Chandrima Dey¹¹, Gloria Narayan¹¹, H Krishna Kumar¹¹, Manash P Borgohain¹¹, Nibedita Lenka² and Rajkumar P Thummer*¹

¹Laboratory for Stem Cell Engineering and Regenerative Medicine (SCERM), Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati (IITG), Guwahati 781039, Assam, India
²National Centre for Cell Science (NCCS), NCCS Complex, Savitribai Phule Pune University Campus, Ganeshkhind, Pune 411007, Maharashtra, India
*These authors contributed equally to this work

Dates: Received: 22 March, 2017; Accepted: 31 March, 2017; Published: 01 April, 2017

*Corresponding author: Rajkumar P Thummer, Indian Institute of Technology Guwahati, Guwa-

hati, Assam, India, Tel: +91-363-258-3208; E-mail: rthu@iitg.ernet.in

Keywords: Cell-penetrating peptide; Recombinant proteins; Protein transduction; Cell reprogramming; Induced pluripotent stem cell; Transgene-free

https://www.peertechz.com

Review Article

Cell-Penetrating Peptides as a Tool to Deliver Biologically Active Recombinant Proteins to Generate Transgene-Free Induced Pluripotent Stem Cells

Abstract

Delivery of biologically active recombinant proteins using cell-penetrating peptides is a useful tool for transduction of molecular cargo into cells. This protein transduction technology is useful to understand the biological function of a specific protein of interest. Hence, it has been recently employed by several groups to understand the molecular functions of stem cell-specific transcription factors in pluripotent stem cells. Pluripotent stem cells have an indefinite self-renewal capacity and can be directed to differentiate into any desired cell type of an adult organism. The groundbreaking discovery of induced pluripotent stem (iPS) cells has revolutionized the field of stem cell research due to its immense potential in in vitro disease modeling, drug screening and regenerative medicine. Although, virus-based gene delivery approaches commonly used to generate iPS cells are robust and highly efficient, however, these methods involve permanent genetic modifications due to viral integrations leading to malignant transformation. Therefore, a protein-based approach to deliver biologically active recombinant proteins in somatic cells to generate iPS cells is safe and will improve the prospects of these cells from bench-to-bedside. This review provides an overview of protein-based somatic cell reprogramming to generate transgene-free iPS cells and gives a glimpse of the bottlenecks associated with this technology.

Abbreviations

CPP, cell-penetrating peptide; ES, embryonic stem; FCS, fetal calf serum; GFP, green fluorescent protein; HDAC, histone deacetylase; iPS, induced pluripotent stem; MEF, mouse embryonic fibroblast; NLS, nuclear localization signal/sequence; OSKM, Oct4, Sox2, Klf4, c-Myc; p-iPS, protein-induced pluripotent stem; RFs, reprogramming factors; RFP, red fluorescent protein; SR, serum replacement; TAT, trans-activator of transcription; TEV, Tobacco etch virus; TLR3, Toll-like receptor 3; VPA, valproic acid

Introduction

Pluripotent stem cells such as embryonic stem (ES) and induced pluripotent stem (iPS) cells are elegant in vitro models for exploring cell fate, commensurate with the developmental hierarchy operational during early embryogenesis. Undoubtedly, their enormous potential in regenerative medicine has revolutionized the field of biology and medicine alike. iPS cells which closely resemble ES cells, are unspecialized cells that have two unique characteristics: self-renewal, the ability to proliferate indefinitely; and pluripotency, the ability to differentiate into numerous specialized cell types of all the three germ layers. These unique characteristics of iPS cells make them promising for autologous cell transplantation, that are devoid of immunological concerns and ethical problems unlike that in case of human ES cells. In addition, these cells can also be used for modeling a disease carrying genetic mutation(s) in a cell culture dish as well as for drug toxicity screening applications.

The remarkable discovery of derivation of iPS cells from adult somatic cells by ectopic expression of stem cell–specific transcription factors [1] has opened new avenues for therapeutic applications of stem cells. This depends on the availability of clinical grade human iPS cells that are not constrained by ethical, technical or immunological considerations. Several integrating (e.g. retro- and lentivirus) and non-integrating (adeno- and sendai-viruses, piggybac transposons, episomal and polycistronic plasmid vectors) approaches have been successfully used to generate iPS cells [2,3]. Integrating
approaches are efficient and robust in reprogramming somatic cells to IPS cells than non-integrating approaches; however, the clinical applicability of such IPS cells is restricted due to the integrated transgenes carrying the risk of insertional mutagenesis and tumor formation [4]. Although, non-integrating approaches reduce genome integration, the DNA of reprogramming cocktail of transcription factors and short vectors after removal can also result in insertional mutagenesis. To circumvent these problems, a virus- and DNA-free safe non-integrative approach of reprogramming is highly desirable. Till date, the promising approaches to generate clinical grade IPS cells are synthetic miRNAs [5,6], small molecules [7], microRNAs [8,9] and recombinant proteins [10–14], which eliminate the possibility of genome alteration by introduced foreign DNA. One of the most fascinating approaches is the transduction of biologically active cell-permeable recombinant proteins, as it provides precise control over dosage, time of application, and gives flexibility to toy with sequence of transcription factor stimulation.

