amplified the full length of the IVV coding domain with EX-Taq DNA Polymerase (Takara, Kyoto, Japan) using the sense primer (gaacaacaacaacaacaaac aacaacaaaatg) and the antisense primer (cgcccttgtcatcgtcatcc).

Cell lines. CHO-K1 (CHO), HeLa, and Jurkat cells were purchased from Dainippon Sumitomo Pharma Co., Ltd. (Osaka, Japan). CHO cells were maintained in DMEM/Ham's F-12 medium (Gibco, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (FBS) and antibiotics. HeLa cells were cultured in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% FBS and antibiotics. Jurkat cells were maintained in RPMI-1640 (Gibco) supplemented with 10% FBS and antibiotics.

Analysis of peptide sequences. The prediction of the secondary structures of the peptides were performed using the Garnier method with EMBOSS (http://www.es.embnet.org/Doc/EMBOSS/Apps/garnier.html). Helical wheel projection analysis of peptides were performed using the pepwheel

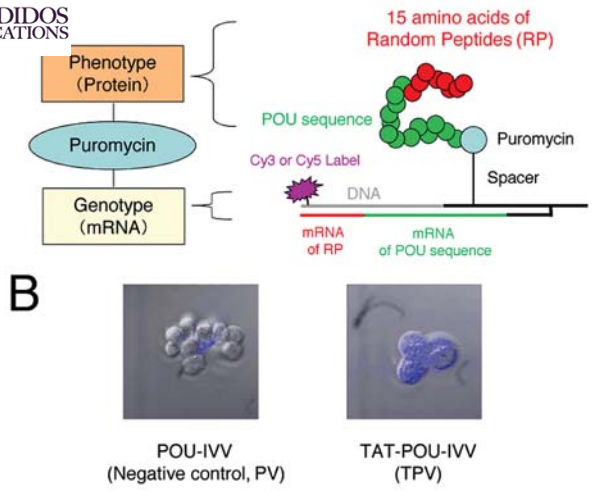


Figure 2. Structure of IVV and translocation of TAT-POU-IVV into CHO cells. (A) Structure of IVV with random peptide (RPV): IVV is a genotype (mRNA) - phenotype (protein) linking molecule with puromycin. IVV had 15 amino acids of random peptide in the N-terminal of the POU sequence and was labeled with Cy3 or Cy5. (B) Detection of translocation of PV or TPV into CHO cells using confocal laser scanning microscopy: A concentration of 45 nM IVV in PBS was added at a volume of 54 μl/well to CHO cells in 96-well plates. After incubation for 1 h, the cells were trypsinized. After washing the cells, the intensity of the fluorescence from cells was detected using confocal laser microscopy. Left, Cy5-labeled PV (negative control); right, Cy5-labeled TPV with a TAT sequence.

function of EMBOSS (http://www.sacs.ucsf.edu/Documentation/emboss/pepwheel.html).

Synthesis of fluorescein-labeled peptides. N-terminal carboxyfluorescein-labeled peptides were purchased from Thermo Electron GmbH (Ulm, Germany). Their purity was 70% or more, which was confirmed by HPLC analysis and mass spectroscopy. The fluorescein-labeled peptides were dissolved with 10% dimethyl sulfoxide (DMSO) to provide a 1 mM solution.

Flow cytometry and confocal laser scanning microscopy analysis. Cells were incubated with various concentrations of fluorescein-labeled peptides in complete medium at 37˚C for 1 h. The cells were washed twice with PBS, and then trypsinized. The cells were collected in complete medium and washed twice with Hanks' balanced salt buffer (HBSS, Gibco). The fluorescence from the cells was analyzed using flow cytometry (FC500, Beckman Coulter, Fullerton, CA, USA) or confocal laser scanning microscopy (Bio-Rad).

Intracellular localization analysis. CHO cells were seeded onto a chamber slide (Nunc, Roskilde, Denmark) at 10^5 cells/ml in complete medium and incubated for 48 h at 37˚C. The cells were treated with 50 μM fluorescein-labeled peptides in complete medium. After incubation at 37˚C for 1 h, the cells were washed with HBSS and analyzed using confocal laser scanning microscopy without fixing.

Results

Confirmation of an intracellular translocation of TAT-POU-IVV (TPV). As a large library of random peptide is required to screen CPPs, we used IVV for a comprehensive search of CPPs. IVV is a protein-nucleic acid complex that consists of mRNA and proteins as its translation product, and they are linked by a covalent bond through puromycin linker (Fig. 2A).

For the screening, we used IVV that expressed random peptide of 15 amino acids at the N-terminal of the POU sequence. To confirm whether we could use IVV for screening CPPs, we prepared Cy5-labeled IVV with a TAT sequence (TPV) and examined whether TPV translocated into cells. We also prepared Cy5-labeled POU-IVV (PV) that did not have a TAT sequence, as negative control.

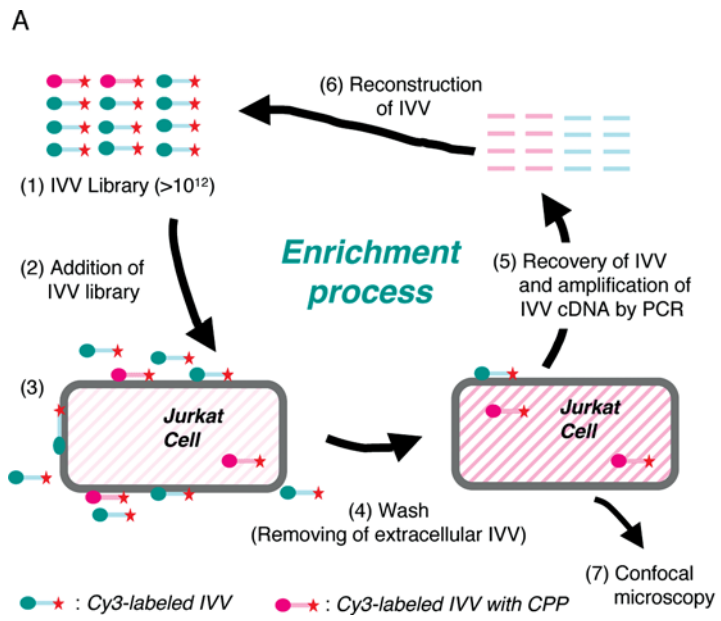
These IVVs were adjusted to 45 nM with PBS and added 54 μl/well to CHO cells that were inoculated to a 96-well plate. After 1 h of incubation, the cells were trypsinized and observed using a confocal laser microscopy. Intracellular fluorescence was detected from almost all cells when we added TPV. However, it was not detected when PV was added to the cells (Fig. 2B). This showed that the intracellular translocation activity was conferred on PV by the TAT sequence. These results suggest that IVV with random peptide could be used to screen CPPs.

