Title: Molecular mechanisms of protein-cholesterol interactions in plasma membranes: Functional distinction between topological (tilted) and consensus (CARC/CRAC) domains

Author: Jacques Fantini Coralie Di Scala Carlos J. Baier Francisco J. Barrantes

PII: S0009-3084(16)30031-7
DOI: http://dx.doi.org/doi:10.1016/j.chemphyslip.2016.02.009
Reference: CPL 4456

To appear in: *Chemistry and Physics of Lipids*

Received date: 25-1-2016
Revised date: 24-2-2016
Accepted date: 26-2-2016

Please cite this article as: Fantini, Jacques, Di Scala, Coralie, Baier, Carlos J., Barrantes, Francisco J., Molecular mechanisms of protein-cholesterol interactions in plasma membranes: Functional distinction between topological (tilted) and consensus (CARC/CRAC) domains.Chemistry and Physics of Lipids http://dx.doi.org/10.1016/j.chemphyslip.2016.02.009

Molecular mechanisms of protein-cholesterol interactions in plasma membranes: Functional distinction between topological (tilted) and consensus (CARC/CRAC) domains

Jacques Fantini[1], Coralie Di Scala[1], Carlos J. Baier[2] and Francisco J. Barrantes[2]*

[1]EA-4674, Interactions Moléculaires et Systèmes Membranaires, Aix-Marseille Université, Marseille, France, [2]Laboratory of Molecular Neurobiology, Institute for Biomedical Research (BIOMED), Faculty of Medical Sciences, UCA-CONICET, Av. Alicia Moreau de Justo 1600, C1107AFF Buenos Aires, Argentina.

*To whom correspondence should be addressed: rtfjb1@yahoo.com

Abstract

The molecular mechanisms that control the multiple possible modes of protein association with membrane cholesterol are remarkably convergent. These mechanisms, which include hydrogen bonding, CH-π stacking and dispersion forces, are used by a wide variety of extracellular proteins (e.g. microbial or amyloid) and membrane receptors. Virus fusion peptides penetrate the membrane of host cells with a tilted orientation that is compatible with a transient interaction with cholesterol; this tilted orientation is also characteristic of the process of insertion of amyloid proteins that subsequently form oligomeric pores in the plasma membrane of brain cells. Membrane receptors that are associated with cholesterol generally display linear consensus binding motifs (CARC and CRAC) characterized by a triad of basic (Lys/Arg), aromatic (Tyr/phe) and aliphatic (Leu/Val) amino acid residues. In some cases, the presence of both CARC and CRAC within the same membrane-spanning domain allows the simultaneous binding of two cholesterol molecules, one in each membrane leaflet. In this review the molecular basis and the functional significance of the different modes of protein-cholesterol interactions in plasma membranes are discussed.

Subheadings:

1. Cholesterol: a singularity in the lipid kingdom. Cholesterol is unique among membrane lipids because it is polycyclic, has a very small polar head group, and does not contain any acyl chain allowing biochemical variability. It is present in each leaflet of the plasma membrane where it may interact with specific lipids such as phosphatidylserine in the inner leaflet [1]. Moreover, tail-to-tail interactions of cholesterol molecules in a symmetric topology have been evidenced in model membranes [2] and probably occur in vivo [3]. In the outer leaflet, cholesterol has a marked preference for sphingolipids, whose apolar section is generally more rigid than that of glycerophospholipids. Hence, there is a lateral distribution of membrane cholesterol which is enriched in sphingolipid domains [4 5]. In these membrane areas, condensed cholesterol/sphingolipids complexes [6] form a specific physical phase referred to as liquid ordered (Lo) [7,8] (**Figure 1**). In contrast, cholesterol is loosely associated with glycerophospholipids (chiefly phosphatidylcholine in the outer leaflet) that constitute a distinct phase referred to as liquid disordered (Ld) [4,9]. The attraction to sphingolipids and at the same time lack of firm adhesion to phosphatidylcholine results in a lipid-based segregation of cholesterol in the Lo phase and a sparse distribution of cholesterol in the Ld phase [4,7] (**Figure 1**). Although this "lipid-exclusive" description does not take into account the important role of membrane proteins in the formation of membrane domains [10,11], it provides a schematic overview of the possible molecular interactions that cholesterol molecules may share in the plasma membrane, especially in the exofacial (outer) leaflet.

From a purely geometrical point of view, there are only two angles of approach of cholesterol to the plasmalemma: from the extracellular space (vertical approach) or from the apolar region of the membrane (lateral approach). The vertical approach implies that

cholesterol has to be accessible to extracellular proteins (or extracellular domains of host membrane proteins). In this case, the "readable" parts of cholesterol, i.e. the atoms of the lipid that are accessible for binding, are generally restricted to the hydroxyl group, which is available only in the Ld phase (**Figures 1 and 2**). Indeed, in the Lo phase, cholesterol is totally masked by the polar head groups of sphingolipids, including both sphingomyelins and glycosphingolipids, through a well-characterized "umbrella effect" [12,13] (**Figure 2A**). The molecular mechanism of this effect is based on the formation of a hydrogen bond network (**Figure 2A**) that involves the -OH group of cholesterol. Therefore, there is little chance of the -OH group of raft-associated cholesterol, bound to and masked by surrounding sphingolipids, being targeted by an extracellular protein. In contrast, the loose association of cholesterol with phosphatidylcholine makes the -OH group both accessible and usually free of bonding, so that it is fully available for extracellular ligands. An important category of proteins that bind cholesterol from the extracellular space are bacterial toxins such as cytolysins [14]. These cholesterol-dependent toxins have a conserved pair of amino acid residues (threonine/leucine) able to detect the OH group of cholesterol and thus promote the initial association of the toxin with the plasma membrane [15]. This process is described in **Figure 2B**.

