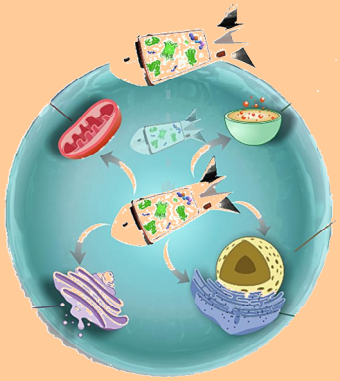
Cellular Biochemistry [918SM]



Module 3

Protein Sorting Pathways I

3_1 Targeting of proteins to the RER

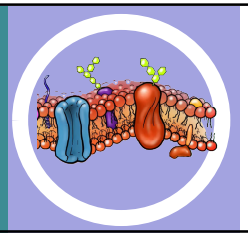
3_2 Targeting membrane proteins

3_3 Folding and PTM of proteins in the ER

A. Tossi atossi@units.it

AA 2026-2027

Cellular Biochemistry [918SM]



Modulo 3-2
Insertion / anchoring of membrane proteins

A. Tossi atossi@units.it

AA 2025-2026

MEMBRANE PROTEINS: GENERAL CHARACTERISTICS

- **Orientation**: All membrane proteins (MP) have a **unique orientation** with respect to the membrane's phospholipid bilayer (irrespective of membrane – *cyto* / *exo*)
- **Integral MP:** synthesized in the RER, they **remain embedded in the membrane** as they move to their final destination (eg. by vesicular traffic or membrane contact)
- **Transport:** During transport, the MP's **orientation (cyto/exo) is preserved**; the **topology** (N-term and C-term location, etc.) is **established during biosynthesis** on the ER membrane.
- **Membrane insertion:** uses the **similar translocation mechanisms as for soluble proteins** insertion into the ER lumen
- **Translocation:** there are **several, partly redundant**, targeting and translocation complexes.
- **Signal sequences:** The translocation process is **guided by SS, but more complex** as they also determine the MP's orientation
- **Topology:** The topological orientation of **MP transmembrane segments** and **cyto/exo hydrophilic regions** depends on the arrangement of specific hydrophobic sequences and flanking charged AA residues.
- **Topogenic sequences:** direct membrane insertion and orientation of various types of integral MP is guided by what are collectively known as '*topogenic sequences*'
- **Classification:** MPs divided into various types/classes based on different topologies,

CLASSES OF INTEGRAL MEMBRANE PROTEINS

- ***Topology* of a membrane proteins. This refers to:**

  1) number membrane spans* by polypeptide chain the; 2) orientation of membrane-spanning segments; 3) location of termini (if N-terminus or C-terminus are *cyto* or *exo*)

Bitopic (1 TM segment)

1. Type I (single pass, C_{cyto} - N_{exo})
2. Type II (single pass, N_{cyto} - C_{exo})
3. Type III (sp, Head-anchored, C_{cyto} - N_{exo})
4. Tail anchored (sp, N_{cyto} - C_{exo})

Polytopic (multipass, >1 TM segments)

5. Type IV

   IV-A (N_{cyto}/ C_{cyto}, even Nº TM, same side)

   IV-B (N_{exo} / C_{cyto}, odd NºTM, opposite side)

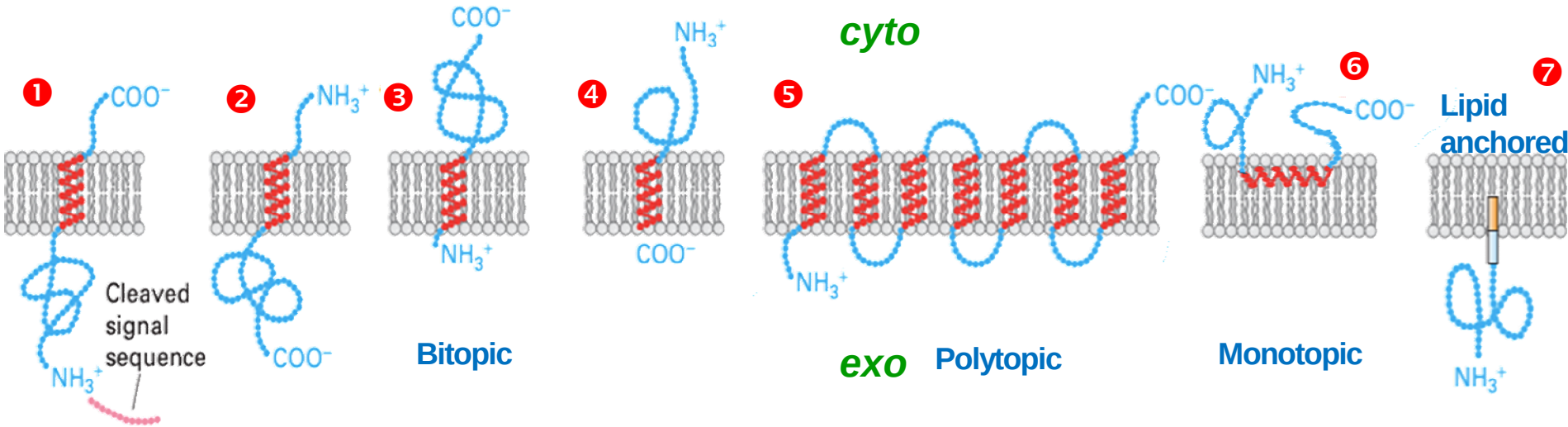
Monotopic (no TM segment)

6. Inserted on one side of membrane

Lipid anchored (to membrane lipids)