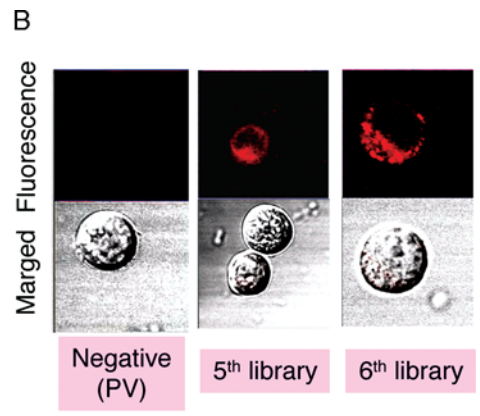
Screening from the random peptide libraries. Random peptide-POU-IVV (RPV) was constructed using random peptides instead of the TAT sequence in TPV. The design of the random peptide followed the methods of Husimi *et al* (29). We constructed Cy3-labeled RPV presenting 15 amino acids of random peptide (Diversity >10^{12}) in the upstream of the POU sequence. DNA sequences of 42 independent clones from RPV were determined and examined their amino acid composition. It was established that the random peptide libraries where each amino acid was expressed in a proportion from 2 to 8% had been built (data not shown).

For comprehensive screening, we used Jurkat cells, a suspension cell line derived from human T lymphocytes. The screening process of the CPPs is shown in Fig. 3A.

To confirm the intracellular transduction of IVV into Jurkat cell, they were analyzed after the addition of IVV using confocal laser microscopy from the third enrichment library. An intracellular fluorescence was observed in cells when the fifth and sixth enrichment library was added to cells (Fig. 3B). The populations of the Cy3 positive cells were 11% in the fifth enrichment library, and 68% in the sixth enrichment library (Fig. 3B). These results suggested that peptides with cell-penetrating activity were concentrated, as the enrichment processes were repeated. However, the intracellular fluorescence was hardly detected when the seventh and eighth enrichment libraries were added to cells (data not shown). There was a possibility that some CPPs were concentrated up to the sixth enrichment library.

Nucleic acid sequences of the IVV libraries were determined after the intracellular fluorescence was detected, and the amino acid sequences of the random peptide region were analyzed. About 70 clones were selected arbitrarily from each enrichment library from the fifth to eighth, and the nucleic acid sequences were determined. However, neither the convergence of sequences nor the known CPP sequences were detected. Fourteen redundant sequences were detected from 267 sequences, and one of them was found 3 times in the same library. The redundant sequences were found from all





Library	Percent of Positive Cells (%)
4th	0
5th	11
6th	68

Figure 3. Enrichment of CPPs using IVV. (A) Scheme of the enrichment process of CPPs with IVVs. (1) Construction of IVV libraries with random peptides. (2) IVV libraries were added to Jurkat cells. (3) IVVs with CPPs translocated into Jurkat cells. (4) IVV, which adhered to the Jurkat cell surface nonspecifically, was removed by trypsinization and washing. (5) The IVV genome was recovered from the Jurkat cells by PCR. (6) IVV libraries were reconstructed using the recovered IVV genome as a template. (7) Since IVVs with CPPs were concentrated, the fluorescence of Cy3 was detected in the cells. IVVs with CPPs were concentrated by repeating the above process. (B) Detection of internalized Cy3-labeled IVV using confocal laser microscopy. Upper layer, fluorescence image; lower layer, merged image.

A

Name	Criteria	Library	Sequence	Reference
No.3	Redundant sequence	7th	QHERSDRMECQDYNV	
No.4	Redundant sequence	8th	LRLQHLKHDRRIETR	
No.8	Helix structure	7th	ADDLENVNEGMHRIH	
No.9	Wheel structure	5th	RYGNRQYQFYEKSLL	
No.13	Wheel structure	5th	YDKECKSQAKMNDSY	
No.14	Wheel structure	6th	KLWMRWYSPTTRRYG	
No.15	Wheel structure	6th	DSLKSYWYLQKFSWR	
TAT	Positive control	—	YGRKKRRQRRR	27
TAT analog	Negative control	—	GRKKRRQPPQ	32

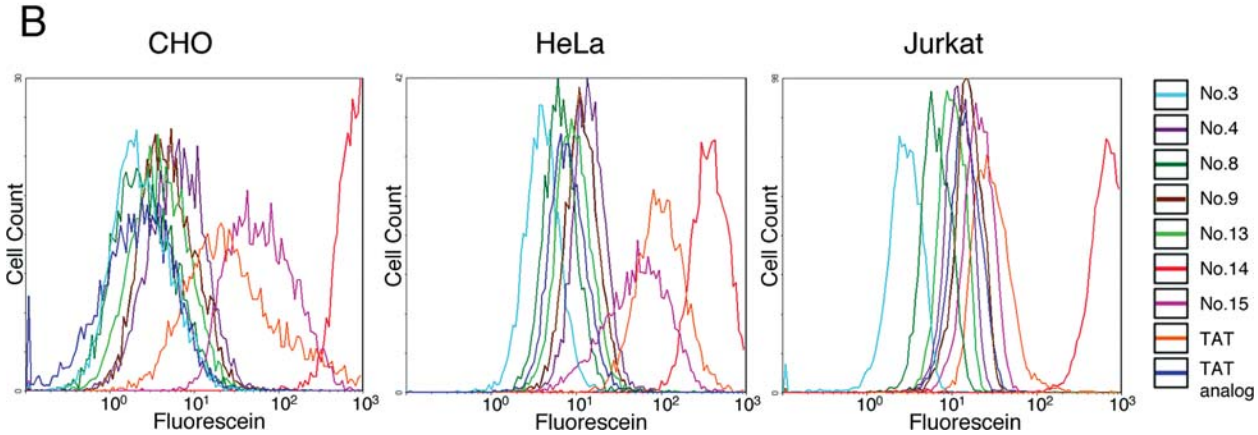


Figure 4. Intracellular translocation of candidate peptides. (A) Sequences of the candidate peptides and control peptides. Seven candidate peptides were selected using three criteria that are common features of known CPPs. (B) Flow cytometry analysis of cells incubated with fluorescein-labeled candidate peptides in CHO, HeLa and Jurkat cells. Three kinds of cells were treated with 50 μM fluorescein-labeled candidate peptides for 1 h at 37˚C and trypsinized. After washing, the intensity of the fluorescence from cells was detected using flow cytometry.

SPANDIDOS PUBLICATIONS

nt libraries; however, they were found in the same nt library, and the same sequence was not found between different enrichment libraries.

There are several possible reasons why the convergence of sequences may not be detected: the number of CPPs might be more than we thought and peptides that could easily bind to the cell surface might exist. It was thought that it might be difficult to get convergence of sequences with the enrichment processes using IVV; therefore, we analyzed the sequences of known CPPs to find their common features. Then, candidate sequences were selected from the enrichment libraries with the common features of known CPPs as criteria and we tried to confirm whether the candidate sequences possessed cell-penetrating activity.

Selection of CPP candidates. A secondary-structure prediction analysis of the amino acid sequences and helical wheel structure analysis (30,31) were performed, to examine whether there were common features among the known CPPs. We used TAT (27), Penetratin (4), HSV VP22 (10-12), MPG (13,14), Pep-1 (15), MAPs (16), HIV-1 Rev (17), pIsl (18), and WR peptide (19) as the known CPPs for setting the criteria to select the candidate sequences. It was expected that most known CPPs formed helical structures from the secondary structure analysis and their cationic amino acids existed in one side on the wheel from helical wheel projection analysis. Then, we set them as criteria for the selection of CPP candidates.