2. Topological motifs: why they are tilted. After binding to the -OH group, proteins must stabilize their association with the plasma membrane. A single interaction with a cholesterol molecule involving only the -OH group of the lipid is obviously far from being sufficient. Under these circumstances the protein has three possibilities: i) disrupt its association with cholesterol and return to the extracellular space; ii) stabilize its adhesion to the membrane

by interacting with other lipid head groups; or iii) penetrate the membrane. The latter outcome corresponds to an insertion process that occurs during the early steps of virus infection, for instance. Numerous viruses fuse their envelope with the plasma membrane of host cells [16]. This early event in the infection cycle is usually cholesterol-dependent [17,18], yet the molecular mechanisms that physically involve cholesterol during virus fusion have been a matter of speculation for years [14]. The viral protein that initiates the process displays a hydrophobic peptide containing a succession of apolar amino acid residues, referred to as "fusion peptide" [16]. A hallmark of fusion peptides is that they insert in the membrane with a tilted orientation, i.e. with an angle that can be as high as 45° [19]. Consequently, the tilt induces a significant disturbance of the organization of the lipid bilayer [20] . The relationship between fusion peptides and cholesterol was not immediately recognized because the tilted orientation is in fact an intrinsic property that does not require a consensus amino acid sequence [21]. However, because they are characterized by an asymmetric distribution of their hydrophobic residues, tilted peptides can be predicted by the application of algorithms that analyze the degree of polarity/apolarity in a continuous protein segment [22] . If it is the nature of a tilted peptide to adopt a tilted orientation when it inserts in the membrane, why would specific lipids be required to favor this kind of orientation? The finding that tilted peptides from various pathogenic proteins can bind cholesterol with high affinity [23] has shed some light on the cholesterol dependence of virus fusion mechanisms [3]. Some typical examples of cholesterol binding to tilted peptides are shown in **Figure 3**. One can note the remarkable geometric complementarity between the apolar domain of cholesterol and the aliphatic residues of the peptide. As discussed above, there is no particular consensus sequence characterizing a cholesterol-binding site on a tilted peptide. More simply, it can be envisioned that the tilted orientation allows an excellent adaptation of the peptide to the

inverted cone shape of cholesterol, a geometric feature that does not exist for lipids other than cholesterol in the exofacial leaflet of the plasma membrane [14]. Viral proteins are not the only pathogenic proteins that possess a tilted cholesterol-binding domain. The amyloid proteins involved in the pathogenesis of neurodegenerative disorders, including Alzheimer and Parkinson diseases, have the capacity to self-oligomerize into highly neurotoxic amyloid pores [24,25]. These proteins display a tilted peptide which interacts with cholesterol during the insertion/assembly of the oligomers in the outer leaflet of the plasma membrane of brain cells [26-28] . Using a combination of in silico and experimental approaches, Di Scala et al. have demonstrated that the tilted orientation of the membrane-associated domain of the amyloid protein favors the oligomerization process driven by the formation of hydrogen bonds linking two vicinal peptide subunits [28]. In this case, functional binding to cholesterol ensures that the amino acid residues involved in the formation of these hydrogen bonds (e.g. Asn-27 and Lys-28 for Alzheimer's β-amyloid peptide) [28] are at the right distance from each other. The oligomerization mechanism subsequently proceeds in a coordinated fashion until a channel-like amyloid pore is formed [24,29,30]. This tilted orientation of the Parkinson's disease-associated protein α-synuclein, when bound to membrane cholesterol, is illustrated in **Figure 3**. It is interesting to note how similar the molecular complexes are between cholesterol and viral (HIV-1, Ebola) and amyloid peptides such as Alzheimer's β-amyloid peptide or α-synuclein [14]. Overall, the insertion process of tilted peptides can be decomposed into two sequential events: i) a first vertical approach during which the "infectious" protein (either viral or amyloid) interacts chiefly with the surface-exposed -OH group of cholesterol [31] ; and ii) a penetration step allowing a lateral interaction between cholesterol and the tilted peptide [23]. The end of the process depends on the length of the protein. In the case of virus proteins, the fusion peptide initiates the insertion; this

subsequently proceeds until the virus envelope comes into close contact with the plasma membrane of the host cell [16,32]. In the case of amyloid pores, the process is stopped when the polar part of the protein, which cannot penetrate the membrane, reaches the polar/apolar interface of the membrane [23,25].

It is interesting to note that tilted peptides act as topological binding sites for cholesterol but do not control the fate of the complex they form with cholesterol. The association of a tilted peptide with cholesterol can be either transient (viruses) or permanent (amyloid pores), depending on the mean polarity of the protein outside the cholesterol-binding motif. These topological cholesterol-binding motifs therefore do not need to be of exceptional affinity: they must be strong enough to catch membrane cholesterol, yet prepared to disengage in accordance with the "pressure" exerted on the complex by the whole protein. Such seemingly opposing (and subtle) requirements have not led to the emergence of a clear-cut consensus motif, but to the myriad of tilted peptides found in a broad range of viruses and amyloid proteins, now collectively referred as "infectious proteins" [14,33]. These domains consist in an endless variety of amino acid combinations that obey a "tilted" code based on the gradient of apolar residues along their axis [19]. Such structural characteristics are also consistent with a functional interaction with cholesterol. Indeed, all tilted peptides identified so far contain both small (Gly, Ala) and branched aliphatic amino acid residues (Val, Leu Ile) (**Figure 3** and **Table 1**), a combination enabling an optimal fit for cholesterol. On the one hand, Gly residues might confer sufficient mobility on the TM domain to enable it to adapt its shape to the cholesterol molecules; and on the other hand, the aliphatic groups of Ala, Val, Leu and Ile could occupy the cavities of cholesterol. Finally, the presence of both small and bulky side chains is consistent with the formation of a groove into which cholesterol can optimally fit (**Figure 3)**. A notable feature of tilted peptides, which may differentiate them

from other cholesterol-binding motifs, is that the presence of aromatic residues is not mandatory. As shown in Table 1, some tilted peptides display 1, 2, or 3 aromatic residues, whereas others do not contain any. The lack of a key sequence signature for tilted peptides is a perfect example of convergence in molecular evolution, in that distinct evolutionary pathways have generated common membrane-associated properties. It is not clear whether evolution has selected tilted peptides primarily for allowing cholesterol binding, or whether this property is a chance consequence of the peculiar geometry of tilted peptides.

3. Consensus motifs: when simplicity hides sophistication. The vast majority of host membrane proteins, including receptors and ion channels, display one or several transmembrane (TM) domains that cross the lipid bilayer [3]. Schematically, a TM domain is a α-helical segment of 20-25 apolar amino acids flanked at each end by polar residues. This symmetric configuration has two roles: i) facilitating the transition between the apolar inside of the membrane and the polar milieu bathing the membrane surface, and ii) allowing the TM domain to interact simultaneously with both the apolar chains and the polar head groups of membrane lipids. Basically, membrane lipids can be divided in three categories: glycerophospholipids, sphingolipids, and cholesterol [4]. If one considers the well-ordered distribution of apolar and polar amino acid residues, it is clear that a TM may functionally interact with any of these three types of lipids, inside or outside ordered lipid ("raft") domains[8] . However, a distinction has to be made between annular lipids, that are exchangeable, and non-annular lipids, that remain tightly bound to the protein [34]. In the first case, the lipid-protein association is not selective, and of sufficiently low affinity to allow (relatively rapid) lipid exchange. In the second case, the TM domain displays a specific

binding motif that selects a given lipid, e.g. cholesterol[35]. Non-annular lipids interact specifically with membrane proteins, influence their conformation, and regulate their functions [34].