Protein transduction via cell-penetrating peptides (CPPs) is an approach to deliver recombinant proteins and various small and large bioactive molecules into cells, and for this reason, this method has the possibility to replace virus-based gene delivery approaches. In this review, we summarize recent published reports on protein-based somatic cell reprogramming to generate IPS cells and highlight the bottlenecks of this technology.

Cell-penetrating peptides

Cell-penetrating peptides (CPPs), also known as protein transduction domains (PTDs) facilitate the delivery of recombinant proteins, peptides, oligonucleotides and other vital molecules into different mammalian cells. Most of the CPPs have been reported to be innocuous and do not impede the functionality of the biomolecules delivered across the cell membrane. Therefore, these CPPs have a wide application in many research areas, including cancer, apoptosis, cell cycle, neurobiology, signal transduction, and many others [15].

CPPs are a valuable class of short peptides with 3–30 positively charged (basic) amino acids, having an ability to translocate across the cell membranes with limited toxicity. Biologically active recombinant proteins when fused in frame with CPPs are translocated into cells via two different entry mechanisms: energy-independent direct penetration and energy-dependent endocytosis (Figure 1) [16]. However, the detailed CPP uptake mechanisms are largely unclear. This may vary substantially according to several parameters such as CPP, CPP length, CPP-cargo, cell types, cell density, concentration, incubation time and temperature, physicochemical parameters and other reasons [15]. Identifying and optimizing the right parameters will overcome the obstacles that might help in the easy uptake of these CPPs by the cell.

Cell penetrating capability of the trans-activator of transcription (TAT) protein of the Human Immunodeficiency Virus (HIV) was originally identified by two independent groups [17,18]. Subsequent studies demonstrated a short sequence in the basic domain of TAT protein (47–57), now recognized as TAT, can rapidly translocate into the cells more efficiently than full length protein and accumulate in the nucleus [19–21]. Later, Penetratin, a CPP comprising of 16 amino acids, was identified in the third helix of the homeodomain of transcription factor Antennapedia of the fruitfly Drosophila melanogaster [22–24]. Both TAT and Penetratin are rich in arginine and lysine residues imparting a high net positive charge at physiological pH. Recombinant proteins fused to these CPPs or synthetic CPPs rich in arginines can easily get internalized into cells. Generally, 6 to 12 poly-arginines exhibit transduction activity as CPPs [25,26], but as recently reported, even three oligo-arginines (3R) are sufficient to deliver transcription factors into the cells [27]. Since the first discovery more than 25 years ago, several short polycationic (Table 1) and amphipathic CPPs [28] have been identified subsequently. Importantly, even after cellular uptake these recombinant proteins are biologically active. This opens up new possibilities for intracellular delivery of various biologically active molecules into mammalian cells. The ability of these molecular delivery vehicles to translocate different small and large bioactive molecules into cells makes them promising candidates for transport of recombinant proteins, peptides, nucleic acids, liposomes, nanoparticles, medical imaging agents and drug delivery applications.

**CPP-mediated protein transduction to generate IPS cells**

Numerous integrating and non-integrating approaches have been employed to successfully derive IPS cells. Non-integrating approaches entirely replace gene delivery during the reprogramming process. These strategies are of major importance as they do not alter the host genome and generate clinically safe IPS cells. One of the most attractive approaches is the direct delivery of CPP-fused cell permeable recombinant proteins. This strategy has been employed by various studies to understand the role of growth factors and stem cell–specific transcription factors in recombinant forms to understand the biological function in pluripotent stem [29–34] and somatic [35] cells. In addition, this strategy has been recently used by few groups to induce ES cell gene expression program by delivering recombinant reprogramming proteins into the somatic cells to generate transgene-free IPS cells (Figure 2). This approach offers precise control over the dose and time of application, and gives complete flexibility to try and test different combinations of reprogramming factors to unravel their role at different stages during the reprogramming process. These CPP-mediated recombinant reprogramming proteins were delivered at a specific concentration, transduction cycle and exposure time on either mouse or human somatic cells in different studies to generate protein-IPS (p-IPS) cells (Table 2).

