MedChemComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/medchemcomm

Peptide HIV Fusion Inhibitors: modifications and conjugations

Wei Liu,[a,b] Jianjun Tan,[*a] Mohammadreza Mohammadzad Mehryar,[a] Zhiping Teng,[b] and Yi Zeng[*a,b]

[a]*College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China*
[b]*State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Viral disease control and prevention, Chinese Centre for Disease Control and Prevention, Beijing 100052, China*

Abstract

HIV fusion inhibitors are a group of virus entry preventing drugs aiming at membrane fusion. Several peptide drugs are screened with potent capacity that blocks virus-host cell fusion efficiently, and researchers have begun to focus on peptide drug design. However, so far T-20 is the only fusion inhibitor that has been approved by FDA and utilised in treating therapy-experienced AIDS patients. The application of conventional peptide drugs is often limited by their short in-vivo half-life and heavy dose subcutaneous injection. Thus, it is necessary to design new type peptides or shift the existing peptides in a modified form to reach lower EC_{50} value and get longer in-vivo half-life. Here, we summarise the inhibitory mechanisms of these peptide fusion inhibitors designed with anti-HIV potency. And we further discuss the recent achievement in these peptide-derivatives. Several approaches have been applied to optimise the peptide properties and we provide examples of successful peptide conjugation strategies to provide hints for additional peptide therapeutic design.

Keywords: HIV, peptide fusion inhibitor, peptide conjugation.

INTRODUCTION

The entry of human immunodeficiency virus (HIV) is complex. It is triggered by the recognition between envelop protein (Env) on the virions and the CD4 receptor[1] on the cell surface. Env is a trimeric glycoprotein composed of three copies of gp120/gp41 heterodimers that origin from precursor protein gp160. The proteolysis of gp160 causes the formation of gp120 and gp41 in a noncovalently associated manner. Gp120 is the surface subunit, responsible for attachment to target cell receptors, whereas gp41 is anchored into the viral membrane as a transmembrane protein to mediate fusion. When gp120 subunit interacts with CD4, the Env spikes aggregate, and the conformational changes of gp120 appear which further drive the recognition by a coreceptor like CXCR4[2] or CCR5[3]. The inner domain of gp120 interacts with gp41, and its spatial rearrangement results in the dissociation of the gp120-gp41 complex and the rearrangement of the gp41 domains in preparation for catalysing membrane fusion.

HIV fusion inhibitors are designed for preventing the entry of the virus into the host cells by block the gp41 conformational transition. T20 is a representative peptide fusion inhibitor and has been put into clinical use. Nevertheless, the high-dose injection and rapid plasma clearance rate render it limit in salvage treatment for cocktail resistant patients[4]. Further HIV fusion inhibitors designed by amino acid substitution are slightly more stable. However, this traditional method cannot additionally increase their inhibitory potency and prolong the serum half-life, and the affinity threshold for maximum activity in sensitive strains has also been reached[5].

Nowadays, neotype HIV inhibitors with enhanced efficiency and prolonged half-life are constructed based on the bio-conjugate techniques. For example, PC-1505, an albumin-conjugated C34 peptide, was designed for less frequent dosing and peptide requirement [6]. C34-cholesterol conjugate exhibits 50-fold more potent than the peptide *per se* according to the IC_{50} in neutralising HXB2 strain [7] [8]. These various modifications and conjugations diversify the forms in developing new drugs and provide us new strategies in drug design.

In this review, we present selected advanced bio-conjugated forms of peptide fusion inhibitors to provide hints not only for the novel anti-HIV peptide drug design but for other protein drug modification as well. These modifications for peptide fusion inhibitors dramatically influence their biophysical and biochemical properties allowing less frequent dosing.

GP41 AND THE MEMBRANE FUSION

HIV entry is a complex process briefly contains four steps: 1) virus-cell attachment; 2) gp120 and CD4 receptor binding; 3) gp120 and coreceptor binding; 4) gp41 mediated membrane fusion.

Understanding the structure of gp41 at each step during HIV fusion is of fundamental significance in fusion inhibitor design for preventing further infection. Gp41 is a 345-residue polypeptide from 512 to 856 of env according to HXB2 HIV-1strain. It can be separated into several functional domains: the fusion peptide (FP), the fusion peptide proximal domain (FPPR), the N-terminal helical heptad repeat region (N-HR), the immunodominant loop region, the C-terminal helical heptad repeat region (C-HR), the membrane proximal region (MPER) and the transmembrane region (TM) as well as a cytoplasmic tail (CP). Gp41 extracellular ectodomain is composed of FPPR, N-HR, the loop region, CHR and MPER. NHR and CHR form the gp41 core. During the fusion reaction, gp41 transits into a pre-hairpin intermediate to open its hydrophobic ectodomain and allow the N-terminal fusion peptide to insert into host cell membrane. In the post-fusion state, the C-terminal helices of the gp41 pack around the N-terminal helices in an anti-parallel fashion, forming a structure called the six-helix bundle (6-HB)[9] [10].

Knowledge of gp41 at high resolution in any pre-fusion step remains to be further elucidated. Fortunately, a study utilising cryo-electron

*Address correspondence to those author at the College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, P. R. China; Tel: (+86)-10-67392724; Fax: (+86)-10-67392837; E-mail: tanjiajun@bjut.edu.cn (Jianjun Tan); Tel: (+86)-10-67391441; Fax: (+86)-10-673962001; E-mail: zengyicdc@sina.com (Yi Zeng)

microscopic captured an activated intermediate whose formation precedes that of pre-hairpin structure[11]. At ~9Å resolution of a cleaved, soluble version of trimeric Env, a structural feature of this intermediate is the open Env conformation with a three-helix motif, composed of α-helical segments that derived from highly conserved, non-glycosylated N-terminal regions of the gp41 trimer. The three gp41 N-HR in this novel, activated Env conformation are held apart by their interactions with the rest of Env, and are less compactly packed than in the post-fusion, six-helix bundle state[12]. During the activated intermediate, the vulnerability of gp41 begins to appear, hence the highly conserved gp41 core regions are exposed and the gp41 N-HR without any glycosylation site shows a larger opening at the side of the target cell and greater burial on the side of the viral membrane. This gp41 arrangement accounts for the great differences observed in the accessibility of different ends of the gp41 N-HR to the large peptides.