So far two linear consensus cholesterol-binding domains have been characterized [3]. The first one is the CRAC domain (an acronym standing for "Cholesterol Recognition/interaction Amino acid Consensus sequence")[36] . CRAC is a short motif which fulfils the simple algorithm (L/V)-X_{1-5}-(Y)- X_{1-5}-(K,R). Although the CRAC motif has been found in various proteins that bind cholesterol[37] , the simplicity of the consensus sequence defined by only three specific amino acids and up to ten undefined residues (thus referred to as "X" in the algorithm) has raised some skepticism about its predictive value [11,38]. However, the cholesterol-binding activity of CRAC has been carefully established by mutational studies. In particular, it has been initially demonstrated that the central tyrosine residue of CRAC cannot be replaced by any other aromatic residue [11,39]. Nevertheless, this rule might not apply to CRAC motifs located in TM domains, since in this specific case both Tyr and Phe-containing motifs are predicted to bind cholesterol with high affinity (**Table 1**, **Figure 4**) [3,40]. The 7th TM domain of the human cannabinoid receptor CB1 is a perfect example of a functional CRAC-cholesterol interaction involving tyrosine as the aromatic residue (**Figure 4**). The sequence of this CRAC motif is 392-**V**NPII**Y**AL**R**-400. In this case, all three residues defined by the algorithm (Val-204, Tyr-209 and Arg-214) are in close contact with the sterol. The whole energy of interaction of this CRAC/cholesterol complex was estimated to be -55 kJ.mol^{-1}. Interestingly, there is also a CRAC motif in the 3rd TM domain of CB1, yet in this case the aromatic residue is Phe instead of Tyr: 204-**V**GSL**F**LTAID**R**-214 (**Figure 4**). Docking studies suggested that cholesterol could also bind to this CRAC motif with high affinity (energy of interaction -60 kJ.mol^{-1}). However, the Phe residue of this CRAC motif is not directly involved in cholesterol

binding. Finally, the 1st TM domain of the human GABA type B receptor (subunit 2), has a CRAC motif with two Phe residues (550-**L**F**F**NI**K**-555) and both of these aromatic residues interact with cholesterol, providing a total energy of interaction of -31 kJ.mol^{-1} (**Figure 4**). Taken together, these data indicate that neither the nature of the aromatic amino acid (Tyr or Phe) nor its physical involvement in cholesterol binding can help to predict the affinity of a CRAC motif for membrane cholesterol.

The second consensus cholesterol-binding motif is a reversed version of the CRAC algorithm, i.e. (K/R)-X_{1-5}-(Y/F)- X_{1-5}-(L,V) that was logically coined "CARC" [3,41] . This new cholesterol-binding motif has been discovered in the human nicotinic acetylcholine receptor (AChR), whose TM domains do not display the CRAC motif. More precisely there are some CRAC motifs in this receptor, but not in TM domains. Similarly, we detected a CRAC domain in the intracellular loop joining the 2nd and 3rd TM domains of the human delta-type opioid receptor [3]. Despite the fact that all these CRAC domains may have the intrinsic capability to bind cholesterol, their location outside the membrane renders such interactions highly unlikely in vivo.

Typical examples of CARC motifs within TM domains of human neurotransmitter receptors are listed in **Table 1**. Physicochemical studies combining Langmuir monolayer studies and NMR approaches have shown that the CARC-cholesterol interaction is of high affinity, lipid specific and saturable [42]. Moreover, in silico studies have shed some light on the possible modes of interaction of the CARC motif with various TM domains [3,41]. In particular, both tilted and non-tilted orientations of the TM domain bound to membrane cholesterol have been evidenced [41] (**Figure 5**). A typical example of a tilted CARC motif is found within the 5th TM domain of the human prolactin-releasing peptide receptor: 221-**R**QL**Y**A**W**GL**LLV**-231

(**Figure 5**). An interesting feature of this CARC motif is that it contains two aromatic residues, Tyr-223 and Trp-225. By itself, the side chain of Trp-225 contributes 18% of the energy of interaction of the domain with cholesterol (-10 kJ.mol^{-1} for Trp-225 compared with -55kJ.mol^{-1} for the whole CARC). In contrast, the CARC motif of the 7th TM domain of human adenosine receptor (265-**K**PSILT**Y**IAIF**L**-276) does not show any tilt when bound to membrane cholesterol (**Figure 5**). Its parallel orientation with respect to the main axis of the cholesterol molecule is nevertheless consistent with high affinity binding (energy of interaction -55 kJ.mol^{-1}).

In most cases, the central aromatic residue of CARC appears to play a key role in the interaction, as assessed by the high energies of interaction predicted from molecular dynamics studies [3,41] . The involvement of an aromatic residue suggests that one of the driving forces of cholesterol binding to these CARC motifs is the CH-π bond[40]. This kind of interaction is particularly efficient when an aliphatic ring stacks onto an aromatic structure, as is the case for the human type 3 somatostatin receptor[3] . The coordinated network of CH-π bonds adopts a typical geometry that is perfectly illustrated in **Figure 6**. In essence, this mechanism is reminiscent of the sugar-aromatic interactions involved in lectin-carbohydrate [43] or protein-glycolipid [44,45] interactions. Because it can adopt distinct geometric orientations that may all be compatible with cholesterol binding, the CARC motif is intrinsically more flexible than most topological domains which, as explained above, have to be tilted. Tilted peptides have typical amino acid residues that are regularly arranged in such a way that they generate a polarity gradient in the protein chain [46] . In this case, the tilt is mandatory because there is no other way to optimize the insertion of a tilted peptide within a membrane bilayer [46] . In contrast, the CARC algorithm is compatible with several possible positions of the aromatic residue (**Figure 6**).