7. Type V GPI-anchored

(NB classification numbering is not fixed – can be confusing)*



LIPID ANCHORED MEMBRANE PROTEINS

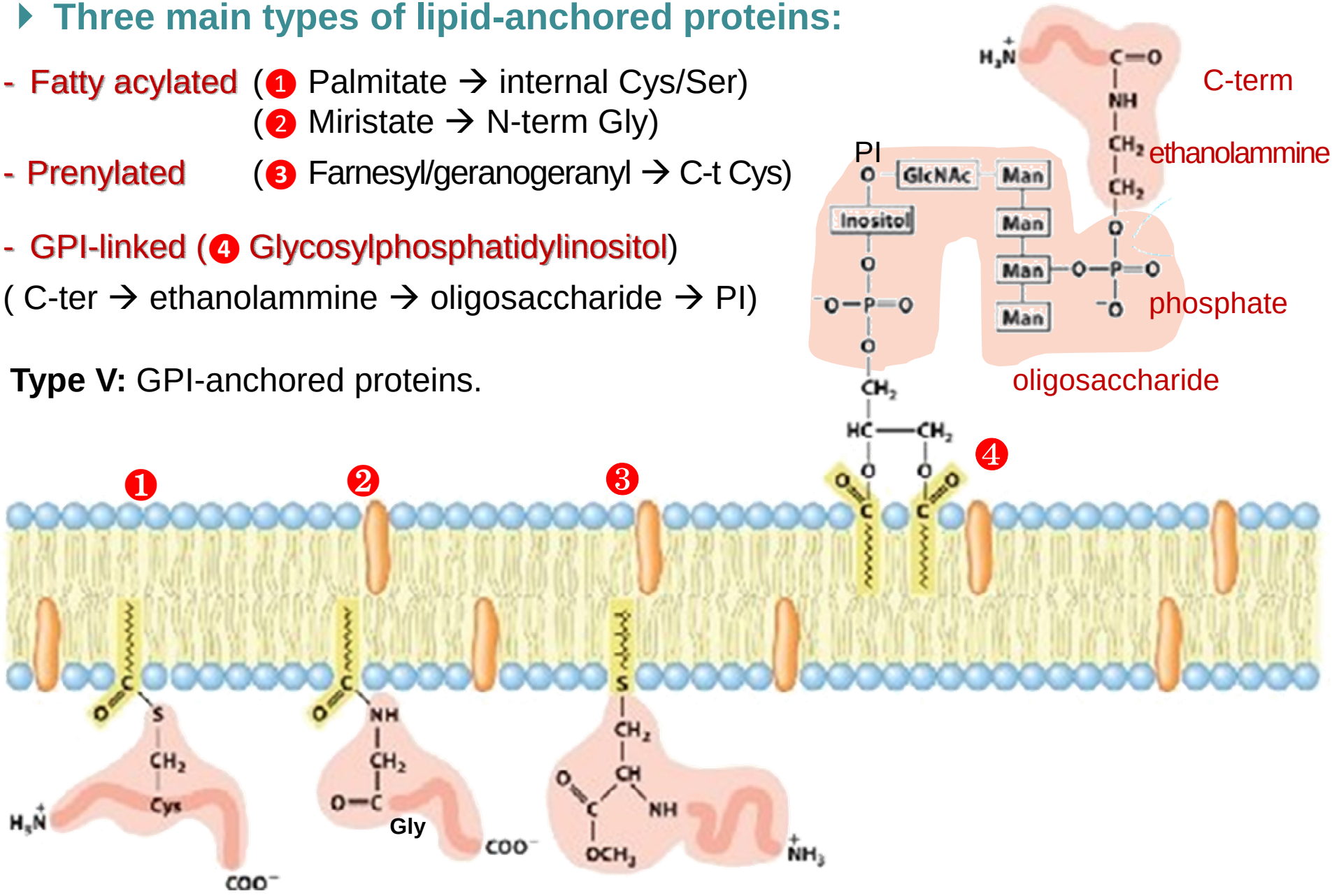
Three main types of lipid-anchored proteins:

- Fatty acylated (❶ Palmitate → internal Cys/Ser)
  (❷ Miristate → N-term Gly)
- Prenylated (❸ Farnesyl/geranogeranyl → C-t Cys)

- GPI-linked (❹ Glycosylphosphatidylinositol)

(C-ter → ethanolammine → oligosaccharide → PI)

Type V: GPI-anchored proteins.