Pre-hairpin structure is suggested to be a putative intermediate that FP inserts into the cell membrane and then forms a bridge between virus and target cell. This state is believed to maintain at least 15 min[13 14]. It is more vulnerable and accessible to neutralising antibodies and peptide fusion inhibitors targeting N-HR or C-HR. The exposure of highly conserved C-HR and N-HR interface during the organization of pre-hairpin intermediate render gp41 an ideal target. The three hydrophobic pocket structures on the surface of the central trimeric coiled-coils of N-HR is composed of 11 highly conserved residues[13 15 16]. C-HR is more variable, but the residues retaining the interaction with N-HR showed highly conservative. These conservative residues are relatively slow in developing mutation, namely the virus effortlessly to escape from drugs aiming at these targets. Additionally, the self-binding feature of gp41 suggests that peptides derive from N- and C- peptide regions are potent inhibitors to interrupt the 6-HB formation through mimic the native peptide functions. Further trials substantiate this view.

Peptide fusion inhibitors

Most of the HIV-1 fusion inhibitors comprise the C- or N- terminal residues of gp41, which originally developed by synthesized peptide to mimic various regions of gp41. They are designated C- or N- peptides respectively. They interact with their counterpart residues of gp41 to form heterogeneous 6-HB and abrogate the membrane fusion. C-peptide inhibitors exhibit more potent than N-peptide inhibitors whereas N-peptides show a considerable ability to abduct the C-peptide resistant viruses.

T20 and other C-peptide fusion inhibitors

T20 (Fuzeon, Enfuvirtide) was licensed by the US Food and Drug Administration (FDA) on March 13, 2003 as a novel antiretroviral medicine. Though the discovery of T20 is several years earlier than the publication of 6-HB structure, the further observation of high-resolution structures of the gp41 core suggests a putative mechanism of T20 and promotes the development of following variants. This hypothesis is that T20 could block the formation of the 6-HB to interfere the formation of the hairpin structure by competitively binding N-HR, suggesting that peptides derive from C-HR or N-HR could be potent to prevent HIV fusion. According to this conjecture, synthesised peptides confirm this view.

T20 is a 36-mer peptide derived from Tyr638 in C-HR to Phe673 in the MPER according to the Env_{HXB2} sequence number, and was predicted to function mainly through binding the N-HR. The detailed anti-viral mechanism of the T20 is not very clear and controversies still exist. Its inhibitory potency not merely depends on the interaction with N-HR. In fact, it has been demonstrated that T20 neither interact with N36, a peptide derived from the HIV gp41 N-HR, to form 6-HB, nor inhibit trimer-of-hairpins formation. The C-terminus MPER derived region of T20, which contains the gp41 lipid binding domain seems that in addition to interfering of 6-HB formation, can potentiate its ability in a rather complex way[17]. It is suggested to have an enhanced membrane binding affinity compared with other peptides like C34 [18]. Further findings indicate that the C-terminal region of T20 functions mainly as an anchor to the cell membrane[19]. Through the WNWF sequence of MPER derived region, the peptide tail of T20 acts on the cell membranes, hence stabilises the binding to gp41. WNWF sequence is crucial in maintaining the inhibitory potency of T20. In the absence of WNWF, T20 would lose its biological activity. The Mutation of WNWF to ANAA had no effect on affinity to the 5-helix, an artificial molecular mimic of the entire gp41 N-HR coiled-coil, suggesting that these amino acids has no contribution in T20 binding to the gp41 N-HR directly [17]. And this substitution diminishes its antiviral ability. However, when the T20 mutant anchored into cell membrane, it showed same inhibitory potency with original T20[20]. Similar observation was found in remarkably reduced activity of truncated T20 due to missing of the lipid binding domain, but the potency could be restored with a fatty acid chain conjugation (DP-C16)[19]. These findings indicate T20 needs both C-HR and MPER derived residues to block HIV-1 entry, and WNWF motif may play a key role in interaction with membrane. However, binding affinity to N-HR does not seem to be the only determinant that decides its antiviral potency[17]. T20 with T639I substitution in C-HR derived region bound 5-helix with significantly higher affinity but did not exhibit higher potency against wild-type HIV-1. Some evidence has further demonstrated that the interaction between T20 and GIV motif of N-HR is not the major contribution towards T20 inhibitory activity[21], although GIV motif was assumed as a target region for T20. A predictive view point in consistent with this observation is that the T20 binding may be stabilised by an interaction of WNMF motif with residues from the gp41 fusion peptide. Another distinct suggestion is that T20 also interacts with the viral membrane and blocks the formation of the membrane fusion pore.

C34 shares a section of overlap sequence with T20 but contain the N-HR cavity binding residues. It could interact with N36 to form a stable 6-HB. X-ray crystallographic studies showed that three C34 peptides bind to the cavities of three N36 molecules formed trimeric core in an anti-parallel manner[13]. Residues Trp628, Trp631, and Ile635 (WWI) of C34 peptide are inserted into the pocket that formed by a cluster of residues in the N-HR coiled-coil [22]. This "WWI" motif is critical for the N-HR and C-HR inter helical interactions and proved to be essential for the peptide's anti-HIV potency. C34 shows very effective potency in interfering 6-HB formation on account of the "WWI" motif, whereas T20 is a weak gp41 core formation inhibitor [18]. This distinction suggests that despite C34 contains most residues of T20; they function in a different manner. Further studies suggest C-peptide fusion inhibitors contain different functional domains may inhibit HIV-1 entry in distinct mechanisms [23]. Overlapping peptide designs demonstrated that peptide which lacks residues from both N-HR pocket binding (PBD) and lipid

binding domains (LBD) has negligible anti-HIV-1 effects[23]. Moreover, addition of PBD and/ or LBD to the HR sequences resulted in significant improvement of anti-HIV activity[24].