From 267 sequences derived from the fifth to eighth enrichment libraries, we selected five sequences that fitted these criteria, and two sequences from 14 redundant sequences that were concentrated by screening (Fig. 4A). We used TAT (YGRKKRRQRRR) (27) as a positive control and TAT analog (GRKKRRQPPQ) as a negative control (32). We prepared seven synthetic peptides, which were labeled on the amino terminal with fluorescein, and examined their cell-penetrating activity with mammalian cells (CHO, HeLa, and Jurkat cells) using flow cytometry and confocal laser microscopy. The results of the flow cytometry analysis are shown in Fig. 4B when fluorescein-labeled peptides at a concentration of 50 μM were added to cells.

As a result, two novel CPPs were found: one peptide (No. 14) whose cell penetrating activity was stronger than TAT in all cells, and the other peptide (No. 15) had cell-penetrating activity, which was better than TAT in CHO cells, although it was slightly inferior to TAT in HeLa and Jurkat cells. In CHO cells, both No. 14 and 15 showed cell-penetrating activity that were stronger than TAT. In particular, the cell-penetrating activity of No. 14 was >10 times that of TAT at a concentration of 50 μM. In HeLa cells, the cell-penetrating activity of No. 15 was inferior to TAT, but the intracellular translocation efficiency of No. 14 was 3.5 times of TAT at a concentration of 50 μM. In Jurkat cells, the cell-penetrating activity of No. 15 was slightly inferior to TAT, but No. 14 showed high cell-penetrating activity, >10 times that of TAT. We confirmed the cell-penetrating activity of both peptides in CHO cells using confocal laser microscopy. A precise intracellular transduction of peptides was detected at a concentration of 50 μM in both peptides and these peptides did not adhere to the surface and were translocated into CHO

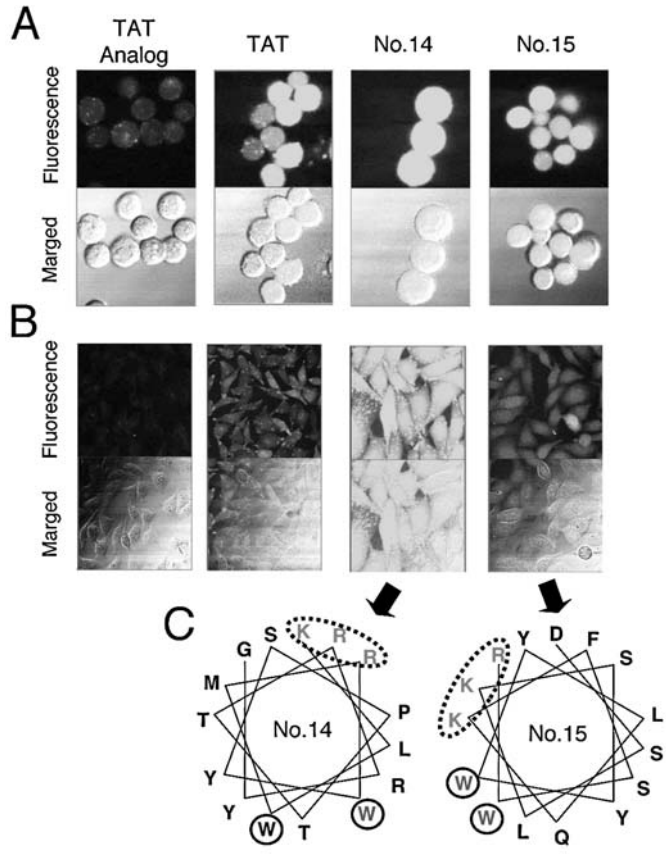


Figure 5. Transduction of No. 14 and 15 in CHO cells and their wheel structure. (A) CHO cells were treated with 50 μM fluorescein-labeled TAT analog, TAT, No. 14 and 15 peptides for 1 h at 37˚C and trypsinized. After washing, the cells were spread onto glass slides and covered with a cover slip. Their fluorescent intensity was measured using confocal laser microscopy. (B) CHO cells were cultured on an 8-well chamber slide (Nunc) for 48 h and treated with 50 μM fluorescein-labeled TAT analog, TAT, No. 14 and 15 peptides for 1 h at 37˚C. After washing, the fluorescence intensity from the cells was detected using confocal laser microscopy. (C) Wheel structure of No. 14 and 15. Dashed circles indicate Arg/Lys clusters and solid circles indicate Trp residues.

cells (Fig. 5A). Moreover, we confirmed their intracellular localization using confocal laser microscopy (Fig. 5B). It was clear that both peptides translocated into CHO cells at a concentration of 50 μM and existed homogeneously within cells.

These two peptides had some common features: Arg and/or Lys were sited on one-side of the wheel structures (Arg/Lys cluster), and they had two Trp residues. As such a structure resembled Penetratin, the possibility that a secondary amphipathic CPP-like Penetratin was concentrated in our screening was suggested. Also, both sequences were derived from the sixth enrichment library, and this fact agreed with the results that intracellular fluorescence was detected from 68% of Jurkat cells when the sixth enrichment library was added. These results suggested that if peptides with the same criteria were selected from the sixth enrichment library more CPPs might be obtained.

Therefore, we determined the nucleic acid sequences of ~940 clones arbitrarily from the sixth enrichment library and estimated the amino acid sequences of the random peptide. Successively, 46 candidate sequences were selected from