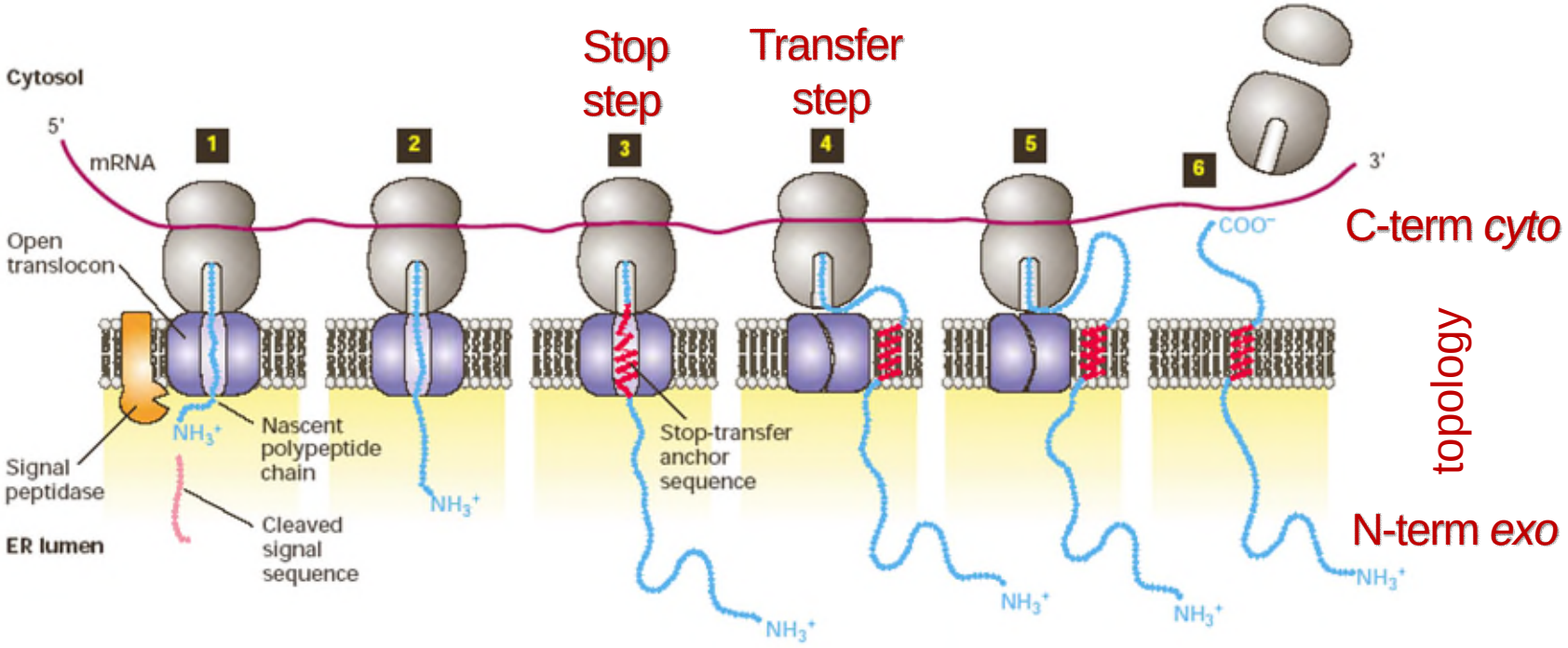
TOPOGENIC SIGNAL SEQUENCES

- **There are 3 types of Signal Sequence that direct TM segment topology**

Topogenic Signal Sequence	Protein Type	SS final location
N-terminal Signal Sequence (N-SS)	• Soluble protein (*exo*) • Type I IMP	✂ cleaved into membrane
Stop-Transfer Anchor seq. (STA)	• Type I IMP	→ TM domain
Signal-Anchor sequence II (SA-II)	• Type II IMP	→ TM domain
Signal-Anchor sequence III (SA-III)	• Type III IMP (head anchored)	→ TM domain
C-term Signal-Anchor (C-SA)	Tail Anchored IMP	→ TM domain

MEMBRANE INSERTION AND ORIENTATION – TYPE I

- **Type I TM proteins - carry two types of topogenic sequences**
  - N-terminal Signal Sequences (SS) (as in soluble proteins) → *cleaved after translation*
  - internal Stop-Transfer Anchor Sequences (STA) ~22 AA hydrophobic stretch → *TMD*



❶ SS recognised by SRP; ribosome → ER Translocon; NC enters ER lumen, SS cleaved

❷ ❸ Chain elongates; hydrophobic STA enters Translocon; Translocation stops (Stop step)

❹ STA is hydrophobic, it moves through the lateral exit of the Translocon (*hydrophobic cleft*) into the membrane bilayer (Transfer step); the Translocon closes.

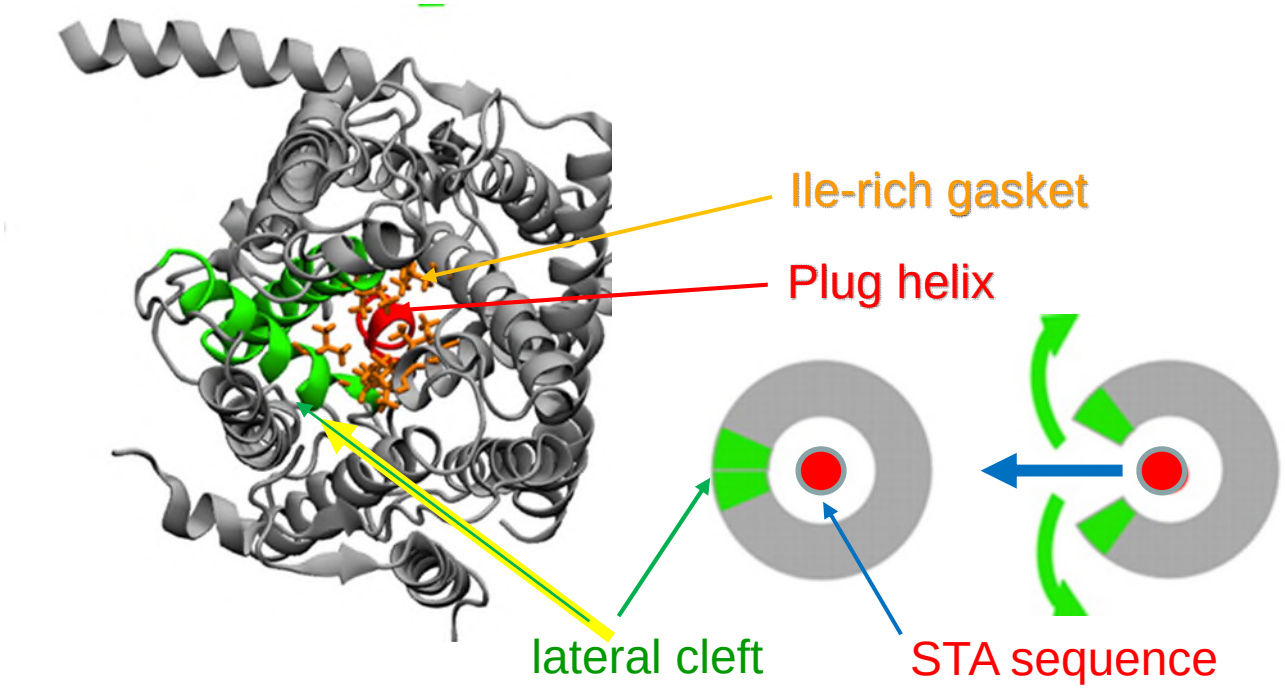
❺ Translation starts again with NC now extruded into the cytoplasm