The very first study to report the generation of p-IPS cells was by reprogramming mouse embryonic fibroblasts (MEF) using a cocktail of recombinant cell–permeant reprogramming factors, Oct4, Sox2, Klf4 and c–Myc (hereafter, OSKM) [11], otherwise also known as “Yamanaka factors” originally identified by Takahashi and Yamanaka [1]. These reprogramming factors were tagged to CPP, poly-arginines (11R), at the C-terminal
Cell-Penetrating Peptides as a Tool to Deliver Biologically Active Recombinant Proteins to Generate Transgene-Free Induced Pluripotent Stem Cells

**Figure 1:** Cellular uptake mechanisms of CPP-fused recombinant proteins

CPPs are short peptide sequences comprising of arginine- and lysine-rich positively charged (basic) amino acids that enables cellular uptake of various molecular cargoes across the cell membrane to elicit a desired function. Recombinant biologically active proteins fused to these CPPs are translocated across the cell membrane via two different entry mechanisms: (A) energy-independent direct penetration, and (B) energy-dependent endocytosis. (A) Energy-independent direct penetration – this phenomenon of direct translocation through cell membrane encompasses different mechanisms (not shown in figure): pore formation, the membrane thinning model, the carpet-like model and inverted micelle formation [16]. The first step in all these mechanisms involves electrostatic interactions of negatively charged components of the membrane with the positively charged arginine head groups of CPPs. This elevates the hydrophobicity of the CPPs, thereby enhancing the potential of the CPPs to interact with lipid bilayer and stimulating the direct and rapid translocation of the peptide across the cell membrane. The successive mechanism of internalization depends greatly on the peptide sequence, peptide concentration and lipid composition in each model membrane study [16]. (B) Energy-dependent endocytosis – endocytosis is divided into two classes: phagocytosis and pinocytosis. Phagocytosis mainly involves cellular uptake of large particles such as other cells, viruses, or bacteria, whereas, pinocytosis involves cellular uptake of small particles. Pinocytosis occurs in all cell types. It is now generally believed that endocytosis is the main pathway involved in the CPP-mediated translocation mechanism. However, it is possible that various mechanisms operate under diverse conditions for all CPPs. Endocytosis of CPP-mediated transduction is further classified into macropinocytosis, clathrin-mediated endocytosis, caveolae-mediated endocytosis and caveolae/clathrin independent endocytosis (not shown in figure) [16], and results in formation of membrane-bound vesicles called endosomes. These endosomes act as a carrier to deliver the CPP-linked cargo into the cytosol. The endosomal release is facilitated by changes in pH and various other reasons, and eventually results in bursting of the endosomal membrane liberating the cargo. However, endosomal entrapment is also a common barrier and is a major challenge in efficient delivery of CPP-linked molecular cargo, which can be overcome by use of lysosomotropic agents, such as sucrose, chloroquine, ammonium chloride, or pH-sensitive fusogenic hemagglutinin-2 (HA2) subunit peptide [54]. In a cell reprogramming paradigm to generate iPS cells via CPP-mediated recombinant protein transduction, few reports also show that misfolded CPP-fused recombinant reprogramming proteins after endosomal release gets localized to cytoplasm and/or have a peri-nuclear region as observed in immunostained images [1, 35, 44, 45, 47-52]. Due to this, they are unable to enter the nucleus to activate its downstream target genes. Nevertheless, a small amount of the biologically active recombinant transcription factors enters the nucleus and binds to its target genes to activate the cell reprogramming machinery.

<table>
<thead>
<tr>
<th>Sl.No.</th>
<th>CPP</th>
<th>CPP origin</th>
<th>CPP sequence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Pep-1</td>
<td>Synthetic</td>
<td>Lys-Glu-Thr-Trp-Glu-Thr-Trp-Thr-Glu-Trp-Ser-Gln-Pro-Lys-Arg-Arg-Arg-Arg-Arg</td>
<td>[51]</td>
</tr>
<tr>
<td>6</td>
<td>MTDs</td>
<td>Synthetic</td>
<td>Ala-Ala-Ala-Pro-Val-Leu-Val-Ala-Ala-Leu-Leu-Leu-Leu-Pro-Val-Pro-Val-Pro-Val-Pro-Val-Pro-Val-Pro</td>
<td>[38]</td>
</tr>
</tbody>
</table>

The table comprises of the list of commonly used CPPs along with their origin, sequence and the respective references. These CPPs are fused to recombinant reprogramming proteins to deliver the latter into mammalian cells to elicit a desired molecular function.
Various integrative and non-integrative approaches have been used to generate IPS cells by overexpression of transcription factors in somatic cells [2,3]. CPP-mediated protein-based somatic cell reprogramming is a promising non-integrative approach to generate transgene-free protein-IPS (p-IPS) cells [10-14]. This strategy involves recopmentation of reprogramming proteins of stem cell-specific transcription factors, fused to CPPs that help in delivering these biologically active recombinant reprogramming proteins into mammalian cells. This approach does not involve any genomic integration, therefore, can be a good cell source for drug screening, disease modelling, tissue engineering, personalized cell therapy and to understand human developmental biology.

Table 2: Summarization of the optimized conditions used for effective CPP-based cell reprogramming.