T20 and C34 have different mechanism because they contain PBD and LBD separately. Distinguished from T20 and C34, T1249 contains both PBD and LBD domain. It is an artificial peptide composed of sequences derived from HIV-1, HIV-2, and simian immunodeficiency virus (SIV). By comparison with T20, T-1249 has greater potency *in vitro* and exhibits activity against most T20-resistant isolates, whereas T-1249 escape variants are not sensitive to T20[25]. Both T20 and T-1249 showed membrane tropism. A fusion inhibitor concentration-dependent decrease on the membrane dipole potential was observed by applying di-8-ANEPPS fluorescence, a lipophilic probe sensitive to the changes in membrane dipole potential. T-1249 showed eight-fold rather higher affinity towards cell membranes, especially the cholesterol-rich domain[26]. Both experimental and simulation methods demonstrated that T-1249 could interact with 1-palmitoyl-2-oleyl-phosphatidylcholine (POPC) and POPC/cholesterol bilayers whereas T20 was shown much weaker and ineffective affinity to POPC/Cholesterol membranes[27 28 26]. LBD and PBD of T-1249 may synergistically reinforce the lipid binding affinity of T1249. These findings suggest the enhanced affinity towards lipid facilitates the drug delivery by blood cells. The membrane may further act as a "catalyst" to the binding between gp41 and peptide fusion inhibitors. The enhanced anti- HIV potency of C34-chol and TRI-999, substantiate this view. They are cholesterol and C18 lipid acids chain conjugated peptide separately derived from two peptide fusion inhibitors C34 and TRI-899[29 7]. Both C34 and TRI-899 lack LBD whereas peptide containing the residues that interact with cell membrane, like T20, did not show enhanced potency with cholesterol conjugation in contrast to C34[7]. Although T-1249 was discontinued to clinical development[30], further studies for peptides containing the PBD and LBD in addition to C-HR sequences provided penetrating insights into structure based rational design of the more effective inhibitors.

N-peptide inhibitors

N-peptides function mainly through two putative mechanisms. One is to interact with C-HR thus to interfere 6-HB formation; the other is to impact within N-helices of gp41, therefore, interfering the formation of the coiled-coil structure [31 32]. However, the potency of linear N-peptides is diminished by their tendency to aggregate in solution[31] and low capacity to form a trimeric coiled-coil. Because of these peptides, like N36, comprise the residues that form the hydrophobic pocket of the N-helix trimer. Nevertheless, appropriate modification of N-peptides could greatly enhance their antiviral potency.

N-peptides can be potent within nanomolar range when presented in the non-aggregating trimeric coiled-coil [31]. IQN17, a designed trimeric coiled-coil GCN4-pI_QI' fused to N17, exhibits two orders of magnitude more potent than N17 per se with an IC_{50} value of 190nM. The antiviral potency of another GCN4-pI_QI' fused inhibitor IQN23 is remarkably enhanced with IC_{50} value of 15nM [31]. IZN17, N17 fused to a different artificial trimeric coiled-coil named IZ, shows the similar low IC_{50} value with 22nM. The inhibitory activity of these peptides is correlated to their thermodynamic stability. CCIZN17, a IZN17 analogue that introduces disulfide bonds into the coiled-coil trimer, forms a highly stable and fully helical structure, $(CCIZN17)_3$, that can efficiently inhibit various HIV isolates with sub-nanomolar activity[32]. Peptides derived from N-HR fused to other trimeric domains could also contain high anti-HIV potency. N36 fused to the N terminus of a natural trimerization domain of T4 bacteriophage fibritin inhibits HIV at nanomolar range and could be put into large-scale production by expressing in *Escherichia coli (E.coli)* cells[33].

Fatty acid not only promotes the antiviral potency of C-peptides but also N-peptides. N36 is a poor HIV fusion inhibitor. Previous studies have demonstrated an IC_{50} value of 16000 ± 2000nM and 584 ± 46nM for acetylated and non-acetylated forms of N36 separately[34 35]. Interestingly, the enhancement of activity shows only a slight difference between the C- and N- terminus fatty acid conjugation of N36, whereas N27, which lacks the pocket sequence, only exhibited potent activity when fatty acids were conjugated to its N-terminus[35 36].

Fig.1 Schematic representation of the HIV-1 HXB-2 gp41 function domains and the corresponding N- and C-peptide sequences.

DERIVATIVES OF PEPTIDE FUSION INHIBITORS: bioconjugate techniques utilised in peptide modifications

Compared with parental peptides, conjugated peptides with diverse forms exhibit various advantages, mainly include: (1) protection of peptide from deactivation and degradation; (2) improvement in pharmacokinetics; (3) amelioration of peptide stability and helicity; (4) possibilities to form a drug-delivery system; (5) enhancement of antiviral potency.

For example, knowledge of the relationship between cell membrane and C-terminus of C-peptides allow the successful design for lipid or cholesterol conjugation with vertical enhanced potency and prolonged half-life. Synergistic effects have been generated by small-molecule fusion inhibitors conjugated peptide, and this conjugated peptide exhibits better stability against protein K digestion [37]. Moreover, most of the organic molecule conjugates exhibit better stability, which is very attractive for further drug design.

Membrane associated conjugates

High affinity towards membrane is a high-performance determinant for peptide fusion inhibitors; hence the enrichment of peptides at the membrane enhances the local concentration to boost efficiency and gives an effective orientation for fusion interruption. Introduction of lipodome components to C-peptides without LBD highly enhances their antiviral potency because these parental peptides have little affinity to interact with membranes. And enhanced potency was also observed in N-peptide conjugates, suggesting that lipid may function similar to the two lipid interactive domains in gp41.

Cholesterol conjugates

Cholesterol group is considered to be the most appropriate lipid anchor for the knowledge of the importance of cholesterol and lipid rafts in HIV entry. Compared with the cell membrane, HIV membrane contains more cholesterol and sphingomyelin that are often segregated in lipid rafts [38, 39]. Many transmembrane proteins and receptors are aggregated in lipid rafts, like CD4, the primary receptor when HIV attaches to the host cell. Therefore, anchoring in lipid rafts for peptide inhibitors may enhance the probability for the interaction between peptide and gp41 core. In addition, the study of T1249 suggests its increased antiviral potency with respect to T-20 is related to the high affinity towards cholesterol-rich membrane [40]. This capacity allows T1249 carried by peripheral blood mononuclear cell (PBMC) and erythrocyte membranes, which contain high level cholesterol [26]. These findings are in consistent with the success of cholesterol conjugated peptide.