❻ Translation ends; the ribosome is released. IMP topology: N_{exo} / C_{cyto}

EXPERIMENTAL EVIDENCE FOR TYPE I TM PROTEIN INSERTION

- **Human growth hormone (HGH) and LDL receptors are Type I TM proteins**

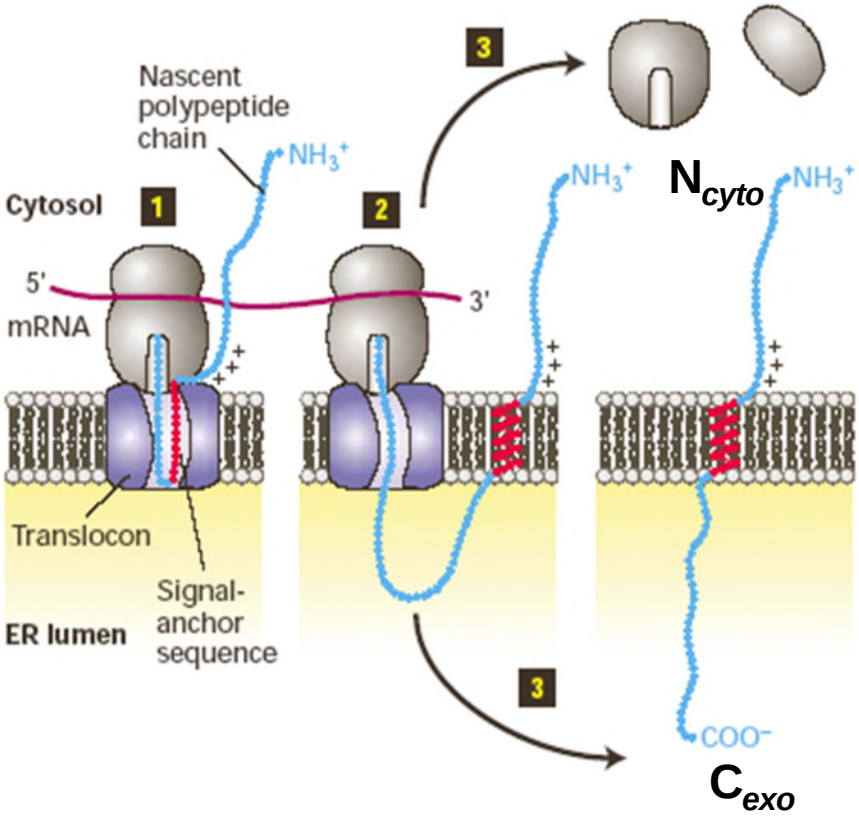
- Experimental studies confirm mode of Type I TM protein insertion
- HGHR & LDLR receptor proteins both reside in the cytoplasmic membrane facing *exo*
- wt *HGHR* cDNA was expressed in cultured cells HGHR → cytoplasmic membrane
- mutant *HGHR* (ΔSTA or altered STA) HGHR → exoplasmic space
- STA moves through the lateral exit of the Translocon