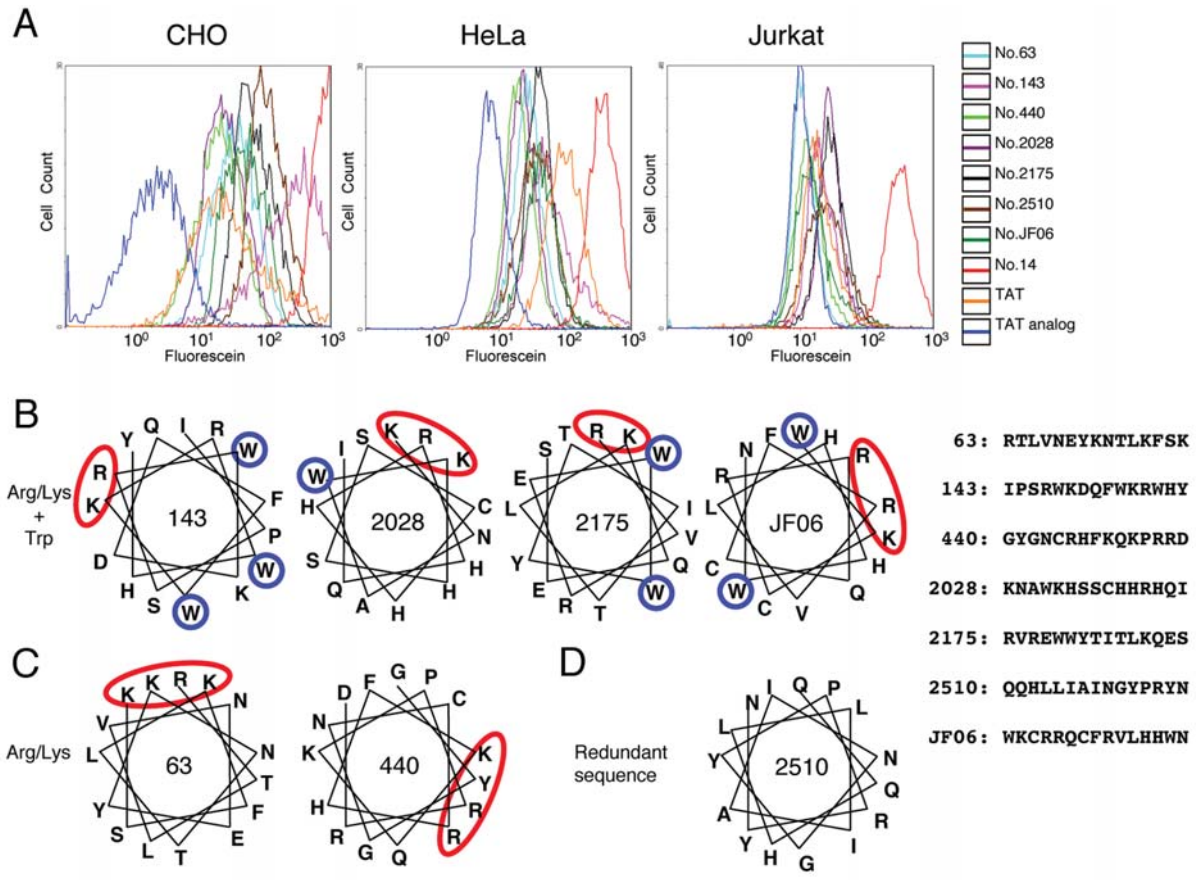


Figure 6. Intracellular translocation of isolated novel CPPs. (A) Flow cytometry analysis of cells incubated with seven fluorescein-labeled peptides in CHO, HeLa and Jurkat cells. The three kinds of cells were treated with 50 μM fluorescein-labeled peptides with candidate sequences for 1 h at 37˚C and trypsinized. After washing, the intensity of the fluorescence from the cells was detected using flow cytometry. (B) Four sequences had an Arg/Lys cluster and Trp. (C) Two sequences had an Arg/Lys cluster. (D) One sequence was redundant, it had only one Arg residue. Red circle, Arg/Lys cluster; blue circle, Trp residue.

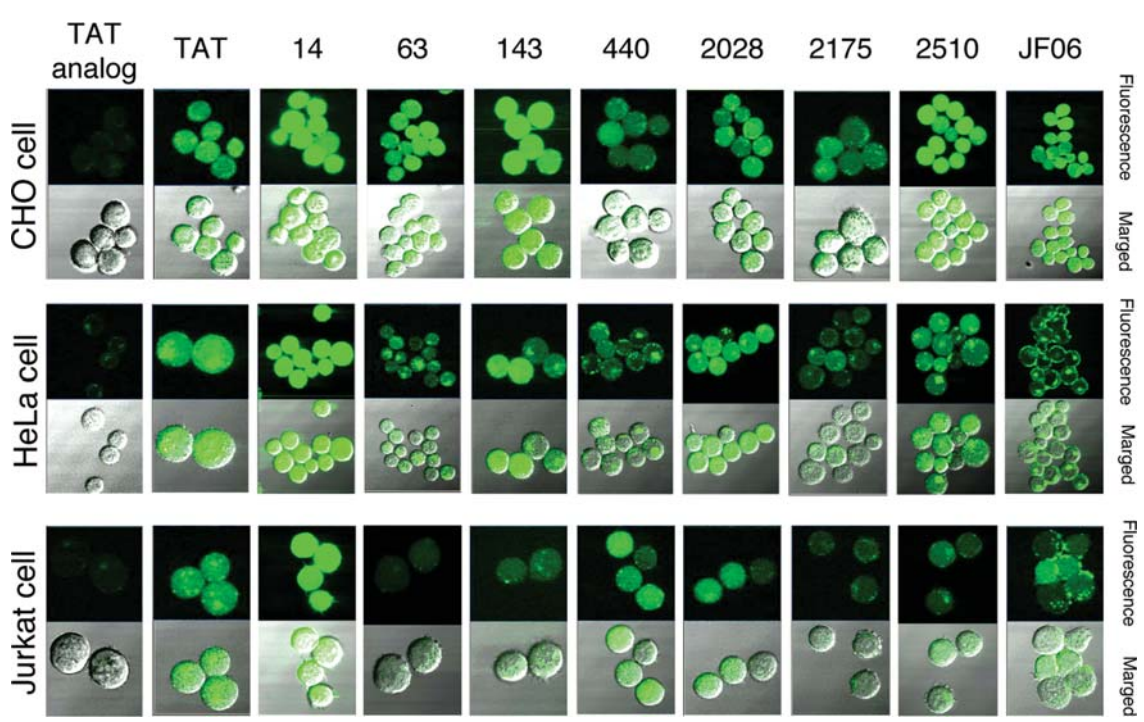


Figure 7. Detection of seven internalized fluorescein-labeled peptides in CHO, HeLa, and Jurkat cells using confocal laser microscopy. Three kinds of cells were treated with 50 μM fluorescein-labeled peptides with seven sequences for 1 h at 37˚C and trypsinized. After washing, the internalization of peptides was detected using confocal laser microscopy. Upper layer, fluorescence image; lower layer, merged image.

SPANDIDOS PUBLICATIONS

common features of No. 14 and 15. We prepared fluorescein-labeled peptides: 25 peptides, which had more than two Arg and/or Lys on one side of the wheel structure (Arg/Lys cluster) and included Trp; 17 peptides, which contained more than two Arg and/or Lys on one side of the wheel structure but did not contain Trp; one peptide that included three Trp; and three peptides that appeared to be redundant. We performed flow cytometry analysis in CHO cells to screen CPPs from these candidate peptides at a concentration of 50 μM (data not shown).

As a result, seven peptides had cell-penetrating activity that was the same or greater than TAT in CHO cells (Fig. 6A). Four peptides among the seven peptides had an Arg/Lys cluster on the wheel structure and Trp (Fig. 6B) and two peptides had an Arg/Lys cluster consisting of more than three Arg or Lys residues (Fig. 6C). When we compared the six peptides with an Arg/Lys cluster, there was a tendency for the peptides including Trp to show higher translocation efficiency. One remaining peptide was a redundant sequence (Fig. 6D). Although No. 2510, one of the redundant sequences, contained only one Arg, it had cell-penetrating activity. As the redundant sequence had cell-penetrating activity, it was suggested some other candidates for novel CPPs might exist in the sixth enrichment library.