Several authors have underscored the fact that the CRAC consensus sequence is very general and hence questioned its predictive value with regards to cholesterol binding [38]. The same remark could also apply to the CARC motif [3]. Nevertheless, one can argue that what might be true for water-soluble proteins is no longer valid when it comes to membrane proteins. As discussed above, the presence of a CARC (or a CRAC) motif within a TM domain is, from a molecular point of view, consistent with a specific interaction with cholesterol. Whether or not this TM domain actually interacts with cholesterol may depend essentially on the availability of cholesterol in the vicinity of the protein. In water-soluble globular proteins displaying CARC/CRAC motifs, the apolar residues of these motifs might be buried in the apolar core of the protein and thus not be accessible to any ligand. More generally, the predictive value of CARC/CRAC motifs for identifying cholesterol binding sites can be considered as particularly high for TM domains, and lower everywhere else[3,41]. A reasonable recommendation is thus to consider not only the presence of a consensus CARC/CRAC motif but also the nature of the protein that contains such motifs: is the protein known to bind cholesterol, is it regulated by cholesterol, is it a membrane protein? In this respect, we have developed a method for identifying cholesterol-binding motifs from the amino acid sequence of membrane proteins[42].

4. Dual presence of CARC/CRAC motifs in the same TM domain: a mirror code?

An interesting feature of the CARC and CRAC algorithms is that these motifs are both vectorial and symmetric. In theory, it is possible that the same TM domain displays both a CARC and a CRAC motif. Let us consider a Type 1 membrane protein whose N-terminus is extracellular and its C-terminus intracellular. The TM domain of this protein could thus have

a CARC motif in the outer leaflet and a CRAC domain in the inner leaflet (**Figure 7**). A similar situation could also apply for the 1st, 3rd, 5th, and 7th TM domains of G-protein coupled receptor with seven TM domains [3]. To assess whether such a "mirror" topology actually exists, we browsed sequence databases and found several cases of receptors with both CARC and CRAC motifs in their TM domains. An example is given in **Figure 7** for the 1st TM domain of the human VIP receptor. Molecular dynamics simulations suggested that this TM domain could perfectly accommodate two cholesterol molecules in a typical tail-to-tail orientation, one bound to CARC and the other to CRAC. The CARC/CRAC mirror motif has been found in a broad range of receptor proteins, including GABA, glutamate, adenosine, TRVP1, and endocannabinoid receptors [42].

Besides its application in the field of membrane receptors, the mirror code could also potentially be applied to the ABC transporters involved in cholesterol transport across cellular membranes. Both CARC and CRAC motifs have been identified in these proteins[42,47-50]. In this case, one can tentatively hypothesize that the simultaneous presence of both CARC and CRAC within the same TM could favor the binding and subsequent translocation of cholesterol along the TM domain[42,50].

A thorough analysis of membrane protein sequences will be necessary to determine whether the dual presence of CARC and CRAC within the same TM domain could be due to a "mirror" code selected by Evolution to allow a functional interaction with a pair of cholesterol molecules, one in each leaflet of the plasma membrane.

5. Conclusions and perspectives.

From a molecular point of view, the mechanisms of protein-cholesterol interactions in plasma membranes are remarkably convergent. Looking at the molecular structure of cholesterol gives a good idea of the nature of these mechanisms: Hydrogen bonding for the polar hydroxyl group, CH-π stacking for the aliphatic rings, Van der Waals (dispersion forces) for the iso-octyl chain. On the protein side, this means that a proper triad of amino acids within a α-helical segment of ca. 15-20 Å may constitute a functional cholesterol-binding domain. The triad should include, in a mandatory order, a basic, an aromatic, and a branched aliphatic residue whose respective roles in cholesterol binding have been recently unraveled [3]. Vectorial motifs such as CARC and CRAC, which fulfill these conditions, generally behave as high affinity, non-annular binding sites for TM domains of numerous membrane proteins whose function is dependent on membrane cholesterol [8]. The amino acid residues that define the linear CARC and CRAC motifs are also present in a spatial cholesterol-binding domain that has been referred as to the 'cholesterol-consensus motif' (CCM)[51] . The CCM consists of a cluster of basic (K, R), aromatic (F, Y, W) and aliphatic (I, V, L) residues that are, in this case, provided by two separate TM domains (basically (W/Y)-(I/V/L)-(K/R) on one TM, and (F/Y/R) on the second one)[51]. From a topological point of view, CARC and CRAC are linear domains whereas the CCM is a more complex three-dimensional structure. Nevertheless, the first principles that govern the interaction of cholesterol with CARC or CRAC also apply for this three-dimensional binding site since all these domains have similar amino acid requirements[3]. Finally, a distinct type of cholesterol-protein interaction has been described for the Kir2.1 channel which does not display any CARC, CRAC or CCM motifs. In this case, cholesterol is predicted to interact with a pocket consisting of several branched aliphatic residues (I, L, and V) in a region that, like the classical consensus motifs, also

contains aromatic and polar residues[52,53]. In contrast, tilted peptides do not have a particular consensus amino acid signature but a unique topology that may also sustain cholesterol binding. A functional distinction between topological (tilted) and consensus (CARC/CRAC) domains can be made, because the former are chiefly found in infectious proteins (either microbial or amyloid) and the latter in resident membrane receptors. Further studies will allow us to clarify the respective roles of topological and consensus motifs in proteins that bind cholesterol either permanently or transiently. Ideally, these studies will combine in silico approaches, structural investigations of cholesterol-protein complexes, and physico-chemical studies of proteins with cholesterol-containing membranes.

Figure legends

Figure 1. Cholesterol and membrane phases. In the Lo phase (raft or sphingolipid domains), cholesterol is masked by sphingolipids such as sphingomyelin (SM) or glycosphingolipids (GSL). In the Ld phase, cholesterol molecules are surrounded by glycerophospholipids such as phosphatidylcholine (PC) and the polar -OH group of cholesterol is accessible to extracellular ligands. In both phases cholesterol molecules are indicated by arrows.

Figure 2. Cholesterol-sphingolipid interactions. A) The glycosydic bond of the glycosphingolipid (GSL), its first sugar, and the -OH group of cholesterol (chol) interact through a coordinated network of H-bonds. The apolar parts of both lipids are stabilized by dispersion (van der Waals) forces. **B)** The initial contact of soluble listeriolysin with the

membrane is mediated by the establishment of an H-bond between the -OH group of cholesterol and residues Thr-515/Leu-516 of the toxin (highlighted in orange). Chol, cholesterol; PC, phosphatidylcholine.

Figure 3. Interaction of tilted peptides with cholesterol. Fusion peptides are represented in surface rendering models (on the right) or in ribbon models (on the left). In both cases cholesterol is in yellow (atomic sphere rendering). The amino acid residues involved in cholesterol binding are indicated in red in the sequence. Docking of cholesterol was performed with the Hyperchem program as previously described [23,41].