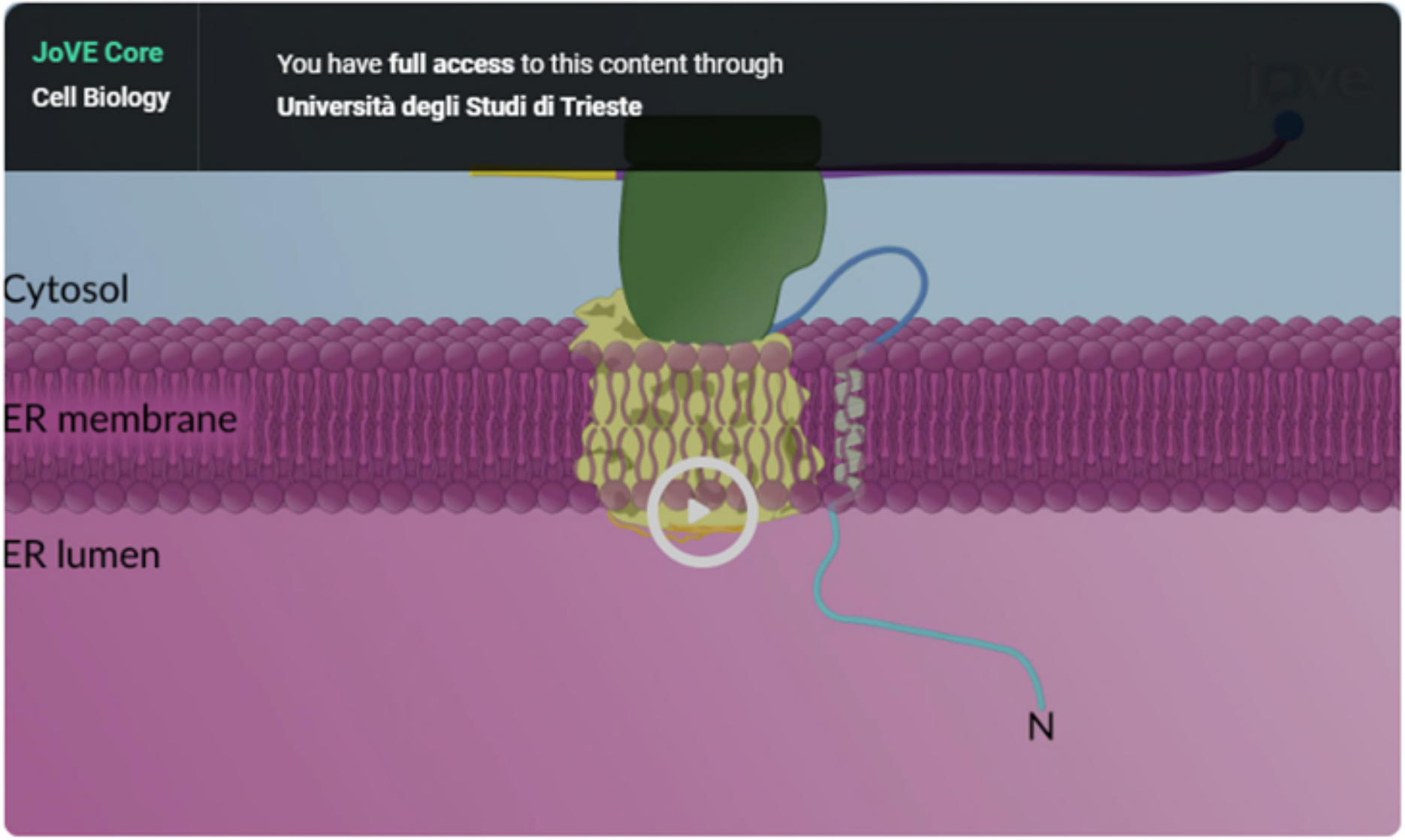
MEMBRANE INSERTION AND ORIENTATION – TYPE II

Type II TM proteins have specific Topogenic Sequences

- N-terminal signal sequence absent (no SS)
- Internal Signal Anchor sequences (SA-II) - hydrophobic sequence → *TM* (NB not STA)
- The ribosome extrudes a nascent chain (NC) into the cytoplasm until the single, internal, hydrophobic SA-II is reached.
- SRP binds SA-II and translation stops.
- SRP guides the NC to the Translocon ❶ and inserts SA-II with its C-term → *exo*
- This orientation is determined by cationic residues N-terminal to the SA-II sequence*
- Being hydrophobic, SA-II moves through lateral exit into the membrane ❷
- NB. the Translocon does not close and translation can restart so the rest of the NC is translocated into the ER lumen (C-*exo*) ❸
- The SA-II sequence is not recognized by *Signal peptidase* , and remains in the protein as the TM domain



Insertion of Single-pass Transmembrane Proteins in the RER

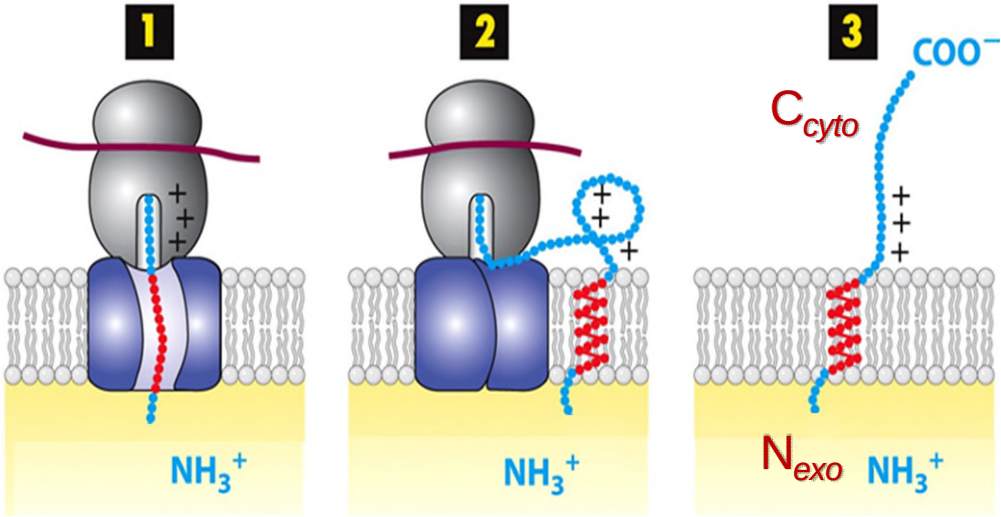


https://app.jove.com/embed/player?id=12127&access=c0faa2dd18&t=1&s=1&fpv=1

MEMBRANE INSERTION AND ORIENTATION – TYPE III

Type III TM proteins are "Head Anchored" IMP

- They carry a hydrophobic Stop-Anchor (SA-III) sequence located close to the N-terminus
- No N-terminal SS, but when SA-III emerges it binds the SRP
- Translation stops ❶ and the SRP guides the NC to the Translocon and inserts it in the N_{exo} direction, due to cationic residues C-term to the SA-III
- SA-III acts as Stop-Transfer sequence ❷, it temporarily blocks translation, moving through the *lateral exit* into the membrane
- The Translocon closes (like in Type I) and translation resumes with the NC extruded into cytoplasm (C_{cyto}) ❸ – Different to Type I TMP, SA-III acts as both SS & STA
- The main aspects determining Type III behaviour are
  - *i)* very short hydrophyllic N-term region (*exo*);
  - *ii)* hydrophobic SA (acts as SS but is not cleaved);
  - *iii*) cationic C-term to SA (C_{cyto})
- Mutating the flanking cationic residues Type II and III proteins to "flip" orientation
- Type III insertion can require auxiliary insertion complexes



MEMBRANE INSERTION AND ORIENTATION – TAIL ANCHORED

Tail-Anchored IMP (TA-MP)

- have a hydrophobic topogenic sequence near the C-terminus (c-SA)
- Most of the NC has been extruded from ribosome to cytoplasm by the time c-SA emerges
- c-SA is not recognized by SRP, and not guided to Translocon Sec61
- dedicated translocation pathway involving *Get3 ATPase* dimer, which binds to c-SA

❶ *a)* *Get3* binds ATP *b)* *Get3*/ATP binds c-SA ; *c)* *Get3* guides TA-MP to the ER membrane *d)* Co-chaperones Sgt2, Get4 and Get5 assist this process