<table>
<thead>
<tr>
<th>Sl.No.</th>
<th>CPP</th>
<th>Recombinant reprogramming proteins and supplements</th>
<th>CPP fused to N/C-terminal of reprogramming proteins</th>
<th>Concentration of fusion protein</th>
<th>Transduction cycles</th>
<th>Exposure period of reprogramming protein</th>
<th>Cell type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11R</td>
<td>OKS/OSKM, VPA</td>
<td>C-terminal</td>
<td>8 μg/ml</td>
<td>4</td>
<td>12 hours*</td>
<td>MEF</td>
<td>[11]</td>
</tr>
<tr>
<td>2</td>
<td>9R</td>
<td>OSKM</td>
<td>C-terminal</td>
<td>NP</td>
<td>6</td>
<td>16 hours*</td>
<td>HNF</td>
<td>[10]</td>
</tr>
<tr>
<td>3</td>
<td>TAT</td>
<td>O (SKC³)</td>
<td>C-terminal</td>
<td>100 nM</td>
<td>14</td>
<td>24 hours</td>
<td>MEF</td>
<td>[40]</td>
</tr>
<tr>
<td>4</td>
<td>TAT Penetratin</td>
<td>K (OSC²)</td>
<td>C-terminal N-terminal</td>
<td>50 nM</td>
<td>4</td>
<td>72 hours</td>
<td>MEF</td>
<td>[44]</td>
</tr>
<tr>
<td>5</td>
<td>TAT</td>
<td>S (OKC³)</td>
<td>C-terminal</td>
<td>400 nM</td>
<td>5</td>
<td>24 hours</td>
<td>MEF</td>
<td>[41]</td>
</tr>
<tr>
<td>6</td>
<td>11R</td>
<td>OSKM, poly I/C</td>
<td>C-terminal</td>
<td>200 nM 100 nM</td>
<td>6 14</td>
<td>24 hours 24 hours</td>
<td>HFF</td>
<td>[12]</td>
</tr>
<tr>
<td>7</td>
<td>TAT</td>
<td>OSKMN, VPA</td>
<td>N-terminal</td>
<td>5-50 nM</td>
<td>6 14</td>
<td>2 hours*</td>
<td>HFF</td>
<td>[13]</td>
</tr>
<tr>
<td>8</td>
<td>TAT</td>
<td>OSKMN</td>
<td>N-terminal</td>
<td>8 μg/ml</td>
<td>4</td>
<td>12 hours</td>
<td>MEF</td>
<td>[14]</td>
</tr>
<tr>
<td>9</td>
<td>MTDs</td>
<td>OSKML, Vit-C</td>
<td>N-terminal</td>
<td>500 nM 100 nM</td>
<td>7 3</td>
<td>3 hours</td>
<td>HDF</td>
<td>[38]</td>
</tr>
</tbody>
</table>

The table contains detailed information of all the studies that generated IPS cells using CPP-fused recombinant reprogramming protein(s) solely or in combination with small molecules or reprogramming factors in viral form.

end and expressed in *E. coli* [11]. Recombinant proteins (Oct4-11R, Sox2-11R, Klf4-11R and cMyc-11R) at concentration 0.5–8 μg/ml were observed to translocate into the nucleus and were stable up to 48 hours [11]. For this reason, OG2/Oct4-GFP (green fluorescent protein) reporter MEF cells were treated in four transduction cycles at every 48 hours at concentration 8 μg/ml for 12 hours [11]. However, this did not yield any IPS cell colonies [11]. The reason could be that large amount of these recombinant proteins got trapped inside the endosomes. Hence, only a small fraction of the recombinant reprogramming proteins could translocate into the nucleus. Concurrently, 1 mM Valproic Acid (VPA), a histone deacetylase (HDAC) inhibitor, was also added to the reprogramming proteins cocktail [11], which was earlier reported to significantly improve reprogramming efficiency [36,37]. Application of VPA prevents histone deacetylation by inhibiting HDACs, inducing an open chromatin structure that permits reprogramming factors to generate an ES cell–specific gene expression profile. Inclusion of VPA in the four reprogramming cocktail yielded only 3 GFP +ve IPS cell colonies, whereas only one GFP +ve colony in case of three reprogramming cocktail (Oct4-11R, Sox2-11R and Klf4-11R) at day 30–35 [11]. Like mouse ES cells, these p-IPS cells expressed all pluripotency markers and successfully tested for the pluripotency assays [11]. However, the reprogramming efficiency in this study was very low compared to viral approaches, primarily due to inadequate translocation of reprogramming factors to the...
nucleus, whereas the majority got confined to endosomes as observed in immunostaining images. Agents stimulating endosomal release of reprogramming proteins will help to improve reprogramming efficiency. Moreover, we hypothesize the masking of nuclear localization signal/sequence (NLS) in recombinant reprogramming proteins, and hence an additional NLS fusion would help them to easily translocate to the nucleus.