Figure 4. Interaction of cholesterol with CRAC motifs. In each cholesterol/CRAC complex the TM domain is represented in ribbon (models on the left) or in electrostatic surface rendering with red ribbon inside (models on the right, with positive charges in blue, negative in red and neutral domains in white). Cholesterol is in yellow. The amino acid residues involved in cholesterol binding are indicated in red in the sequence. CB1, human cannabinoid receptor 1; $GABA_B$, human GABA type B receptor subunit 2.

Figure 5. Interaction of cholesterol with CARC motifs. The cholesterol/CARC complexes are represented in surface rendering as explained in the caption of Figure 3. Pro-RP Rec., human prolactin-releasing peptide receptor; A1, human adenosine A1 receptor; TRVP1, human transient receptor potential cation channel subfamily V member 1.

Figure 6. Role of CH-π stacking in cholesterol recognition in the CARC context. The model on the left shows the interaction of cholesterol with the CARC motif of human type 3 somatostatin receptor: 203-**R**AG**F**IIYTAA**L** – 213. The phenyl ring of F-206 stacks onto one of the rings of cholesterol as illustrated by the models in the middle of the figure. This

coordinated CH-π stacking interaction is typical of many cholesterol-TM complexes. The cartoon on the right side illustrates the fact that the aromatic residue of CARC can be located at various positions along the TM axis, consistent with the establishment of CH-π stacking interaction with any of the four cycles of cholesterol. In contrast, the location of first basic residue of CRAC (Lys/Arg) is always located at the surface of the membrane (snorkeling effect[54]).

Figure 7. A mirror CARC/CRAC code for TM domains? The respective topologies of CARC and CRAC are consistent with the presence of a dual CARC/CRAC motif within the same TM domain (cartoon on the left). A typical example of such a situation is given by the 1st TM domain of the human VIP receptor (model on the right with the cholesterol bound to CARC in yellow and the other one bound to CRAC in green).

References

1 Leventis, P. A. & Grinstein, S. The distribution and function of phosphatidylserine in cellular membranes. *Annu. Rev. Biophys.* **39**, 407-427 (2010).

2 Harris, J. S., Epps, D. E., Davio, S. R. & Kezdy, F. J. Evidence for transbilayer, tail-to-tail cholesterol dimers in dipalmitoylglycerophosphocholine liposomes. *Biochemistry* **34**, 3851-3857 (1995).

3 Fantini, J. & Barrantes, F. J. How cholesterol interacts with membrane proteins: an exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. *Front. Physiol.* **4**, 31 (2013).

4 Fantini, J., Garmy, N., Mahfoud, R. & Yahi, N. Lipid rafts: structure, function and role in HIV, Alzheimer's and prion diseases. *Exp. Rev. Mol. Med.* **4**, 1-22 (2002).

5 Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569-572.

6 Radhakrishnan, A., Anderson, T. G. & McConnell, H. M. Condensed complexes, rafts, and the chemical activity of cholesterol in membranes. *Proc. Natl. Acad. Sci. USA* **97**, 12422-12427 (2000).

7 Brown, D. A. & London, E. Functions of lipid rafts in biological membranes. *Annu. Rev. Cell Dev. Biol.* **14**, 111-136 (1998).

8 Fantini, J. & Barrantes, F. J. Sphingolipid/cholesterol regulation of neurotransmitter receptor conformation and function. *Biochim. Biophys. Acta* **1788**, 2345-2361 (2009).

9 Brown, R. E. Sphingolipid organization in biomembranes: what physical studies of model membranes reveal. *J. Cell Sci.* **111**, 1-9 (1998).

10 Pike, L. J., Han, X., Chung, K. N. & Gross, R. W. Lipid rafts are enriched in arachidonic acid and plasmenylethanolamine and their composition is independent of caveolin-1 expression: a quantitative electrospray ionization/mass spectrometric analysis. *Biochemistry* **41**, 2075-2088 (2002).

11 Epand, R. F. *et al.* Juxtamembrane protein segments that contribute to recruitment of cholesterol into domains. *Biochemistry* **45**, 6105-6114 (2006).

12 Ramstedt, B. & Slotte, J. P. Interaction of cholesterol with sphingomyelins and acyl-chain-matched phosphatidylcholines: a comparative study of the effect of the chain length. *Biophys. J.* **76**, 908-915 (1999).

13 Fantini, J., Yahi, N. & Garmy, N. Cholesterol accelerates the binding of Alzheimer's beta-amyloid peptide to ganglioside GM1 through a universal hydrogen-bond-dependent sterol tuning of glycolipid conformation. *Front. Physiol.* **4**, 120 (2013).

14 Fantini, J., & Yahi, N. *Brain lipids in synaptic function and neurological disease. Clues to innovative therapeutic strategies for brain disorders.*, Elsevier Academic Press, San Francisco (2015).

15 Farrand, A. J., LaChapelle, S., Hotze, E. M., Johnson, A. E. & Tweten, R. K. Only two amino acids are essential for cytolytic toxin recognition of cholesterol at the membrane surface. *Proc. Natl. Acad. Sci. USA* **107**, 4341-4346 (2010).

16 Harrison, S. C. Viral membrane fusion. *Virology* **479-480**, 498-507 (2015).

17 Wudiri, G. A. *et al.* Molecular requirement for sterols in herpes simplex virus entry and infectivity. *J. Virol.* **88,** 13918-13922 (2014).

18 Schroeder, C. Cholesterol-binding viral proteins in virus entry and morphogenesis. *Subcell. Biochem.* **51**, 77-108 (2010).

19 Brasseur, R., Pillot, T., Lins, L., Vandekerckhove, J. & Rosseneu, M. Peptides in membranes: tipping the balance of membrane stability. *Trends Biochem. Sci.* **22**, 167-171 (1997).

20 Brasseur, R. Tilted peptides: a motif for membrane destabilization (hypothesis). *Mol. Membr. Biol.* **17**, 31-40 (2000).

21 Crowet, J. M. *et al.* Tilted properties of the 67-78 fragment of alpha-synuclein are responsible for membrane destabilization and neurotoxicity. *Proteins* **68**, 936-947 (2007).

22 Lins, L. & Brasseur, R. Tilted peptides: a structural motif involved in protein membrane insertion? *J. Pept. Sci.* **14**, 416-422 (2008).

23 Fantini, J., Carlus, D. & Yahi, N. The fusogenic tilted peptide (67-78) of alpha-synuclein is a cholesterol binding domain. *Biochim. Biophys. Acta* **1808**, 2343-2351 (2011).

24 Di Scala, C., Chahinian, H., Yahi, N., Garmy, N. & Fantini, J. Interaction of Alzheimer's beta-amyloid peptides with cholesterol: Mechanistic Insights into Amyloid Pore Formation. *Biochemistry* **53**, 4489-4502 (2014).