Cholesterol conjugation enhances the antiviral potency by the improvement of the interaction with membranes and specific enrichment in the lipid rafts where HIV-cell fusion occurs [7]. Peptides with LBD that already drives insertion into lipid membranes are not considerable archetypes for this kind of modification because cholesterol maybe redundant, and the increase for potency may be limited. In fact, the decreased antiviral potency was observed in T-20. Hence, C34 based conjugates were constructed to investigate whether the anti-viral of peptides without LBD could be enhanced by addition of a cholesterol group. C34 is a well skeleton because it is a well-studied peptide which contains higher potency with PBD but not LBD, namely this conjugation may ascribe C34 more potent with high-performance of cell membrane enrichment and 6-HB interruption.

Consistent with its anticipated mechanism, C34-Chol with potency in a picomolar range, a C34 C-terminal added conjugate, is 25 to 100 fold more potent than parental C34, 50- to 400- fold more potent than T-20, and 15- to 300-fold more potent than T1249 according to different strain tested [7]. C34-chol is constructed by introducing a cysteine via a flexible GSG linker adding at the end of the C-terminus of C34. Notably, conjugation site is crucial to potency maintenance and enhancement because N-terminal conjugation of C34 shows a 50-fold decrease in potency compared with its unconjugated form. Interpretation of this result is that the wrong orientation towards the hairpin intermediate is made after binding to cell membrane. However, when cholesterol was added to the C-terminal of T-20, synergistic effect between LBD and cholesterol was not observed to reinforce the antiviral potency but showed to be detrimental, suggesting interference may occur between the

two lipophilic moieties[7]. This finding substantiates the opinion that cholesterol conjugation not fits peptides that contain LBD. C34-Chol exhibits an unexpected persistence during the cell assay. When washing steps added before triggering fusion, the antiviral potency only showed 7 fold decreases whereas parental C34 could not retain its potency when pre-incubated followed by wash[7], since C34-Chol perhaps anchored into the lipid raft during incubation and washing step could not remove all the peptides. Cell membranes could be the drug reservoir for C34-Chol [8]. Moreover, C34-Chol showed an enhanced half-life. It was still detectable in plasma after 24-hour injection in contrast with unconjugated form which under detectable during 6 hours in mouse model [7]. Further studies substantiate the membranotropic behaviour and interaction with circulating blood cells membranes [8]. It was demonstrated that C34-Chol could decrease the membrane dipole potential of erythrocytes and PBMCs[8] while HIV-1 was known to associate to the surface of those two kinds of cells in circulation[41, 42], suggesting a facilitate interaction and drug delivery.

Further cholesterol conjugates were established by adding a PEG linker to the C-terminus of C34 in order to form a cholesterol tagged monomer and dimer named HIVP3 and HIVP4 [43]. Multimerization and PEG addition could increase the potency of some fusion and entry inhibitors, and optimization of the length of the linker may provide ideal enhancement in potency. HIVP3 and HIVP4 exhibit a similar ability to interact with human erythrocytes and PBMCs compared with C34-chol[44]. However, the monomeric and dimeric cholesterol-tagged inhibitors exhibit comparable potency with a difference among variant strains. The increase stability of these peptides does not translate into antiviral potency. Early study suggests the extra-affinity may counteract the onset of resistance caused by affinity-disrupting mutations [45].

Sphingopeptides

Dihydrosphingomyelin is an unusual sphingolipid formed by sphinganine. It was described as a major component of biological membranes only in human lens extracts[46]. Further study identified that HIV lipidome shows a remarkable enrichment with dihydrosphingomyelin[39]. This observation enabled the research to investigate the capacity of sphinganine on incorporation into the site of fusion reaction. Sphinganine was conjugated to the highly conserved short fragments within 20 residues from Env core, which originally possess no antiviral potency [47]. Sphinganine endowed these inactive short C- and N- peptides, DP19 and N17, significant antiviral potency whereas other lipid moieties, including cholesterol and palmitic acid could not. This finding demonstrated that sphinganine was the only lipid component which could shift the inactive peptides within 20 residues to be active. Although lipid acids also can ascribe activity to both C- and N-peptides[48, 19], these peptides are much longer. These observations substantiate sphinganine could endow inactive peptide antiviral potency but whether it could reinforce the potency of active inhibitors, such as C34, remains to be elucidated.

Fatty acid conjugates

C-peptide conjugates

Fatty acid conjugation for C-peptides was initially constructed to identify the function of T-20 C-terminal region and demonstrate it functions mainly as an anchor to the cell membranes[19]. The abolished antiviral potency of the C-terminal truncated T-20 variants could be reserved by further C-terminal fatty acid conjugation and their inhibitory capacity is correlated with the length of the fatty acid chain. DP-C16, the minimal sequence that enables to create the gp41 hetero-functional-core conjugated with palmitic acid (C16) was the most active compared with the octanoic acid (C8) and dodecanoic acid (C12) conjugates[19]. The inhibitory potency of DP-C16 is almost identical to that of T-20, suggesting fatty acid could substitute the C-terminal residues which contain the membrane binding capacity.

Another very potent C-peptide conjugate is TRI-999, which derived from TRI-899, a 36 amino acid peptide from HIV-1_{LIA} with 24 amino acid residues overlaps with T-20[29]. The conjugation strategy is to add octadecanoic acid (C18) onto the $lysine_{30}$ residue through a PEG linker. TRI-899 possesses very potent activity and high genetic barrier to resistance[29]. Compared with TRI-899, TRI999 achieves desirable pharmacokinetic properties and elevated potency as well as resistance profile. Unfortunately, TRI-999 did not meet all the project criteria to enter clinical development.

N-peptide conjugates

As mentioned before, fatty acid could dramatically enhance the antiviral potency of N-peptides. Interestingly, N-terminal conjugates of N-peptides exhibit elevated activity rather than C-terminal conjugates. This observation is in consistent with the potency maintenance of C-terminal fatty acid substitution for C-peptide LBD. Native N-terminal extension residues of these N-peptides in gp41 form the FP. It is the N-terminal hydrophobic region of gp41, which directs fusion via anchoring the protein into the cell membrane. Fatty acids that attached in the N-terminus of the N-peptide simulate the FP to enhance the fusogenic activity toward phosphatidyl choline membrane. This finding demonstrates that fatty acid may act as the original membrane binding residues both in C- and N- peptides. However, different from C-peptide, the conjugation site of N-peptide is not as criticle as C-peptide: both C- and N-terminal conjugates exhibit enhanced inhibitory potency, suggesting a planar orientation of N-peptides as well as endogenous N-HR region on the cell membrane[35]. C-terminal conjugation of N36 was also observed with enhanced inhibitory potency but lower than N-terminal conjugation.