Simultaneously, another study reported successful generation of p-iPS cells by reprogramming human fibroblasts using OSKM fused to poly-arginines (9R) [10]. Preliminary experimental designs were based on examining whether the red fluorescent protein (RFP) when fused to 9 poly-arginines (RFP-9R) was able to penetrate into COS-7 cells and human newborn fibroblasts (HNF). Encouraging results were observed showing effective delivery of RFP-9R into both the cell types within few hours, even with the treatment of whole cell extracts [10]. Based on this data, stable HEK293 cell lines were generated expressing each of the four reprogramming factors fused with 9R [10]. These stable cell lines showed high expression of these proteins as determined by western blot analysis. Further treatment of HNF with the cell extracts from HEK293 resulted in efficient intracellular translocation of the reprogrammable proteins within 8 hours [10]. However, single dose of protein transduction did not generate any iPS cell colony, therefore, repeated treatments for six consecutive days for 16 hours was performed [10]. This eventually gave rise to p-iPS cell colonies after 8 weeks, but with low reprogramming efficiency (0.001%) and slow kinetics (8 weeks) [10]. These p-iPS cell colonies generated in this study were indistinguishable from human ES cells in morphology, proliferation and expression of pluripotency markers [10]. The low efficiency and kinetics could be due to treatment with whole cell extracts from the expressing cells thereby limiting the concentration of reprogramming factors delivered into target cells. Hence, treatment of fibroblasts with purified reprogramming factors might facilitate generation of iPS cells with higher efficiency and kinetics.

Reprogramming of mouse or human somatic cells to generate iPS cells using CPP-mediated recombinant protein transduction approach has remained highly inefficient [10,11]. Addressing this problem, a fascinating study recently reported that Toll-like receptor 3 (TLR3) signaling enables efficient induction of pluripotency due to the innate immune response to viral or modified mRNA [12]. However, this vital signaling is not activated in CPP-mediated protein transduction, which could be the probable reason for its low reprogramming efficiency. To activate this signaling, addition of TLR3 agonist (Polyinosinic-polycytidylic acid (Poly I:C)) to CPP-fused reprogramming cocktail (OCT4-11R, SOX2-11R, KLF4-11R, and c-MYC-11R) resulted in histone acetylation, down-regulating expression of various HDACs (HDAC 1,2,5,7) in CPP-treated human fibroblasts to generate p-iPS cells with high reprogramming efficiency and faster kinetics [12]. Therefore, TLR3 stimulation by Poly I:C is critical for activation of downstream target genes of the reprogramming factors and this TLR3 activation may trigger rapid and global changes in the expression of epigenetic modifiers that promote an open chromatin configuration and favor nuclear reprogramming [12]. Screening highly efficient novel TLR3 agonists that activate innate immune response would be desirable to improve efficiency and kinetics of CPP-mediated protein-based somatic cell reprogramming and further enhance the robustness of this technology.

Using macromolecule intracellular transduction technology (MITT), Lim et al. applied to deliver biologically active reprogramming factors OSKM and NANOG (Group A) or OSKM and LIN28 (Group B) into human skin fibroblasts to generate iPS cells [38]. This technology is based on the ability of hydrophobic macromolecule transduction domains (MTDs) to facilitate bidirectional transfer of recombinant proteins and peptides through the cell membrane. Seven MTDs were selected on the basis of high transduction ability. They were tagged to the reprogramming factors using a 6-histidine tag and nuclear localization signal/sequence (NLS) for nuclear localization [38]. However, neither Group A nor Group B gave rise to stably expandable iPS cell colonies, even though the colonies resembled morphologically to ES cells [38]. Hence, further investigation and optimization of this technology is required to deliver reprogramming factors to generate stably expandable human p-iPS cells.

Comparing p-iPS cell generation from two different genetic backgrounds, Nemes et al. attempted to generate mouse iPS cells using recombinant proteins from inbred (C57BL6) and outbred (ICR) strains [14]. They linked TAT with recombinant proteins to facilitate cell penetration. In addition, they fused NLS to promote nuclear delivery, thereby minimizing the quantity of proteins getting trapped in endosomes, and eventually preventing their degradation by hydrolytic enzymes. A short linker between TAT and NLS was also added to provide the necessary flexibility to the protein structure. MEFs from inbred and outbred strains were treated with recombinant TAT-fused OSKM for four times with an interval of 48 hours [14]. Nevertheless, only the outbred MEFs were successfully reprogrammed [14]. The possible reason, in addition to lower efficiency of protein transduction, could be the variation in the effect of modifier genes from two different genetic backgrounds [39]. Generated p-iPS cells of outbred strain were positive for pluripotency markers and qRT-PCR analysis expressed a small increase in the level of nuclear and membrane markers in the cell line [14]. Further to investigate iPS cell generation at different O 2 levels, they performed cell reprogramming at atmospheric and low (5%) O 2 levels [14]. Nevertheless, only the outbred MEFs were successfully reprogrammed [14]. This could be mainly due to low stability of recombinant reprogramming proteins and/or poor endosomal release of these proteins after cellular uptake, which ultimately leads to incomplete reprogramming.