25 Fantini, J. & Yahi, N. Molecular insights into amyloid regulation by membrane cholesterol and sphingolipids: common mechanisms in neurodegenerative diseases. *Exp. Rev. Mol. Med.* **12**, e27 (2010).

26 Fantini, J. *et al.* Bexarotene blocks calcium-permeable ion channels formed by neurotoxic Alzheimer's ss-amyloid peptides. *ACS Chem. Neurosci.* **5,** 216-224 (2014).

27 Di Scala, C. *et al.* Biochemical identification of a linear cholesterol-binding domain within Alzheimer's beta amyloid peptide. *ACS Chem. Neurosci.* **4**, 509-517 (2013).

28 Di Scala, C. *et al.* Mechanism of cholesterol-assisted oligomeric channel formation by a short Alzheimer beta-amyloid peptide. *J. Neurochem.* **128,** 186-195 (2014).

29 Lal, R., Lin, H. & Quist, A. P. Amyloid beta ion channel: 3D structure and relevance to amyloid channel paradigm. *Biochim. Biophys. Acta* **1768**, 1966-1975 (2007).

30 Quist, A. *et al.* Amyloid ion channels: a common structural link for protein-misfolding disease. *Proc. Natl. Acad. Sci. USA* **102**, 10427-10432 (2005).

31 Fantini, J. & Yahi, N. The driving force of alpha-synuclein insertion and amyloid channel formation in the plasma membrane of neural cells: key role of ganglioside- and cholesterol-binding domains. *Adv. Exp. Med. Biol.* **991**, 15-26 (2013).

32 Blumenthal, R., Durell, S. & Viard, M. HIV entry and envelope glycoprotein-mediated fusion. *J. Biol. Chem.* **287**, 40841-40849 (2012).

33 Jang, H. *et al.* Alzheimer's disease: which type of amyloid-preventing drug agents to employ? *Phys. Chem. Chem. Phys.* **15**, 8868-8877 (2013).

34 Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62-87 (2004).

35 Paila, Y. D., Tiwari, S. & Chattopadhyay, A. Are specific nonannular cholesterol binding sites present in G-protein coupled receptors? *Biochim. Biophys. Acta* **1788**, 295-302 (2009).

36 Li, H. & Papadopoulos, V. Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. *Endocrinology* **139**, 4991-4997 (1998).

37 Epand, R. M. Proteins and cholesterol-rich domains. *Biochim. Biophys. Acta* **1778**, 1576-1582 (2008).

38 Palmer, M. Cholesterol and the activity of bacterial toxins. *FEMS Microbiol. Lett.* **238**, 281-289 (2004).

39 Epand, R. M. Cholesterol and the interaction of proteins with membrane domains. *Progress in lipid research* **45**, 279-294, doi:10.1016/j.plipres.2006.02.001 (2006).

40 Nishio, M., Umezawa, Y., Fantini, J., Weiss, M. S. & Chakrabarti, P. CH–π hydrogen bonds in biological macromolecules. *Phys. Chem. Chem. Phys.* **16,** 12648-12683 (2014).

41 Baier, C. J., Fantini, J. & Barrantes, F. J. Disclosure of cholesterol recognition motifs in transmembrane domains of the human nicotinic acetylcholine receptor. *Sci. Rep.* **1**, 69 (2011).

42 Fantini, J., Di Scala, C., Evans, L. S., Williamson, P. T. & Barrantes, F. A mirror code for protein-cholesterol interactions in the two leaflets of biological membranes. *Sci. Rep.* **6**, 21907 (2016).

43 Quiocho, F. A. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu. Rev. Biochem.* **55**, 287-315 (1986).

44 Fantini, J. How sphingolipids bind and shape proteins: molecular basis of lipid-protein interactions in lipid shells, rafts and related biomembrane domains. *Cell. Mol. Life Sci.* **60**, 1027-1032 (2003).

45 Maresca, M., Derghal, A., Caravagna, C., Dudin, S. & Fantini, J. Controlled aggregation of adenine by sugars: physicochemical studies, molecular modelling simulations of sugar-aromatic CH-pi stacking interactions, and biological significance. *Phys. Chem. Chem. Phys.***10**, 2792-2800 (2008).

46 Charloteaux, B., Lorin, A., Brasseur, R. & Lins, L. The "Tilted Peptide Theory" links membrane insertion properties and fusogenicity of viral fusion peptides. *Protein Pept. Lett.* **16**, 718-725 (2009).

47 Gal, Z. *et al.* Mutations of the central tyrosines of putative cholesterol recognition amino acid consensus (CRAC) sequences modify folding, activity, and sterol-sensing of the human ABCG2 multidrug transporter. *Biochim. Biophys. Acata* **1848**, 477-487 (2015).

48 Telbisz, A., Hegedus, C., Varadi, A., Sarkadi, B. & Ozvegy-Laczka, C. Regulation of the function of the human ABCG2 multidrug transporter by cholesterol and bile acids: effects of mutations in potential substrate and steroid binding sites. *Drug Metab. Dispos.* **42**, 575-585 (2014).

49 Clay, A. T., Lu, P. & Sharom, F. J. Interaction of the P-Glycoprotein Multidrug Transporter with Sterols. *Biochemistry* **54**, 6586-6597 (2015).

50 Sharpe, L. J. *et al.* Cholesterol sensing by the ABCG1 lipid transporter: Requirement of a CRAC motif in the final transmembrane domain. *Biochim. Biophys. Acta* **1851**, 956-964 (2015).

51 Hanson, M. A. *et al.* A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure* **16**, 897-905 (2008).

52 Rosenhouse-Dantsker, A., Noskov, S., Durdagi, S., Logothetis, D. E. & Levitan, I. Identification of novel cholesterol-binding regions in Kir2 channels. *J. Biol. Chem.* **288**, 31154-31164 (2013).