Conjugation of fatty acids to N36 and N54 significantly enhanced their fusogenic activity in the lipid mixing assay, although N54 was already fusogenic before fatty acid attachment. Decanoic acid significantly increased its fusogenic activity, especially at the low peptide: lipid molar ratios. However, fatty acid did not show a strong ability to enhance the affinity of N-peptides towards zwitterionic membrane [48]. A correlation between the enhanced antiviral potency and the length of fatty was observed [35]. Similar to C-peptide conjugates, longer fatty acid chain also means lower IC_{50} value for N-peptides. The inhibitory potency of N36 is increased up to 100 fold in a lipid conjugated dependent manner [35]. Due to the antiviral potency did not vary much between C- and N- terminal N36 conjugates; it had been demonstrated the orientation of the N-peptides caused by different terminal fixation into membrane was not critical to their anti-viral potency. However, significant distinction was observed in N27 conjugates [36]. N27 is a 27-mer peptide derived from N36 but without the pocket residues. It showed potent inhibitory capacity only when fatty acids were attached to its N-terminus not C-terminus. Among C8, C12 and C16 conjugate of N27, C16-N27 is the most potent inhibitor that targets C-HR rather than internal N-HR coiled-coil.

Albumin conjugates

Serum albumin is the most abundant protein and possesses an extraordinary long half-life in a range of 2-4 weeks, which makes it distinct from other plasma proteins [49]. The long-duration feature of albumin is due to the recognition of neonatal Fc receptor (FcRn)[50]. When taken up by cells, albumin will bind to the FcRn in a pH-dependent manner in the acidic environment of the early endosome. This binding prevents albumin from degradation in lysosome and allows it to return to plasma membrane, on which it is released back into the blood stream.

During the past decade, numerous proteins have been conjugated or genetically fused to albumin in order to delay the degradation of the protein drugs. And a lot of protein drugs associated with human serum albumin (HSA) have been approved for clinical use or developed into clinical trials. A phase Ⅲ trials for an interferon α-2b fusion protein for hepatitis C virus treatment demonstrated an every 2 weeks drug administration[51]. However, it has been suggested that fusion with HSA is not feasible for C-peptide inhibitors because Gp41 N-HR trimer is poorly accessible to protein cargoes of increasing size [52].

Fortunately, covalent attachment by chemical coupling did not show significantly decreased antiviral potency in the C34 conjugate named PC-1505[6]. C34 moiety of this novel molecule was functionalised with 3-maleimidopropionic acid (MPA) at its N-terminus, which allows the formation of a stable thioester bond between peptide and albumin. A 1:1 complex was achieved by attachment of the peptide to the accessible cysteine 34 of albumin. Cysteine 34 is a conservative residue among mammals and possesses the only free thiol group in albumin. These features attribute cysteine 34 of special interest for conjugation. PC-1505 exhibits equipotent ability to original C34 *in vitro* and retains antiviral potency even for a single preexposure dose for T-20 resistant virus caused by infrequent dosing *in vivo* in the SCID- hu Thy/Liv mouse model[6]. One single preexposure dose of PC-1505 could reduce viral RNA in HIV-1infected SCID-hu Thy/Liv mice with 3.3 $\log_{10}$ and prevent T-cells from depletion by the virus. These data showed a higher potency than Truvada, a prevention strategy proved to be effective for men who have sex with men in preexposure prophylaxis [53].

Based on the albumin-conjugated strategy utilised in PC-1505, another C34-albumin conjugate, FB006M, is designed superior to PC-1505. Three residues of the original C34 which are not involved in target binding were substituted to improve the solubility and antiviral activity. The 13th residue serine substituted with lysine was determined as the conjugation site to be modified with MPA and the other two residues were substituted by glutamic acid. This modified peptide could rapidly conjugate with albumin after intravenous injection *in vivo*. In rhesus monkey, FB006M showed a terminal half-life of 102.4h, which sustained 9.4-fold longer than general peptide [54]. FB006M, also known as Albuvirtide (ABT), exhibits broad spectrum potency in anti-HIV-1 pseudo viruses. It possessed potent inhibitory potency against A, B, B' and C subtypes which are prevalence in the world [55]. In addition, compared with C34 and PC-1505, ABT possessed significant enhanced potency in the interruption of 6-HB formation and lower IC_{50} to inhibit HIV virus. Impressively, variants with single or double mutations in gp41 were sensitive to ABT while they are highly resistant to T-20 and cross-resistant to C34. This spontaneous combination strategy utilised in albumin conjugation provides a novel insight into albumin associated drug design.

Other conjugates

Besides the approaches that associated the peptide inhibitors to cell membranes or serum albumin, other original strategies have induced into peptide modifications. For example, hybrid molecules of small-molecule fusion inhibitors covalently attached to C-peptide[37]; chimeric peptide that composed of two different target peptides[56]; polymers[57] [32]and bivalent peptides[58] were designed for novel drugs. The designs of these conjugates share one common idea that is to make several drug moieties synergistically work in one molecule and develop new drugs with higher potency than utilising several drugs together.