The intracellular translocation efficiency of these seven peptides varied among cells when we examined their translocation into HeLa and Jurkat cells using flow cytometry (Fig. 6A). Although No. 143, 2028, 2175, and 2510 showed more translocation efficiency than TAT in Jurkat cells, there was no peptide that showed more translocation efficiency than TAT in HeLa cells.

Moreover, the intracellular translocation of these seven peptides was confirmed using confocal laser microscopy in three kinds of cells (Fig. 7). In No. JF06, although the cell-penetrating activity was found in both CHO and Jurkat cells, it was not found in HeLa cells because No. JF06 adhered to the cell surface. It is possible that an aggregation of the peptide might occur because of the inclusion of two Cys residues, and substituting other amino acids for Cys might enhance the cell-penetrating activity.

We were, therefore, able to identify some novel CPPs by the following processes, although the possibility that the intracellular translocation efficiency of the peptides varied among cells was suggested. A random peptide library was concentrated to some extent by screening using IVV. The number of the candidate sequences was reduced using the criteria that are common to known CPPs. Their cell-penetrating activity was examined using flow cytometry and confocal laser microscopy. We obtained No. 14, which had higher cell-penetrating efficiency than TAT in three kinds of cells. It was thought that No. 14 could be used as a novel CPP in place of TAT because No. 14 had cell-penetrating activity that far exceeded that of TAT in any of the cells studied.

Modification and alanine scanning of No. 14. Since No. 14 had better cell-penetrating activity than TAT in CHO, Jurkat and HeLa cells, we tried to modify No. 14 and performed flow cytometry analysis in CHO cells. Although the intracellular translocation of No. 14 at a concentration of

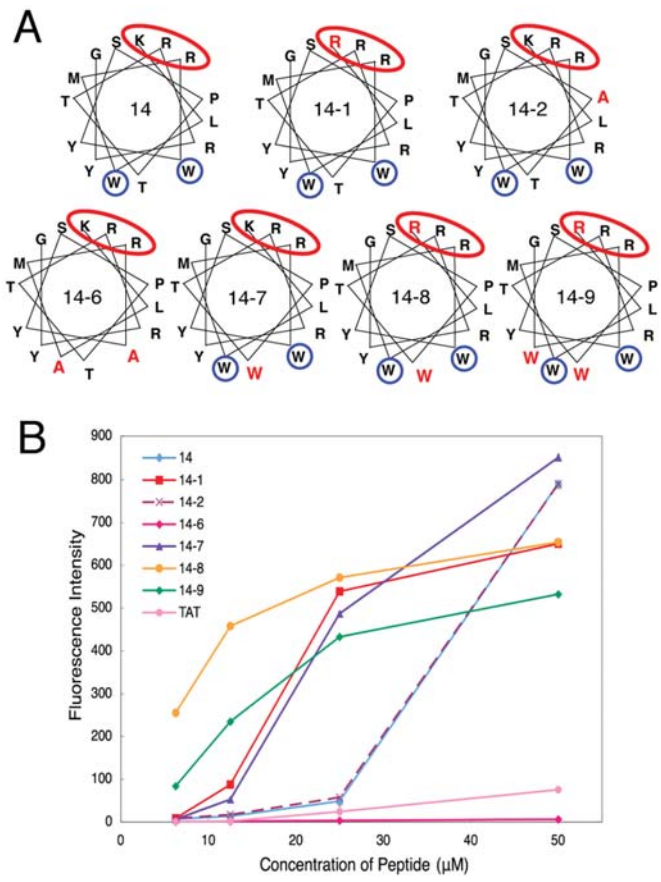


Figure 8. Analysis of the modified peptides of No. 14. (A) Wheel structure of modified peptides of No. 14. Modified amino acid residues are shown by red letters. (B) Results of flow cytometry analysis of modified peptides of No. 14 in CHO cells. Four concentrations (6.25, 12.5, 25 and 50 μM) of fluorescein-labeled modified peptides of No. 14 were added to CHO cells and incubated for 1 h. The cells were collected by trypsinization and analyzed using flow cytometry. Mode values are plotted for each concentration.

50 μM was >10 times that of TAT, it was almost the same level as TAT at a concentration of 25 μM (Fig. 8B). No. 14 would be a useful CPP if it could maintain high intracellular translocation efficiency even at a low peptide concentration.

We tried to modify No. 14 on the basis of its Arg/Lys cluster and two Trp (Fig. 8A). Because it has been reported that cationic amino acids play an important role in the intracellular translocation of CPP and Arg is more effective than Lys, Lys^{1} was substituted with Arg. As a result, the intracellular translocation at 12.5 and 25 μM concentrations improved drastically (14-1). When Pro^{9} was substituted with Ala (14-2), their translocation efficiencies did not change at all, in spite of our prediction that Pro might influence the secondary structure of No. 14.

Next, the importance of the two Trp residues was examined. As an intracellular translocation was not detected at all when we substituted Ala for the two Trp (14-6), it is suggested that these two Trp were essential to the intracellular translocation of No. 14. Moreover, when one Trp (14-7) was added, the translocation efficiency was the same as 14-1. Since the cellular uptake was expected to increase by combining these two modifications of 14-7 and 14-1, Lys^{1} of 14-7 was substituted with Arg (14-8). As a result, the intracellular

A

Name	Sequence
14	KLWMRWYSPTTRRYG
14-1	RLWMRWYSPTTRRYG
14-11	ALWMRWYSPTTRRYG
14-12	RAWMRWYSPTTRRYG
14-13	RLAMRWYSPTTRRYG
14-14	RLWARWYSPTTRRYG
14-15	RLWMAWYSPTTRRYG
14-16	RLWMRAYSPTTRRYG
14-17	RLWMRWASPTTRRYG
14-18	RLWMRWYAPTTRRYG
14-20	RLWMRWYSPATRRYG
14-21	RLWMRWYSPTARRYG
14-22	RLWMRWYSPTTARYG
14-3R	RLWMRWYSPTTRAYG
14-23	RLWMRWYSPTTRRAG
14-35	RLWMRWYSPTTRRYA

B

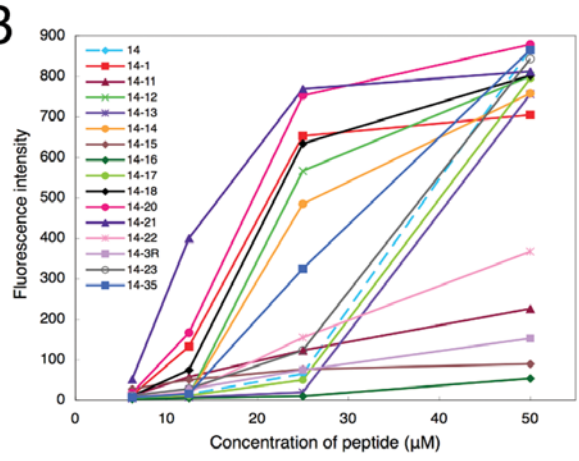


Figure 9. Alanine scanning of 14-1. (A) Alanine scanning products of 14-1. Substituted amino acid residues are shown by red letters. (B) Results of flow cytometry analysis of Ala substitutes of 14-1 peptide in CHO cells. Four concentrations (6.25, 12.5, 25, and 50 μM) of fluorescein-labeled Ala substitution peptides of 14-1 were added to CHO cells, which were then incubated for 1 h. The cells were collected by trypsinization and analyzed using flow cytometry. Mode values are plotted for each concentration.