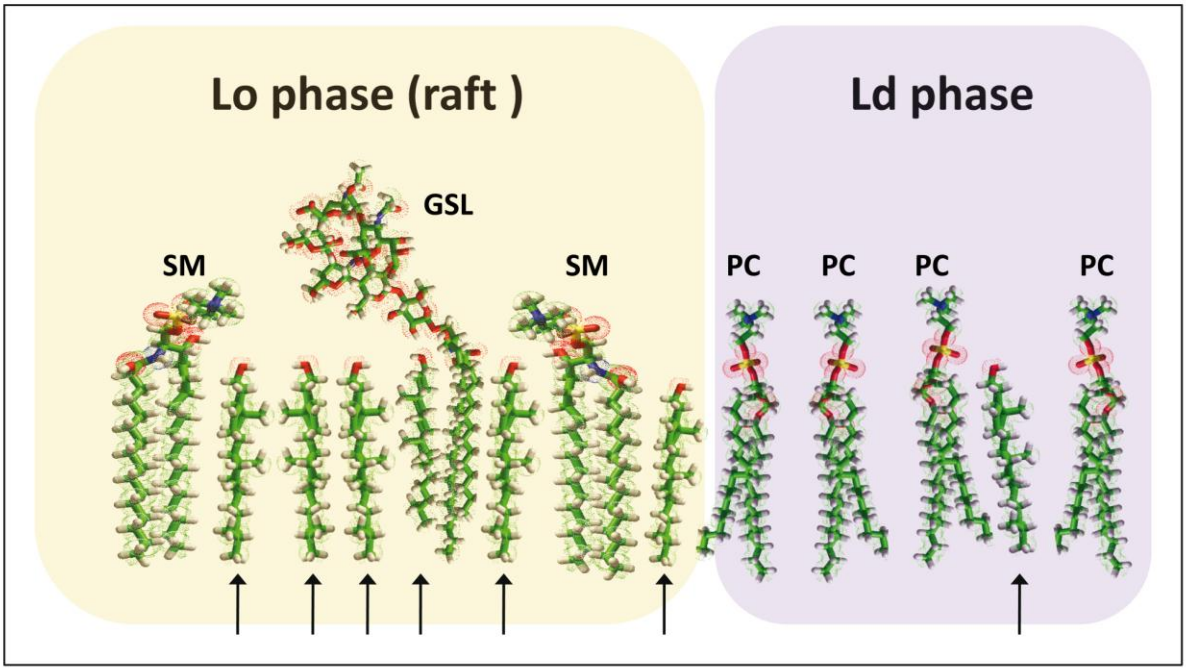
53 Levitan, I., Singh, D. K. & Rosenhouse-Dantsker, A. Cholesterol binding to ion channels. *Front. Physiol.* **5,** 65 (2014).

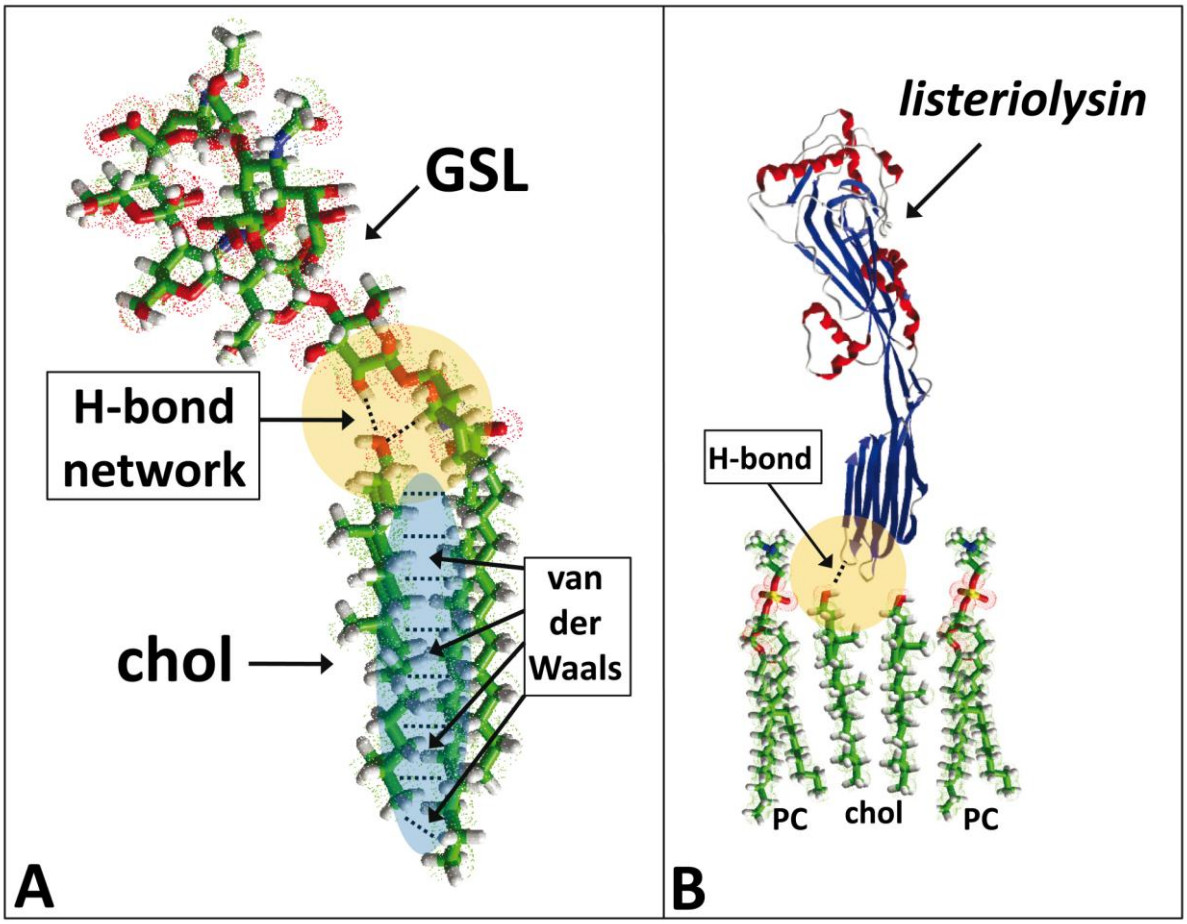
54 Strandberg, E. & Killian, J. A. Snorkeling of lysine side chains in transmembrane helices: how easy can it get? *FEBS letters* **544**, 69-73 (2003).