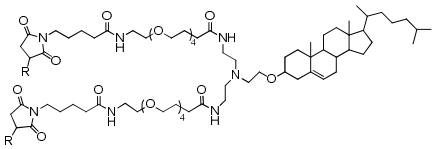
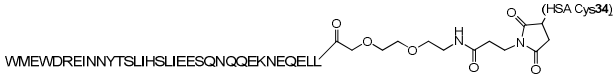
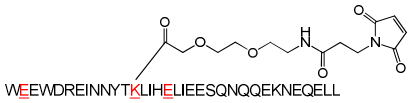
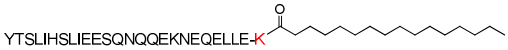
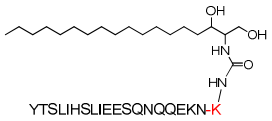
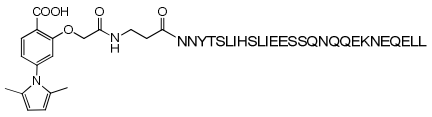
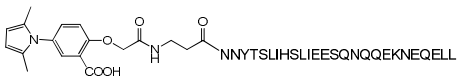
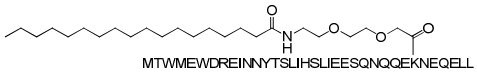
Highly petent small molecule-peptide conjugates

The primary cavity of gp41 coiled-coil motif is highly conserved and plays a key role in membrane fusion so that it becomes a very attractive target for small-molecule drug design. Moreover, it is also the target for C34 pocket binding domain. These observations provided a new thinking in considering whether small-molecule drugs aiming at this target could substitute the PBD of C34 and synergistically act on HIV with the rest of C34 residues. Two N-(carboxyphenyl) pyrrole derivatives, NB-2 and A_{12} were modified and covalently coupled with a C34 truncated peptide[37] named P26. In this kind of neotype drug molecules, P26 moiety was inactive due to PBD truncation and the small molecules targeted the primary cavity but showed a relatively weak potency in a micromolar range. Strikingly, even the antiviral potency is limit by application of the peptide and non-peptide parts separately, the conjugates Aoc-βAla-P26 and Noc-βAla-P26 showed a low nanomolar IC_{50} in the cell-cell fusion assay. Furthermore, the hybrid molecules exhibited better stability against proteinase K than T-20 and C34[37].

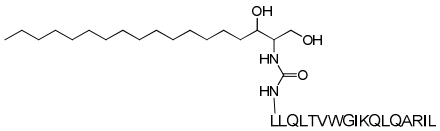
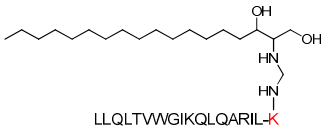
Chimeric bispesific peptide inhibitor

Two highly potent chimeric HIV entry inhibitors are designed with the bispecific feature. This kind of inhibitors is composed of two parts; one is a CCR5-targeting RANTES (regulated on activation normal T cell expressed and secreted) variant named 5P12 and 5P14; the other is a gp41 targeting C-peptide, C37[56]. The C-terminus of the RANTES peptides was fused to C37 by a 10-mer flexible glycine/serine linker "GGGGSGGGGS" and the fusion proteins were expressed in *E.coli*. The recombinant inhibitors 5P12-linker-C37 and 5P14-linker-C37 exhibited extremely high inhibitory potency with IC_{50} at low picomolar range in single cycle and replication-competent viral assays against R5-tropic viruses and this inhibition showed strain-dependent and was up to 100-fold better than the RANTES variant or with unlinked C37[56]. Moreover, the antiviral potency of C37 moiety induced in the chimeric inhibitor in targeting X4-tropic viruses was fully retained. 6000-fold potency enhancement against X4-tropic virus was observed, although C37 was the only part that target in this kind of virus. In addition, when target cells of the X4-tropic virus co-express CCR5 and CXCR4 receptor, the chimeric inhibitor can inhibit viral entry more efficiently[56].

Table 1. Peptides developed in bio-conjugates

Peptide Designation	sequence	Conjugate Designation	Conjugate Form
C-peptide			
C34	MEWDREINNYTSLIHSLIEESQNQQEKNEQELL	C34-Chol	
		HIVP3	
		HIVP4	
		PC-1505	
		FB006M (ABT)	
C37	HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELL	5P12-linker-C37	5P12-RANTES-$(GGGGS)_2$-HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELL
DP	YTSLIHSLIEESQNQQEKNEQELLE	DP-C16	
DP19	YTSLIHSLIEESQNQQEKN	DP19-Sphing	
P26	NNYTSLIHSLIEESQNQQEKNEQELL	Noc-βAla-P26	
		Aoc-βAla-P26	
Tri-899	MTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELL	Tri-999	

N-peptide

N36	SGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARIL	C16-N36	
		N36M-C16	
N54	STMGARSMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARIL	C10-N54	
N27	LLSGIVQQQNNLLRAIEAQQHLLQ	C16-N27	
N17	LLQLTVWGIKQLQARIL	Sphing-N17	
		N17K-Sphing	

Peptide fusion inhibitor sequences applied in study of peptide bioconjugation and their conjugated forms. Letters in red represent amino acid substitution or extra addition for conjugation reaction.

Table 2. Partial values for Anti-HIV activity of the peptide conjugates

Conjugate Designation	Archetype Designation	IC50 or EC50		HIV isolate	Target Cell	Reference
		Archetype	conjugate			
C34-Chol	C34	0.205±0.059	0.004±0.001	HXB2	P4–2/R5	[7]
Chol-C34			9.515±3.172			
HIVP3	C34	8.25±6.15	0.08±0.07	III_B	TZM-bl	[43, 44]
HIVP4			0.01±0.01	BaL		
PC-1505	C34	0.6	1.8	III_B	PBMC	[6, 53]
FB006M	FB006	1.4	3.9	III_B	PBMC	[54, 55]
5P12-linker-C37	C37	6.1±0.3	0.001±0.0003	HXB2	TZM-bl	[56]
		7.0±2.5	0.44±0.02	III_B	PBMC	
DP-C8	DP	2864±1492	261±64	HXB2	TZM-bl	[19]
DP-C12			9.9±0.39			
DP-C16			1.3±0.16			
DP19K-Sphinganie	DP19K	>4000	350±60	HXB2	TZM-bl	[47]
Noc−βAla−P26	P26	2540±630	14.5 ± 0.5	III_B	MT-2	[37]
Aoc−βAla−P26			14.1 ± 0.2			
C8-N36	N36	488±119	222±56	----	Jurkat E6-1 and Jurkat HXBc2	[35]
C12-N36			190±21			
C16-N36			72±27			
N36M-C16			159±47			
C8-N27	N27	338 ± 16	293±27	----	Jurkat E6–1 and Jurkat HXBc2	[36]
C12-N27			182±15			
C16-N27			10±1			
N27-C16			>1000			
Sphinganine-N17	N17	>4000	121±36	HXB2	TZM-bl	[47]
N17K-Sphinganine			287±143			
C34 trimer	C34	152	1.28	NL4-3	TZM-bl	[57]
T20 trimer	T20	465	54.1			