❷ *Get3* / TA-MP complex docks onto a membrane complex formed by Get1/Get2

❸ ATP hydrolysis drives the c-SA tail into the membrane, assisted by Get1/Get2

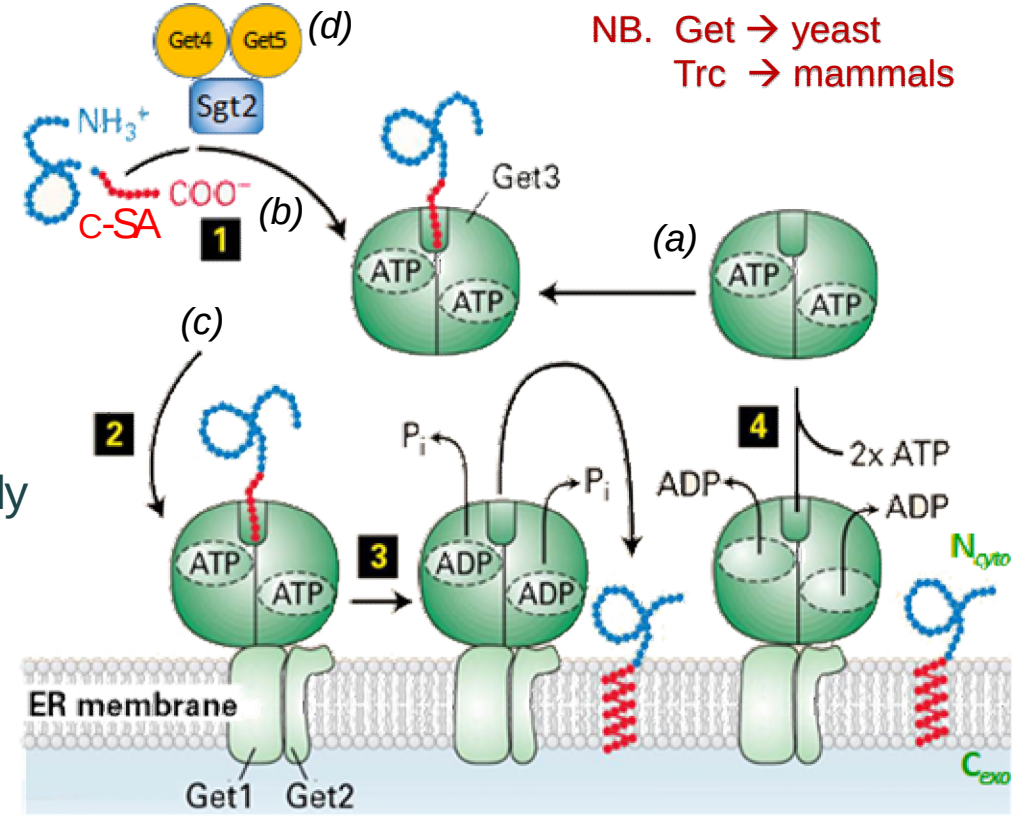
❹ ADP is released, *Get3* detaches from Get1/Get2, binds ATP and recycles to the cytosol.

- Alternative form of targeting, mechanistically similar to SRP delivery, but:

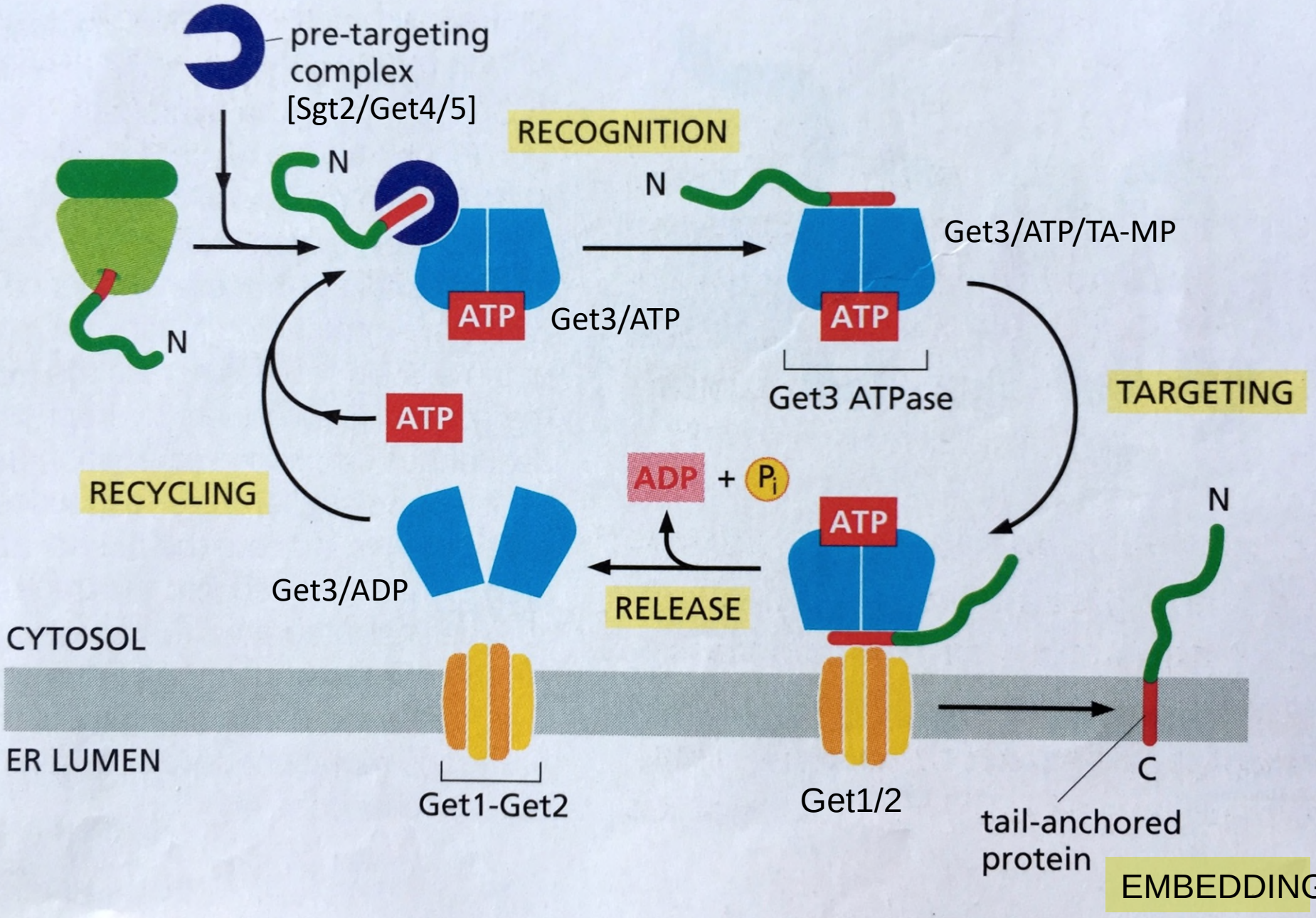
i) specific for c-SA

ii) <u>Translocon independent</u>

iii) driven by ATP (not GTP) hydrolysis



MEMBRANE INSERTION – TAIL ANCHORED MEMBRANE PROTEINS



Tail-anchoring of Proteins in the ER Membrane

https://app.jove.com/embed/player?id=12129&access=d7cebe6fa2&t=1&s=1&fpv=1

MEMBRANE INSERTION – TYPE IV 'MULTIPASS' TM PROTEINS

Many IMP proteins have more than one TM domains

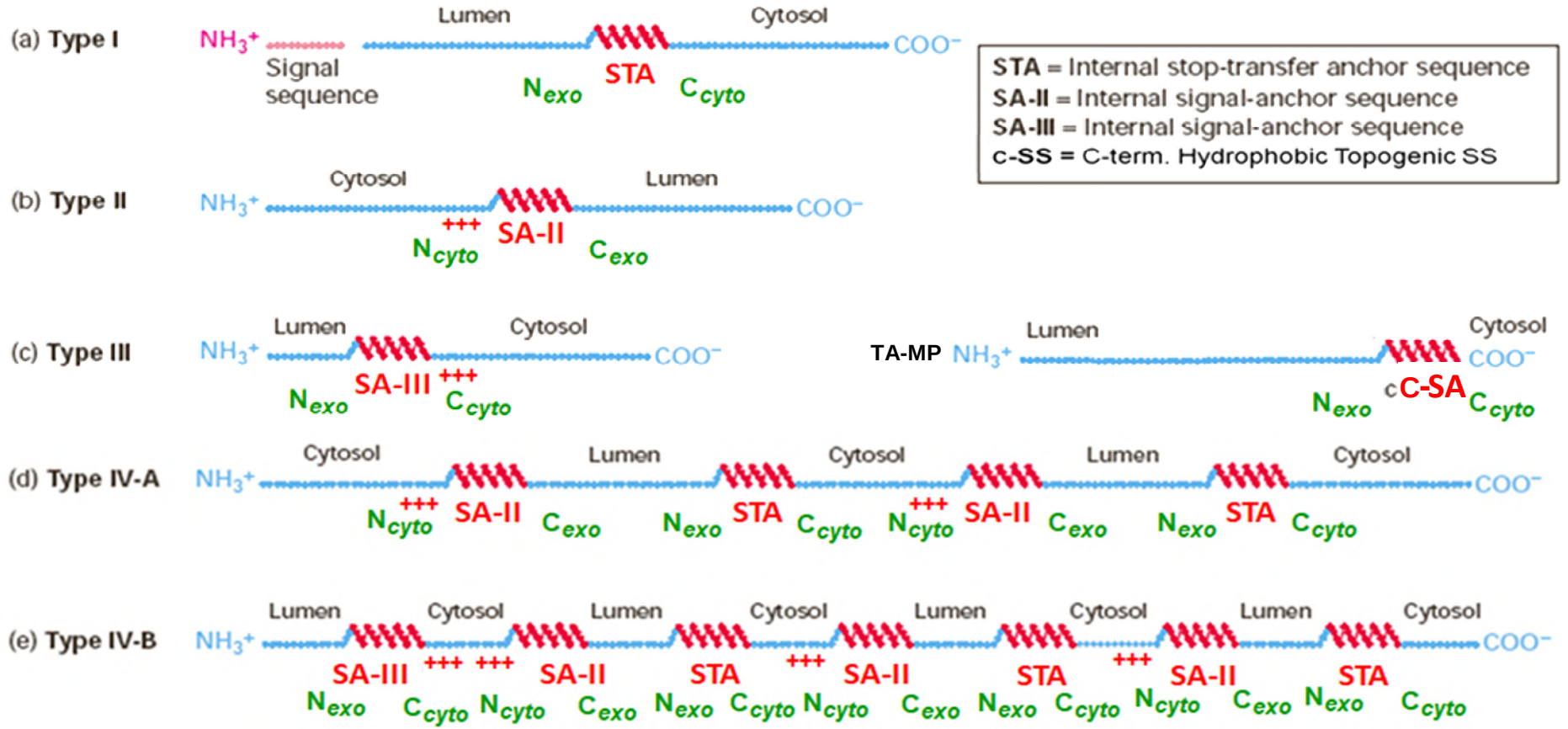
- More complex than single-pass IMP but TM domains topology follows the same principles
- **Type IV** IMP use a combination of SA-II, SA-III ,STA

Type IV-A: N_{cyto}

Type IV-B: N_{exo}

The topology of the C-terminus depends on if the Nº of TMD. If it is even they are on the same side, if odd they are on opposite sides

- Often 12 or more segments (e.g. some receptors, ion channels, transporters, pumps)