To overcome these major bottlenecks, Thier et al., showed how media conditions could be optimized to achieve enhanced stability and endosomal release to maximize entry of biologically active protein into somatic cells [40,41]. In two separate studies, they have purified recombinant Oct4 and Sox2 fused to NLS, TAT and Histidine (His) tag at the C-terminus from...
E. coli, namely Oct4-TAT [40] and Sox2-TAT [41] respectively. Purified Oct4-TAT and Sox2-TAT showed limited solubility under cell culture conditions and their stability was found to be compromised [40,41]. Previously serum was used to increase the stability of the recombinant fusion proteins [42,43]. To improve Oct4-TAT and Sox2-TAT stabilization, serum or serum replacement (SR) supplemented with lipid–rich albumin fraction (Albumax) was tested [40,41]. Though SR increased the transduction efficiency than serum, but it compromised with the stability of the proteins. Moreover, the long–term activity of Oct4–TAT and Sox2–TAT was much improved when fetal calf serum (FCS), SR and Albumax were used together at a specific proportion. After detailed examination, 2.5% of FCS and Albumax (2.5%) for stability, and 7.5% of SR for transduction efficiency was found to be optimal [40,41]. This combination unblocks transduction inhibiting effects of FCS and provides better transduction efficiency and stability. Ultimately these supplements satisfy the criteria needed for the successful protein transduction media conditions for reprogramming. Even though the stability of recombinant proteins was improved, the endosomal entrapment of these proteins became another huge barrier to its successful delivery. Lysosomotropic agents that affects the endosomal integrity such as sucrose, was used to enhance the endosomal release of Oct4–TAT [40].

Apart from the enhancement of endosomal release, sucrose also facilitates the concentration of purified Oct4–TAT up to 500 μg/ml. Using “3 viruses + 1 protein” assay, where Oct4 virus was substituted by Oct4-TAT recombinant protein and applied to SKM infected MEF cells under optimal media conditions, it showed reprogramming efficiency in the same order of magnitude to OSKM viral transduction [40]. However, the reprogramming efficiency was one order lower in case of Sox2–TAT replacing Sox2 virus in reprogramming OSKM infected MEF cells [41]. This could be because no sucrose was used to facilitate release of Sox2–TAT recombinant protein from the endosomes. In addition, the efficacy of reprogramming directly depends on the concentration and exposure time period of Oct4–TAT and Sox2–TAT. Both Oct4–TAT– and Sox2–TAT–IPS cells expressed pluripotency markers and differentiated into all three germ layers [40,41]. Further animal studies are needed to be carried out for final confirmation of pluripotency in vivo.

Employing the same “3 viruses + 1 protein” assay, recombinant Klf4 coupled with CPP was used to generate iP cells from OSM infected MEF cells [44]. TAT and Penetratin were used to deliver recombinant Klf4 in this study and various combinations of CPPs, Klf4 and a reporter Discosoma red fluorescent protein (DsRed) (TRK, TR, TK, KT, PRK, PR and PK constructs) were evaluated for maximum reprogramming efficiencies. The TAT peptide fused with C-terminus of Klf4 (KT) showed reprogramming activity, unlike that fused with N-terminus (TK) [44]. Similarly, low reprogramming efficiency was observed, due to cationic property of Penetratin in which it was fused with N-terminus of Klf4 (PK). This could be one reason why Pan et al. also failed to generate human iP cells due to N-terminal direct fusion of TAT to reprogramming factors OSKM [45]. Direct linking of TAT/Penetratin to N-terminus negatively interferes with the clusters of acidic residues rich region of Klf4 which interacts with p300/CBP coactivators, components of the histone acetyltransferase complex [46]. Presumably, this results in masking of transcription activation domain of Klf4 due to the strong cationic/amphipathic property of TAT/Penetratin leading to poor reprogramming activity. Generation of the constructs with DsRed between TAT and Klf4 (TRK) exhibited maximum reprogramming efficiency as it unveiled the transcriptional activation domain of Klf4 [44]. However, this enhanced reprogramming efficiency was only due to a small amount of recombinant TRK entering the nucleus, as most of the TRK transduced was localized in cytoplasm and/or peri–nuclear areas. Initially it was assumed that cytoplasmic localization was due to inappropriate construct design. Later it was found that inadequate intracellular trafficking due to misfolding of recombinant Klf4 and limited endosomal release could be the cause. Hence, it represents one of the major bottlenecks for the recombinant protein transduction as reported and observed in immunostained images of different studies [35,45,47–52]. Overall, Klf4 transduction mediated reprogramming directly depends on dose, frequency, transduction cycles and exposure time. Here they reported that four cycles with 72 hours transduction interval of 50nM TRK proteins were required for an effective generation of iPS cells from somatic cells [44]. Essentially, Klf4 concentration used here was 3-fold lower than that used by Zhou et al., 2009. In addition, reprogramming time to generate iP cells was less with Klf4 protein transduction than that required by Klf4 substituted small molecule, Kenpaullone [44]. Importantly, TRK proteins can even further enhance the reprogramming efficiency if the major bottlenecks such as cytoplasmic/peri–nuclear localization and endosomal entrapment are thoroughly addressed.