Summary and perspectives

Although IC_{50} values of T-20 are in low nanomolar range in vitro, the effective treatment of HIV-1 infection need high level serum concentrations that have been estimated to be nearly 1 μM. This observation may result from non-target binding and rapid degradation by the kidney. An ideal HIV-1 fusion inhibitor should be inexpensive, stable under plasma degradation, and possesses a long serum half-life and orally active[4]. However, not only general HIV peptide inhibitors but other protein drugs could not easily meet any of these requirements. But protein therapeutic is essential to diseases' control. They are highly specific to their target and work in a rather complex and subtle manner, which lower the side effects. Bio-conjugates for HIV treatment mentioned above give hints for further protein drug design to improve their efficacy and prolong drug half-life. In the future, we can anticipate that wild-type protein drugs perhaps will be substituted by specific modified proteins; that an even faster development of different bio-conjugation techniques will be generated for better protein therapeutics.

Acknowledgements

The authors thank Chinese Natural Science Foundation project (21173014).

1. Q. J. Sattentau and R. A. Weiss, *Cell*, 1988, **52**, 631-633.
2. B. J. Doranz, M. J. Orsini, J. D. Turner, T. L. Hoffman, J. F. Berson, J. A. Hoxie, S. C. Peiper, L. F. Brass and R. W. Doms, *Journal of virology*, 1999, **73**, 2752-2761.
3. P. D. Bieniasz, R. A. Fridell, I. Aramori, S. S. Ferguson, M. G. Caron and B. R. Cullen, *The EMBO journal*, 1997, **16**, 2599-2609.
4. F. Naider and J. Anglister, *Current opinion in structural biology*, 2009, **19**, 473-482.
5. J. J. Dwyer, K. L. Wilson, D. K. Davison, S. A. Freel, J. E. Seedorff, S. A. Wring, N. A. Tvermoes, T. J. Matthews, M. L. Greenberg and M. K. Delmedico, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 12772-12777.

6. C. A. Stoddart, G. Nault, S. A. Galkina, K. Thibaudeau, P. Bakis, N. Bousquet-Gagnon, M. Robitaille, M. Bellomo, V. Paradis, P. Liscourt, A. Lobach, M. E. Rivard, R. G. Ptak, M. K. Mankowski, D. Bridon and O. Quraishi, *The Journal of biological chemistry*, 2008, **283**, 34045-34052.
7. P. Ingallinella, E. Bianchi, N. A. Ladwa, Y. J. Wang, R. Hrin, M. Veneziano, F. Bonelli, T. J. Ketas, J. P. Moore, M. D. Miller and A. Pessi, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 5801-5806.
8. A. Hollmann, P. M. Matos, M. T. Augusto, M. A. Castanho and N. C. Santos, *PloS one*, 2013, **8**, e60302.
9. K. Tan, J. Liu, J. Wang, S. Shen and M. Lu, *Proceedings of the National Academy of Sciences of the United States of America*, 1997, **94**, 12303-12308.
10. W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel and D. C. Wiley, *Nature*, 1997, **387**, 426-430.
11. A. Merk and S. Subramaniam, *Current opinion in structural biology*, 2013, **23**, 268-276.
12. E. E. Tran, M. J. Borgnia, O. Kuybeda, D. M. Schauder, A. Bartesaghi, G. A. Frank, G. Sapiro, J. L. Milne and S. Subramaniam, *PLoS pathogens*, 2012, **8**, e1002797.
13. D. C. Chan, D. Fass, J. M. Berger and P. S. Kim, *Cell*, 1997, **89**, 263-273.
14. I. Munoz-Barroso, S. Durell, K. Sakaguchi, E. Appella and R. Blumenthal, *The Journal of cell biology*, 1998, **140**, 315-323.
15. D. M. Eckert and P. S. Kim, *Annual review of biochemistry*, 2001, **70**, 777-810.
16. Y. He, S. Liu, W. Jing, H. Lu, D. Cai, D. J. Chin, A. K. Debnath, F. Kirchhoff and S. Jiang, *The Journal of biological chemistry*, 2007, **282**, 25631-25639.
17. K. Champagne, A. Shishido and M. J. Root, *The Journal of biological chemistry*, 2009, **284**, 3619-3627.
18. Y. Kliger, S. A. Gallo, S. G. Peisajovich, I. Munoz-Barroso, S. Avkin, R. Blumenthal and Y. Shai, *The Journal of biological chemistry*, 2001, **276**, 1391-1397.
19. Y. Wexler-Cohen and Y. Shai, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 2007, **21**, 3677-3684.
20. S. G. Peisajovich, S. A. Gallo, R. Blumenthal and Y. Shai, *The Journal of biological chemistry*, 2003, **278**, 21012-21017.
21. S. Liu, H. Lu, J. Niu, Y. Xu, S. Wu and S. Jiang, *The Journal of biological chemistry*, 2005, **280**, 11259-11273.
22. H. Mo, A. K. Konstantinidis, K. D. Stewart, T. Dekhtyar, T. Ng, K. Swift, E. D. Matayoshi, W. Kati, W. Kohlbrenner and A. Molla, *Virology*, 2004, **329**, 319-327.
23. S. Liu, W. Jing, B. Cheung, H. Lu, J. Sun, X. Yan, J. Niu, J. Farmar, S. Wu and S. Jiang, *The Journal of biological chemistry*, 2007, **282**, 9612-9620.
24. W. Shi, Z. Qi, C. Pan, N. Xue, A. K. Debnath, J. Qie, S. Jiang and K. Liu, *Biochemical and biophysical research communications*, 2008, **374**, 767-772.
25. J. P. Lalezari, N. C. Bellos, K. Sathasivam, G. J. Richmond, C. J. Cohen, R. A. Myers, Jr., D. H. Henry, C. Raskino, T. Melby, H. Murchison, Y. Zhang, R. Spence, M. L. Greenberg, R. A. Demasi, G. D. Miralles and T. S. Group, *The Journal of infectious diseases*, 2005, **191**, 1155-1163.
26. P. M. Matos, M. A. Castanho and N. C. Santos, *PloS one*, 2010, **5**, e9830.
27. A. M. Martins do Canto, A. J. Carvalho, J. P. Ramalho and L. M. Loura, *International journal of molecular sciences*, 2013, **14**, 14724-14743.
28. A. M. Martins do Canto, A. J. Palace Carvalho, J. P. Prates Ramalho and L. M. Loura, *Computational and mathematical methods in medicine*, 2012, **2012**, 151854.