A

Name	Sequence
14	KLWMRWYSPTTRRYG
14-1	RLWMRWYSPTTRRYG
14-24	RLLMRLYSPTTRRYG
14-25	RLFMRFYSPTTRRYG
14-26	RLIMRIYSPTTRRYG
14-29	RLVMRVYSPTTRRYG
14-30	RLYMRYYSPTTRRYG

B

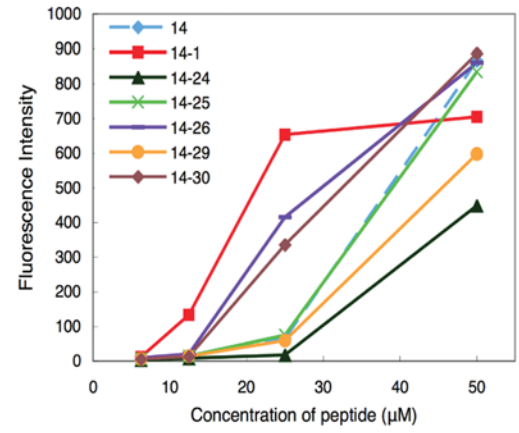


Figure 10. Substitution of Trp residues of 14-1 with other bulky or hydrophobic amino acids. (A) Sequences of Trp residue modified peptides of 14-1 peptide. Two Trp residues were substituted with bulky or hydrophobic amino acids that are indicated with red letters. (B) Results of flow cytometry analysis of the Trp residue modified peptides of 14-1 in CHO cells. Four concentrations (6.25, 12.5, 25, and 50 μM) of fluorescein-labeled modified peptides of 14-1 were added to CHO cells, which were then incubated for 1 h. The cells were collected by trypsinization and analyzed using flow cytometry. Mode values are plotted for each concentration.

translocation efficiency of 14-8 was approximately equal to 14-1 at 25 and 50 μM concentrations, but it remarkably improved at 6.25 and 12.5 μM concentrations. Moreover, when the number of Trp increased to four (14-9), the cellular uptake was lower than 14-8 because its solubility might be poor.

To examine the importance of the other amino acids of No. 14, we performed alanine scanning on the basis of 14-1 because the intracellular translocation at low peptide concentrations was considerably higher than that of No. 14 (Fig. 9). As a result, it was suggested that the modified peptides could be classified into three groups by intracellular translocation efficiency: a high efficiency type (14-12, 14-14, 14-18, 14-20, 14-21, 14-35) where the translocation efficiency was approximately equal to 14-1, an intermediate type (14-13, 14-17, 14-23) where the translocation efficiency was approximately equal to No. 14, and a low efficiency type (14-11, 14-15, 14-16, 14-22, 14-3R) where the intracellular translocation was lower than No. 14 (Fig. 9B). Although two Trp residues were thought important for the intracellular translocation of No. 14 from the result of 14-6, there was a difference in contributions to the intracellular translocation between Trp^3 and Trp^6. The cell-penetrating activity of 14-16 (Trp^6→Ala) was hardly detected, while the cellular uptake was detected at a concentration of 50 mM equal to 14-1 but hardly detected at concentrations <25 mM in 14-13 (Trp^3→Ala). From these results, it was thought Trp^6 was more important for cell-penetrating activity than Trp^3. Therefore, it is thought that the important amino acids to cellular uptake were in the order Trp^3<Arg^{12}<Arg^1<Arg^5, Arg^{13}<Trp^6.

Only 14-20 (Thr^{10}→Ala) and 14-21 (Thr^{11}→Ala), which belonged to the high efficiency type, had slightly higher cell-penetrating activity than 14-1, and Thr was substituted with Ala in both cases. Since 14-12, 14-14, 14-18, and 14-35, which belonged to the high efficiency type, did not show much difference in cellular uptake to 14-1, it was suggested that Leu^2, Met^4, Ser^8, and Gly^{15} were not so important for the translocation efficiency of 14-1. As the intracellular translocation efficiency of both 14-17 and 14-23 belonging to the intermediate type decreased by substituting Tyr with Ala, it was thought that Tyr was also important for the intracellular translocation.

Since Trp, which is a hydrophobic amino acid, was very important for cellular uptake from the result of 14-6, we examined whether the cellular uptake was maintained even if we substituted Trp with the other amino acids (Phe, Tyr, Ile, Val, Leu) that had a hydrophobic side chain (Fig. 10). However, there were no modified peptides superior to 14-1, a

SPANDIDOS PUBLICATIONS

acellular translocation was observed even when we substituted Trp with Ile (14-26) or Tyr (14-30). When we substituted Trp with Phe (14-25) or Val (14-29), their cell-penetrating activity reduced to almost that of No. 14, and when we substituted Trp with Leu (14-24), most of its cell-penetrating activity at a concentration of 25 μM was not observed. Among the hydrophobic amino acids, the effectiveness for intracellular translocation was in the order Leu<Val, Phe<Tyr, Ile<Trp.

Discussion

Until now, trials to screen CPPs from random peptides have been performed using methods such as the phage-display technique (23-26). However, effective screening of CPPs using phage-display is difficult because a bias may occur in expressed sequences, the library size is small ($<10^9$), phages bind on cells nonspecifically, and the sequences obtained by screening are not reproducible (33).

Since IVV allowed $>10^{12}$ molecules to be screened and the bias of the library was small (33,34), we attempted to find a better method for screening CPPs using the IVV system. However, we could not obtain a great enough convergence of the cell-penetrating sequences using IVV. This may be because the known CPPs including many cationic amino acids such as Arg or Lys and the fusion proteins tend to be insoluble. Since the proportion of cationic amino acids increased during the enrichment process of CPPs, the solubility of the concentrated peptides decreased when we reconstructed IVVs; therefore, an effective concentration of CPPs might be difficult to obtain. In addition, since peptides that contain many cationic amino acids have a high affinity with nucleic acids (35), it was thought that the effective presentation of peptides was not possible because a complex of nucleic acids and peptides was formed. However, as IVVs that translocated into cells increased up to the sixth enrichment library, we thought that peptides with cell-penetrating activity might be concentrated to some extent. Therefore, a better strategy for screening CPPs from these enrichment libraries was required. Hence, we examined the characteristics of known CPPs and tried to use the common features as the criteria for further screening.