Poly-arginines (9R/11R) and TAT have been commonly used by different studies for p-iPS cell generation [10–12,14,40,41,44]. To compare these two commonly used CPPs for cell reprogramming, Zhang et al. carried out protein–based reprogramming of somatic cells and compared the efficiency of TAT fused reprogramming factors (TAT–RFs) with poly–arginines (11R) fused reprogramming factors (11R–RFs) [13]. Incubation time and concentration for efficient transduction was tested and was found to be 2 hours and 25–400 nM respectively for three TAT–RFs, TAT–Klf4, TAT–Sox2 and TAT–c–Myc. Treatment of TAT–Oct4 at higher concentrations caused reduced transduction and cell death, and therefore 5–50 nM was found to be the suboptimal concentration. Notably, they demonstrated that TAT–RFs were transcriptionally more active than the respective 11R–RFs in activating their target genes. This could be mainly due to significant positive charges in 11R CPP which may negatively interfere with transcriptional activation domains of eukaryotic transcription factors often rich in acidic amino acids. This indicated that TAT CPP could be a better choice than 11R for protein–based iPS cell generation [13]. Further, the reprogramming activity of TAT–RFs were investigated using “3 viruses + 1 protein” assay. The addition of TAT–RFs increased the efficiency to multiple folds as compared to the untreated cells. This indicated that TAT–RFs can replace, but, were found to be less transcriptionally active when compared to their respective retrovirally transduced RFs [13]. For this reason, these TAT–RFs failed to reprogram.
human foreskin fibroblasts (HFF) to generate human iPS cell colonies [13]. Further improvements can be made in the design of their constructs by adding a linker or a reporter gene between TAT and reprogramming factor(s) providing flexibility. Tagging TAT at the C-terminal end would also help to improve the reprogramming efficiency with mouse fibroblasts and also to generate human p-iPS cells. Direct tagging of TAT to reprogramming factors at the N-terminal end has been observed to cause hindrance as reported earlier [14,40,44,41,52]. NLS fusion would also help to promote nuclear localization, reducing protein degradation by the acidic pH and digestive hydrolytic enzymes in endosomes. Human iPS cell colonies were obtained only after inclusion of TAT–Nanog and 125 μM VPA in the four reprogramming cocktail, having efficiency and kinetics at par with four retroviruses alone [13].

Strategies to improve protein transduction and biological activity

The ultimate aim of CPP-mediated protein transduction is to deliver desired biological activity of transduced proteins without affecting the viability of the transduced cells. To improve biological activity, Konno et al. demonstrated that removal of CPP from intracellularly delivered recombinant reprogramming proteins can help regain lost or reduced biological activity [53]. Presence of CPPs in reprogramming proteins is known to interfere with proper folding inside the cell and thereby decreasing the biological activity. In order to overcome this problem, they employed a site-specific protease derived from tobacco etch virus (TEV), to excise out the CPPs after the intracellular delivery of reprogramming proteins [53]. For this, the authors fused poly-arginines (11R) with reprogramming proteins, Sox2 and Oct4. TEV protease recognition site was inserted between 11R and Sox2 or Oct4 [53]. Further, they showed that deletion of CPPs from reprogramming proteins by TEV protease after delivery into the cells substantially restored their biological activity, whereas the presence of CPPs compromised their functions [53]. Moreover, they discovered that TEV proteases did not affect the endogenous proteins; therefore, ES cell identity remained unaltered [53]. This study clearly demonstrated that CPPs interfered with the biological activity of reprogramming proteins, which can be restored by intracellular excision of CPP using a specific protease. The application of this strategy in derivation of p-iPS cells is anticipated to improve both reprogramming efficiency and kinetics, and work efficiently with cell lines which are challenging to reprogram.