29. H. Zhang, S. E. Schneider, B. L. Bray, P. E. Friedrich, N. A. Tvermoes, C. J. Mader, S. R. Whight, T. E. Niemi, P. Silinski, T. Picking, M. Warren and S. A. Wring, *Organic Process Research & Development*, 2008, **Vol. 12**.
30. L. Martin-Carbonero, *AIDS reviews*, 2004, **6**, 61.
31. D. M. Eckert and P. S. Kim, *Proceedings of the National Academy of Sciences of the United States of America*, 2001, **98**, 11187-11192.
32. E. Bianchi, M. Finotto, P. Ingallinella, R. Hrin, A. V. Carella, X. S. Hou, W. A. Schleif, M. D. Miller, R. Geleziunas and A. Pessi, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 12903-12908.
33. X. Chen, L. Lu, Z. Qi, H. Lu, J. Wang, X. Yu, Y. Chen and S. Jiang, *The Journal of biological chemistry*, 2010, **285**, 25506-25515.
34. C. A. Bewley, J. M. Louis, R. Ghirlando and G. M. Clore, *The Journal of biological chemistry*, 2002, **277**, 14238-14245.
35. Y. Wexler-Cohen and Y. Shai, *PLoS pathogens*, 2009, **5**, e1000509.
36. Y. Wexler-Cohen, A. Ashkenazi, M. Viard, R. Blumenthal and Y. Shai, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 2010, **24**, 4196-4202.
37. C. Wang, W. Shi, L. Cai, L. Lu, Q. Wang, T. Zhang, J. Li, Z. Zhang, K. Wang, L. Xu, X. Jiang, S. Jiang and K. Liu, *Journal of medicinal chemistry*, 2013, **56**, 2527-2539.
38. R. C. Aloia, F. C. Jensen, C. C. Curtain, P. W. Mobley and L. M. Gordon, *Proceedings of the National Academy of Sciences of the United States of America*, 1988, **85**, 900-904.
39. B. Brugger, B. Glass, P. Haberkant, I. Leibrecht, F. T. Wieland and H. G. Krausslich, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**, 2641-2646.
40. A. S. Veiga, N. C. Santos, L. M. Loura, A. Fedorov and M. A. Castanho, *Journal of the American Chemical Society*, 2004, **126**, 14758-14763.
41. Z. Beck, B. K. Brown, L. Wieczorek, K. K. Peachman, G. R. Matyas, V. R. Polonis, M. Rao and C. R. Alving, *PloS one*, 2009, **4**, e8297.
42. C. Hess, T. Klimkait, L. Schlapbach, V. Del Zenero, S. Sadallah, E. Horakova, G. Balestra, V. Werder, C. Schaefer, M. Battegay and J. A. Schifferli, *Lancet*, 2002, **359**, 2230-2234.
43. A. Pessi, A. Langella, E. Capito, S. Ghezzi, E. Vicenzi, G. Poli, T. Ketas, C. Mathieu, R. Cortese, B. Horvat, A. Moscona and M. Porotto, *PloS one*, 2012, **7**, e36833.
44. M. T. Augusto, A. Hollmann, M. A. Castanho, M. Porotto, A. Pessi and N. C. Santos, *The Journal of antimicrobial chemotherapy*, 2014.
45. B. D. Welch, J. N. Francis, J. S. Redman, S. Paul, M. T. Weinstock, J. D. Reeves, Y. S. Lie, F. G. Whitby, D. M. Eckert, C. P. Hill, M. J. Root and M. S. Kay, *Journal of virology*, 2010, **84**, 11235-11244.
46. W. C. Byrdwell, D. Borchman, R. A. Porter, K. G. Taylor and M. C. Yappert, *Investigative ophthalmology & visual science*, 1994, **35**, 4333-4343.
47. A. Ashkenazi, M. Viard, L. Unger, R. Blumenthal and Y. Shai, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 2012, **26**, 4628-4636.
48. N. Lev and Y. Shai, *Journal of molecular biology*, 2007, **374**, 220-230.
49. R. E. Kontermann, *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy*, 2009, **23**, 93-109.
50. D. C. Roopenian and S. Akilesh, *Nature reviews. Immunology*, 2007, **7**, 715-725.
51. V. K. Rustgi, *Current medical research and opinion*, 2009, **25**, 991-1002.
52. A. E. Hamburger, S. Kim, B. D. Welch and M. S. Kay, *The Journal of biological chemistry*, 2005, **280**, 12567-12572.

53. C. A. Stoddart, G. Nault, S. A. Galkina, N. Bousquet-Gagnon, D. Bridon and O. Quraishi, *Antimicrobial agents and chemotherapy*, 2012, **56**, 2162-2165.

54. D. Xie, C. Yao, L. Wang, W. Min, J. Xu, J. Xiao, M. Huang, B. Chen, B. Liu, X. Li and H. Jiang, *Antimicrobial agents and chemotherapy*, 2010, **54**, 191-196.

55. H. Chong, X. Yao, C. Zhang, L. Cai, S. Cui, Y. Wang and Y. He, *PloS one*, 2012, **7**, e32599.

56. B. Zhao, M. K. Mankowski, B. A. Snyder, R. G. Ptak and P. J. Liwang, *The Journal of biological chemistry*, 2011, **286**, 28370-28381.

57. W. Nomura, C. Hashimoto, T. Suzuki, N. Ohashi, M. Fujino, T. Murakami, N. Yamamoto and H. Tamamura, *Bioorganic & medicinal chemistry*, 2013, **21**, 4452-4458.

58. Y. Ling, H. Xue, X. Jiang, L. Cai and K. Liu, *Bioorganic & medicinal chemistry letters*, 2013, **23**, 4770-4773.