Table1: Comparison of the amino acid sequences of tilted peptides, CARC and CRAC motifs

Tilted peptides	
HIV-1 (P04578)	512-AVGIGAL**F**LG**F**LGAAGSTMGA-532
HTLV-I (P03381)	313-AVPVAV**W**LVSALAMGAGVAGG-333
Ebola virus (P87671)	524-GAAIGLA**W**IP**YF**GPAA-539
Bovine leukemia virus (P25506)	304-AAALTLGLALSVGLTGINVAV-324
Prion protein (P04156)	118-AGAVVGGLGG**Y**MLGSAMS-135
Aβ (P05067)	29- GAIIGLMVGGVVIA-42
α-synuclein (P37840)	67- GGAVVTGVTAVA-78
CARC motifs	
AChR (P02708)	254-**R**LPL**YF**I**V**N**V**-263
5HT-3 (P46098)	241-**R**PL**FY**V**V**S**LLL**-251
CB1 (P21554)	376-**K**TV**F**A**F**CSM**L**C**L**-387
Tachykinin NK1 Rec. (P25103)	63-**R**TVTN**YFLV**N**L**-73
mGluR1 (Q14416)	603-**R**ELC**Y**I**LL**-610
β2-adrenergic Rec. (P07550)	304-**RK**EV**Y**I**LL**-311
$GABA_B$ Rec. 1 (Q9UBS5)	590-**K**L**F**IS**V**S**VL**-598
CRAC motifs	
β2-adrenergic Rec. (P07550)	212-**LV**IMV**F**V**Y**S**R**-221
$GABA_B$ Rec. 1 (Q9UBS5)	783-**L**GI**F**LA**Y**ET**K**-792
Somatostatin Rec. Type 3 (P32745)	311-**L**YG**F**LS**YRFK**-320
Histamine H1 Rec. (P35367)	204-**LL**M**LWFY**A**K**I**YK**-215
Adenosine Rec. A1 (P30542)	287-**VY**A**FR**IQ**K**-294
mGluR5 (P41594)	812-**V**A**L**GCM**F**VP**K**-821
5HT-7 (P34969)	103-**V**IS**V**C**F**V**KK**L**R**-113

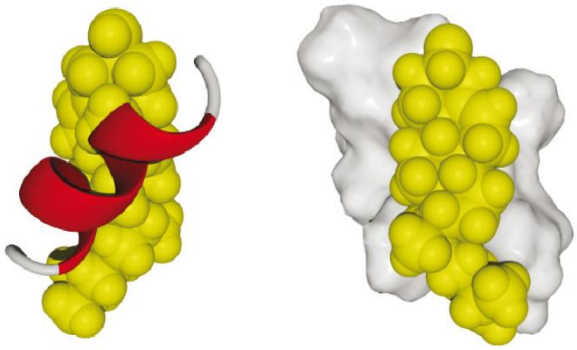
The Uniprot code of each protein is noted in the brackets. Aromatic residues are written in red and highlighted in cyan. Basic residues (Lys/Arg) defining the consensus CARC and CRAC motifs are highlighted in yellow. Aliphatic residues (Leu/Val) defining CARC and CRAC are highlighted in green.





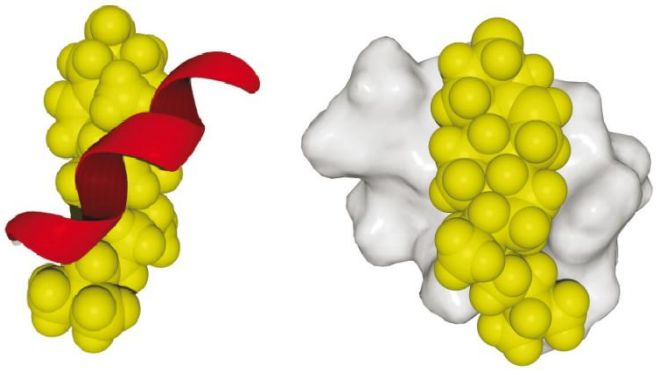
HIV-1 fusion peptide

AVGIGALFLGFLGAAGST



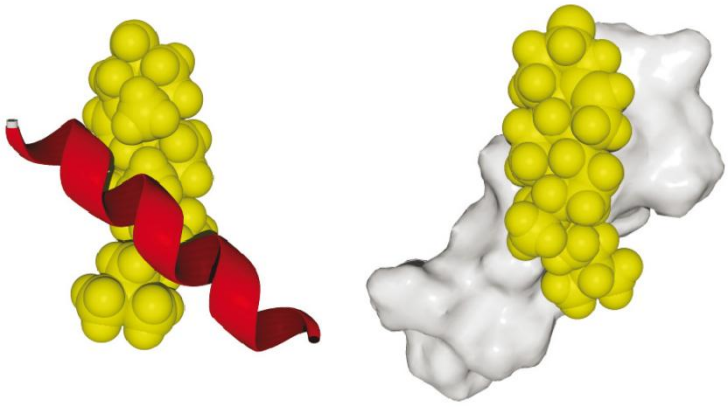
Ebola fusion peptide

GAAIGLAWIPYFG



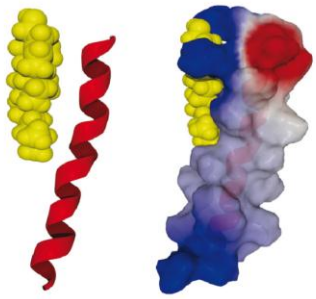
α-synuclein

GGAVVTGVTAVAQ



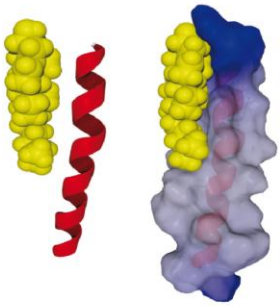
CB1 TM3

392-VGSLFLTAIDR-400



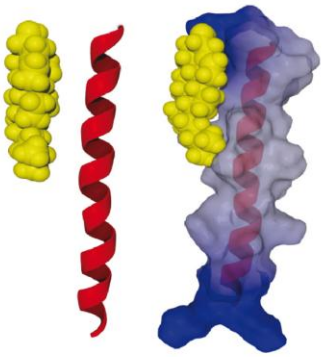
CB1 TM7

204-VNPIIYALR-214



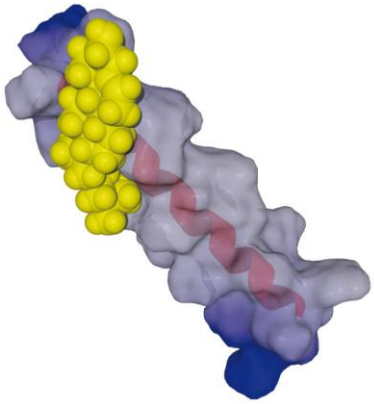
$GABA_B$ TM1

550-LFFNIK-555



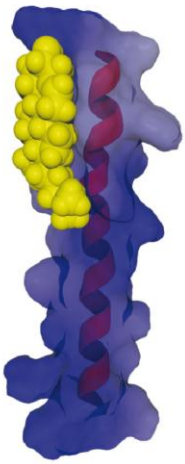
Pro-RP Rec. TM5

221-RQLYAWGLLLV-231



A1 Rec. TM7

265-KPSILTYIAIFL-276



TRVP1

468-KTGDYFRV-475