Recently, Ryu and colleagues, employed a novel CPP of non-viral origin to increase efficiency of intracellular delivery of reprogramming factors by overcoming the issues related to their solubility and stability [52]. They engineered *Bombyx mori* 30Kc19 CPP from silk and tagged it to the C-terminus of the reprogramming proteins, generating Oct4–30Kc19, Sox2–30Kc19, Klf4–30Kc19, c-Myc–30Kc19 and L-Myc–30Kc19 constructs [52]. They discovered that 30Kc19 CPP conjugated reprogramming factors exhibited improved soluble expression and stability in vitro. In contrast, unconjugated reprogramming factors aggregated in the inclusion bodies [52]. The authors concluded that enhanced solubility of recombinant reprogramming proteins was due to the C-terminus fusion of the CPP [52]. Similarly, increased stability of reprogramming proteins was attributed to the shielding effect of 30Kc19 [52], without addition of lipid–rich albumin supplements or serum to the cell culture media. Further, on incubation with human dermal fibroblasts (HDFs), they found that the rate of intracellular delivery and transcriptional activity of these reprogramming proteins were remarkably enhanced [52]. However, immunofluorescence images showed that still a large amount of proteins were trapped in endosomes and also observed in the peri–nuclear region [52]. These proteins can be released from endosomes using lysosomotrophic agents such as chloroquine, ammonium chloride, sucrose, or pH-sensitive fusogenic hemagglutinin–2 (HA2) subunit peptide [54]. Treatment with these agents will result in the endosomal release of these reprogramming proteins increasing the amount of these proteins entering the nucleus, further improving the transcriptional activity. Many groups have also demonstrated delivery of CPP-cargo into the nucleus with an additional NLS to overcome peri–nuclear entrapment [14,38,40,41,50]. Therefore, 30Kc19 can be used as a potential non-toxic carrier of non-viral origin enhancing solubility, stability and transcriptional activity of recombinant reprogramming proteins. Further role in generation of p-iPS cells by delivering these recombinant reprogramming proteins in active form fused to this CPP is yet to be investigated.

As discussed earlier and highlighted in Figure 3, several bottlenecks associated with CPP-mediated protein transduction to deliver reprogramming factors are required to overcome for successful generation of p-iPS cells. The studies listed above addressed only few of these major bottlenecks in their experimental design. As a result, these studies failed to generate stable p-iPS cells using only OSKM reprogramming factors [11,14,38,45] or resulted in poor reprogramming efficiency [10]. Subsequently, inclusion of mutagenic small molecules such as VPA [11,13] and/or additional reprogramming factors such as Nanog [13] has been used to generate p-iPS cells. A comprehensive study addressing all the bottlenecks is required to successfully generate p-iPS cells with high reprogramming efficiency and faster kinetics, at par with robust and highly efficient viral approaches. Combinatorial approaches such as using CPP with chemical platform that would activate/inhibit specific signaling cascade or vitamin C for curtailing cellular senescence [55] and also depletion of Mbd3 [56] may provide plausible insight and solutions in this regard. This will pave the way for p-iPS cell implications in basic research and biomedical applications. Moreover, the implications of CPP may go well beyond iPS cell generation with lower risk of teratoma formation [11] to their directed differentiation and gene editing [57].

Conclusion

iPS cells represent a valuable source to derive patient-specific cells for disease modeling, drug toxicity screening and cell replacement therapy applications. Introduction of

exogenous genes via viral vectors, plasmids and nucleic acids in somatic cells to generate iPS cells cannot completely eliminate the possibility of genomic integration. Therefore, novel safe approaches should be explored to generate clinically safe iPS cells. In this review, we have highlighted reports that demonstrated the potency of CPP-mediated protein transduction technology, which involves delivery of transcription factors in recombinant forms into the somatic cells to generate iPS cells. Delivery of this CPP-mediated core stem cell-specific recombinant proteins can regulate gene expression, and function transiently without any genomic alteration. In addition, this protein transduction technology will help us to understand the mechanism of cell reprogramming phenomenon. This is attainable as it offers greater control over concentration and time of application. Moreover, it provides ample flexibility to experiment various combinations of reprogramming factors with small molecules using a sliding window approach to unravel their role at different stages during cell reprogramming. Importantly, this technology does not involve any genomic integration and induce an innate immune response. Henceforth, it also permits us to generate clinical grade transgene-free iPS cells for biomedical applications. Although several questions remain to be answered, and existing limitations on the use of CPP-mediated protein transduction are yet to overcome, it is clear that this evolving realistic protein transduction technology has much to offer in a cell reprogramming paradigm. This technology undoubtedly reinforces the significant role of CPPs and its immense potential in delivery of reprogramming proteins to generate transgene-free iPS cells. In conclusion, this technology is devoid of any possible risk of permanent genomic integrations and/or mutations and therefore will translate iPS cell related applications from laboratory to the clinic.

Acknowledgment

We thank all the members of Laboratory for Stem Cell Engineering and Regenerative Medicine (SCERM) for their excellent support. This work was supported by grant funded to Rajkumar P Thummer and Nibedita Lenka by North Eastern Region – Biotechnology Programme Management Cell (NER-BPMC), Department of Biotechnology, Government of India (BT/PR16655/NER/95/132/2015) and also supported by Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India (Early Career Research Award; ECR/2015/00193) funded to Rajkumar P Thummer.

References