It is not known that the homology of the sequences of known CPPs was high (36), but there are some studies that suggest that they form an alpha helix structure (30,36). We used secondary-structure prediction and helical wheel projection analysis for further screening. We selected candidates that were predicted to have a helical structure by the secondary-structure analysis, or identified by the helical wheel projection analysis to have Arg and/or Lys on one side of the wheel structure. When we chose peptides with these features from the sixth enrichment library and examined their cell-penetrating activity using their fluorescein-labeled peptides, peptides with more than two, preferably three, Arg and/or Lys on one side of the wheel structure, had cell-penetrating activity, and the activity was higher where Trp was included. These results suggest that the secondary amphipathicity of the peptides was very important to their cell-penetrating activity, as observed in Penetratin (37). In other words, it is thought that intracellular translocation might occur effectively with the interaction between the cell membrane and hydrophobic amino acids such as Trp after the binding of positive charges such as Lys or Arg to heparin sulfate on the cell surface (30,36). However, it was thought that there was an influence of amino acids other than Arg, Lys, and Trp because all peptides with such characteristics did not have cell-penetrating activity. Modifying the amino acids of No. 14, might possibly clarify which amino acids are important to its cell-penetrating activity. The Lys and Arg content of novel CPPs obtained in this study tended to be lower than in reported CPPs (36), it was thought that peptides containing many of these amino acids might not be concentrated because such peptides became insoluble during the enrichment process of IVV. Actually, when we examined the change of the amino acid composition of libraries during the enrichment process, the Arg content did not change, but the Lys content clearly decreased. No. 2510, a redundant sequence, had cell-penetrating activity although it contained only one Arg and no Trp residues. Since this peptide contained many hydrophobic amino acids, such as Ile, Tyr, and Leu, it was thought that it was concentrated because it had high affinity against cell membrane (38).

We investigated whether the cell-penetrating activity of No. 14 was improved by modifying them. The intracellular translocation activity of No. 14 was very high at a concentration of 50 μM, but it decreased at concentrations <25 μM, even so, it was still better than TAT. In a recent review (39), comparatively high concentrations of secondary amphipathic CPPs such as No. 14 were needed for intracellular translocation. We tried to modify No. 14 to improve its intracellular translocation efficiency at low concentrations. When the peptide concentration of No. 14 was high, the influence of cationic amino acids was significant and it bound easily to the cell membrane. However, it was thought that the effect suddenly decreased when the concentration was low, as shown in Fig. 8. Since the intracellular translocation in the low concentration remarkably improved by substituting Lys[1] with Arg, it was suggested that Arg had a stronger affinity for the cell surface than Lys. This was shown by a comparison between PenArg and PenLys (40).

The hydrophobic interaction was also important for cell-penetrating activity because the decrease of cellular uptake at low peptide concentrations improved by increasing the number of Trp residues. Since Trp residues tend to be buried in the cell membrane (41-43), the cell-penetrating activity of peptides with many Trp residues may not decrease at low peptide concentrations. In low peptide concentrations, two Trp residues of No. 14 were not enough to the effective translocation, but three Trp enhanced cellular uptake (14-7). However, when the number of Trp residues of 14-1 was increased to four (14-9), its intracellular translocation activity fell below that of 14-8. This may be due to a decrease in solubility of the peptide due to the increased number of Trp residues.

In addition, we carried out Alanine scanning on the basis in 14-1, because the intracellular translocation at low peptide concentrations was considerably higher than that of No. 14 (Fig. 8). Ala substitution products were classified into three types by intracellular transduction: a low efficiency type, an intermediate type, and a high efficiency type. As the intra-

cellular uptake remarkably decreased in the low efficiency type in comparison with 14-1, it was suggested that both cationic amino acids and Trp^6 residues were essential for the cellular uptake. Since the cellular uptake decreased by substituting Tyr or Trp^3 with Ala in the intermediate type, it was suggested that the cellular uptake in low peptide concentrations decreased by reducing hydrophobicity. Since the cellular uptake did not change even if Leu^2, Met^4, Ser^8, Gly^{15} were substituted with Ala, it is considered that they are not important for the cell-penetrating activity. It is possible that cellular uptake varies by substituting them with amino acids other than Ala.

Fig. 9 shows that when we substituted Trp with other amino acids with hydrophobic side chains, the modified peptides did not have intracellular translocation activity that was superior to Trp. However, there was some intracellular translocation activity of these modified peptides and the results of Ile substitution (14-26) and the Tyr substitution (14-30) were better than No. 14. In addition, the amino acids that had bulky hydrophobic groups were thought to be important for intracellular translocation because no cellular uptake was found in 14-6, although Ala had a hydrophobic group. This is supported by the observation that the cell-penetrating activity decreased in both 14-17 and 14-23 by substituting Tyr with Ala. Moreover, because cellular uptake was enhanced a little by substituting Thr, a hydrophilic amino acid, with Ala, a hydrophobic amino acid (14-20, 14-21), the amphipathic nature of the cluster of hydrophobic amino acids on the opposite side of the of the wheel structure to the Arg/Lys cluster was thought to be important for the cell-penetrating activity of 14-1. In other words, because 14-1 was bound to the cell membrane tightly by amino acids with bulky hydrophobic side chains such as Trp and Tyr after combining with glycosaminoglycans on cell surface through the Arg/Lys cluster (39), 14-1 might have strong cell-penetrating activity.

From the examination of No. 14 and its modified peptides, it is thought that we can predict which amino acids should be modified to improve the cell-penetrating activity of the other secondary amphipathic CPPs by plotting them on a Wheel structure to some extent. However, future studies are required to determine the cargo-carrying abilities of the novel CPPs identified in this study.

Acknowledgements

We received generous support from Dr Toru Nakajima and Dr Hiroaki Inoue (Mitsubishi Tanabe Pharma Corp.). Mr. Kouji Kanzaki (Mitsubishi Tanabe Pharma Corp.) gave insightful comments and suggestions.

References

1. Fawell S, Seery J, Daikh Y, Moore C, Chen LL, Pepinsky B and Barsoum J: Tat-mediated delivery of heterologous proteins into cells. Proc Natl Acad Sci USA 91: 664-668, 1994.
2. Mann DA and Frankel AD: Endocytosis and targeting of exogenous HIV-1 Tat protein. EMBO J 10: 1733-1739, 1991.
3. Vivès E, Brodin P and Lebleu B: A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. J Biol Chem 272: 16010-16017, 1997.
4. Derossi D, Calvet S, Trembleau A, Brunissen A, Chassaing G and Prochiantz A: Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. J Biol Chem 271: 18188-18193, 1996.
5. Fischer R, Fotin-Mleczek M, Hufnagel H and Brock R: Break on through to the other side-biophysics and cell biology shed light on cell-penetrating peptides. Chembiochem 6: 2126-2142, 2005.
6. Schwarze SR, Hruska HA and Dowdy SF: Protein transduction: unrestricted delivery into all cells? Trends Cell Biol 10: 290-295, 2000.
7. Shen WC and Ryser HJ: Conjugation of poly-L-lysine to albumin and horseradish peroxidase: a novel method of enhancing the cellular uptake of proteins. Proc Natl Acad Sci USA 75: 1872-1876, 1978.
8. Ryser HJ, Drummond I and Shen WC: The cellular uptake of horseradish peroxidase and its poly (lysine) conjugate by cultured fibroblasts is qualitatively similar despite a 900-fold difference in rate. J Cell Physiol 113: 167-178, 1982.
9. Mitchell DJ, Kim DT, Steinman L, Fathman CG and Rothbard JB: Polyarginine enters cells more efficiently than other polycationic homopolymers. J Pept Res 56: 318-325, 2000.
10. Dilber MS, Phelan A, Aints A, Mohamed AJ, Elliott G, Smith CI and O'Hare P: Intercellular delivery of thymidine kinase prodrug activating enzyme by the herpes simplex virus protein VP22. Gene Ther 6: 12-21, 1999.
11. Liu CS, Kong B, Xia HH, Ellem KA and Wei MQ: VP22 enhanced intercellular trafficking of HSV thymidine kinase reduced the level of ganciclovir needed to cause suicide cell death. J Gene Med 3: 145-152, 2001.
12. Cheng WF, Hung CF, Hsu KF, Chai CY, He L, Polo JM, *et al*: Cancer immunotherapy using Sindbis virus replicon particles encoding a VP22-antigen fusion. Hum Gene Ther 13: 553-568, 2002.
13. Morris MC, Vidal P, Chaloin L, Heitz F and Divita G: A new peptide vector for efficient delivery of oligonucleotides into mammalian cells. Nucleic Acids Res 25: 2730-2736, 1997.
14. Morris MC, Chaloin L, Méry J, Heitz F and Divita G: A novel potent strategy for gene delivery using a single peptide vector as a carrier. Nucleic Acids Res 27: 3510-3517, 1999.
15. Morris MC, Depollier J, Mery J, Heitz FF and Divita G: A peptide carrier for the delivery of biologically active proteins into mammalian cells. Nat Biotechnol 19: 1173-1176, 2001.
16. Hällbrink M, Florén A, Elmquist A, Pooga M, Bartfai T and Langel U: Cargo delivery kinetics of cell-penetrating peptides. Biochim Biophys Acta 1515: 101-109, 2001.
17. Futaki S, Suzuki T, Ohashi W, Yagami T, Tanaka S, Ueda K and Sugiura Y: Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. J Biol Chem 276: 5836-5840, 2001.
18. Magzoub M, Kilk K, Eriksson LE, Langel U and Gräslund A: Interaction and structure induction of cell-penetrating peptides in the presence of phospholipid vesicles. Biochim Biophys Acta 1512: 77-89, 2001.
19. Prochiantz A: Messenger proteins: homeoproteins, TAT and others. Curr Opin Cell Biol 12: 400-406, 2000.
20. Murriel CL and Dowdy SF: Influence of protein transduction domains on intracellular delivery of macromolecules. Expert Opin Drug Deliv 3: 739-746, 2006.
21. Kaplan IM, Wadia JS and Dowdy SF: Cationic TAT peptide transduction domain enters cells by macropinocytosis. J Control Release 102: 247-253, 2005.
22. Futaki S, Nakase I, Tadokoro A, Takeuchi T and Jones AT: Arginine-rich peptides and their internalization mechanisms. Biochem Soc Trans 35: 784-787, 2007.
23. Gao C, Mao S, Ditzel HJ, Farnaes L, Wirsching P, Lerner RA and Janda KD: A cell-penetrating peptide from a novel pVII-pIX phage-displayed random peptide library. Bioorg Med Chem 10: 4057-4065, 2002.
24. Mukai Y, Sugita T, Yamato T, *et al*: Creation of novel Protein Transduction Domain (PTD) mutants by a phage display-based high-throughput screening system. Biol Pharm Bull 29: 70-74, 2006.
25. Kamada H, Okamoto T, Kawamura M, *et al*: Creation of novel cell-penetrating peptides for intracellular drug delivery using systematic phage display technology originated from Tat transduction domain. Biol Pharm Bull 30: 218-223, 2007.
26. Nomura T, Kawamura M, Shibata H, *et al*: Creation of a novel cell penetrating peptide, using a random 18mer peptides library. Pharmazie 62: 569-573, 2007.
27. Nagahara H, Vocero-Akbani AM, Snyder EL, Ho A, Latham DG, Lissy NA, Becker-Hapak M, Ezhevsky SA and Dowdy SF: Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration. Nat Med 4: 1449-1452, 1998.

M, Husimi Y and Nemoto N: Solid-phase translation and protein fusion: a novel approach for folding quality control and direct immobilization of proteins using anchored mRNA. Nucleic Acids Res 34: E140, 2006.
29. Tabuchi I, Soramoto S, Ueno S and Husimi Y: Multi-line split DNA synthesis: a novel combinatorial method to make high quality peptide libraries. BMC Biotechnol 4: 19, 2004.
30. Ho A, Schwarze SR, Mermelstein SJ, Waksman G and Dowdy SF: Synthetic protein transduction domains: enhanced transduction potential in vitro and in vivo. Cancer Res 61: 474-477, 2001.
31. Bogoyevitch MA, Kendrick TS, Ng DC and Barr RK: Taking the cell by stealth or storm? Protein transduction domains (PTDs) as versatile vectors for delivery. DNA Cell Biol 21: 879-894, 2002.
32. Langel U: Pharmacology and toxicology: basic and clinical aspects. In: Cell-Penetrating Peptides: Processes and Applications. CRC Pr I Llc, pp3-7, 2002.
33. Roberts RW and Szostak JW: RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc Natl Acad Sci USA 94: 12297-12302, 1997.
34. Horisawa K, Doi N and Yanagawa H: Use of cDNA tiling arrays for identifying protein interactions selected by in vitro display technologies. PLoS One 3: E1646, 2008.
35. Meade BR and Dowdy SF: Enhancing the cellular uptake of siRNA duplexes following noncovalent packaging with protein transduction domain peptides. Adv Drug Deliv Rev 60: 530-536, 2008.
36. Wadia JS and Dowdy SF: Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer. Adv Drug Deliv Rev 57: 579-596, 2005.
37. Dom G, Shaw-Jackson C, Matis C, *et al*: Cellular uptake of Antennapedia Penetratin peptides is a two-step process in which phase transfer precedes a tryptophan-dependent translocation. Nucleic Acids Res 31: 556-561, 2003.
38. Thorén PE, Persson D, Esbjörner EK, Goksör M, Lincoln P and Nordén B: Membrane binding and translocation of cell-penetrating peptides. Biochemistry 43: 3471-3489, 2004.
39. Ziegler A: Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans. Adv Drug Deliv Rev 60: 580-597, 2008.
40. Amand HL, Fant K, Nordén B and Esbjörner EK: Stimulated endocytosis in penetratin uptake: effect of arginine and lysine. Biochem Biophys Res Commun 371: 621-625, 2008.
41. Caesar CE, Esbjörner EK, Lincoln P and Nordén B: Membrane interactions of cell-penetrating peptides probed by tryptophan fluorescence and dichroism techniques: correlations of structure to cellular uptake. Biochemistry 45: 7682-7692, 2006.
41. Brattwall CE, Lincoln P and Nordén B: Orientation and conformation of cell-penetrating peptide penetratin in phospholipid vesicle membranes determined by polarized-light spectroscopy. J Am Chem Soc 125: 14214-14215, 2003.
42. Henriques ST and Castanho MA: Translocation or membrane disintegration? Implication of peptide-membrane interactions in pep-1 activity. J Pept Sci 14: 482-487, 2008.