STA = Internal stop-transfer anchor sequence
SA-II = Internal signal-anchor sequence
SA-III = Internal signal-anchor sequence
c-SS = C-term. Hydrophobic Topogenic SS

MEMBRANE INSERTION – TYPE IV 'MULTIPASS' TM PROTEINS

Key principles that govern the assembly of most multipass proteins:

1) TM segments pass sideways from the Translocon into the membrane (Lateral exit)

2) This occurs co-translationally in the sequential order in which TM emerge from ribosome

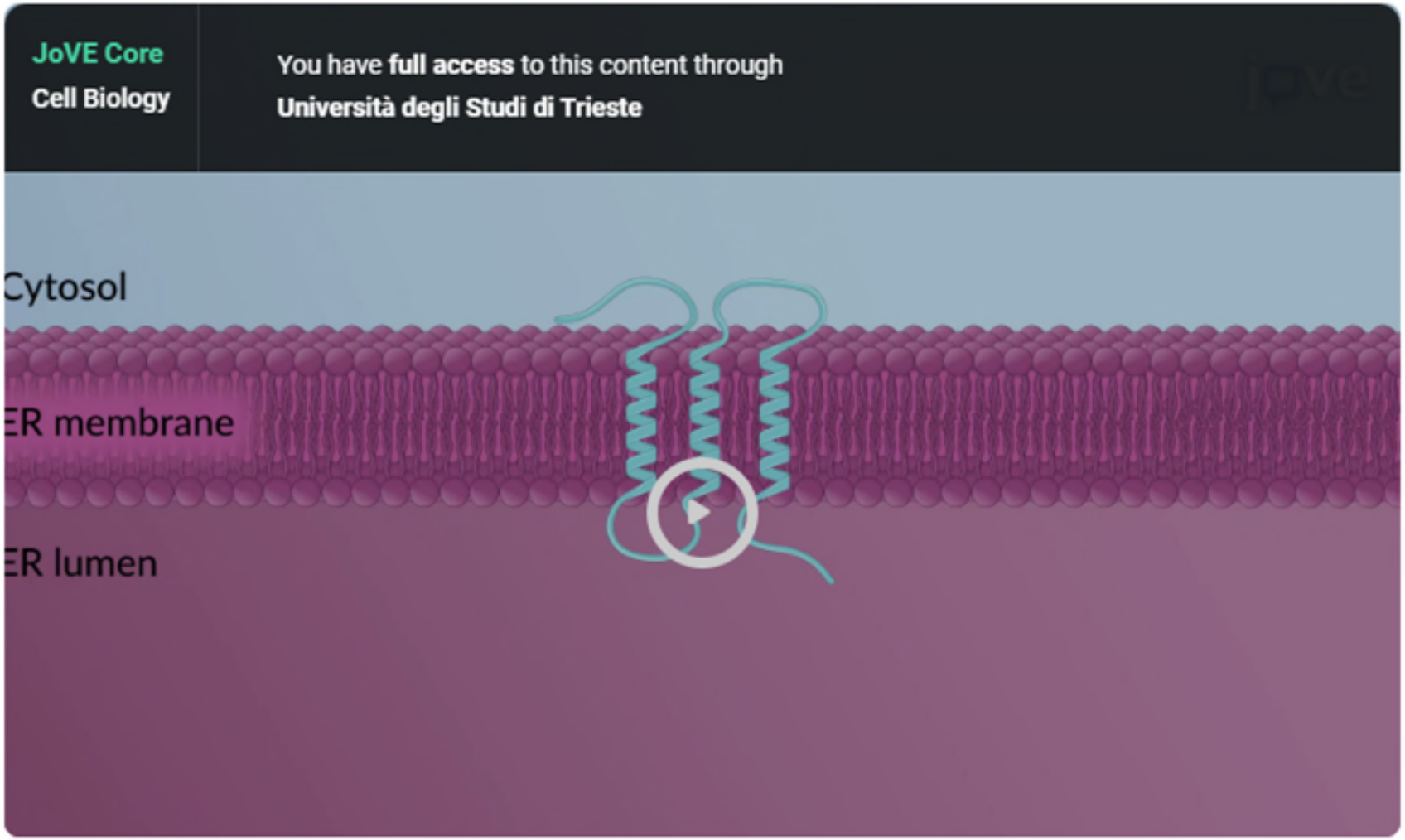
3) 1st topogenic segment engages the SRP (acts as SS) – The SRP introduces it into the Translocon in an SRP/receptor-dependent manner

4) All subsequent TM sequences engage the Translocon in an SRP-independent manner

Therefore:

- 1st TM segment acts as topogenic sequence as described for Type I (STA) , II (SA) or III (SA)
- This establishes the orientation of the 1st TM segment (N_{exo}– C_{cyto} or C_{exo}– N_{cyto})
- Successive TM segments orient in opposite orientations in an alternating manner (each TM segment in an opposite orientation to the previous one)
- The topology of the C-terminus (on the same side or opposite side of the membrane to the N-terminus) depends on the N° of TM segments (odd or even).
- most Type IV MP don't have a cleavable SS (more similar to Type III than Type I)

Insertion of Multi-pass Transmembrane Proteins in the RER



https://app.jove.com/embed/player?id=12128&access=5acaed4f72&t=1&s=1&fpv=1

Simplified model for the insertion of multipass proteins

MEMBRANE INSERTION – TYPE IV A 'MULTIPASS' PROTEINS

TYPE IV-A MP have both SA-II and STA sequences

❶ 1st SA-II emerges from Ribosome and is recognized by SRP, which stops translation and carries the NC to the SRPR and inserts it into Translocon, with N_{cyto} (≡ Type II)

❷ Translation starts, SA-II passes through the lateral exit of the Translocon, into the membrane

❸ Translation continues and STA segment emerges in the NC and is inserted into the translocon N_{exo} (≡ Type I) and moves through the cleft to the membrane

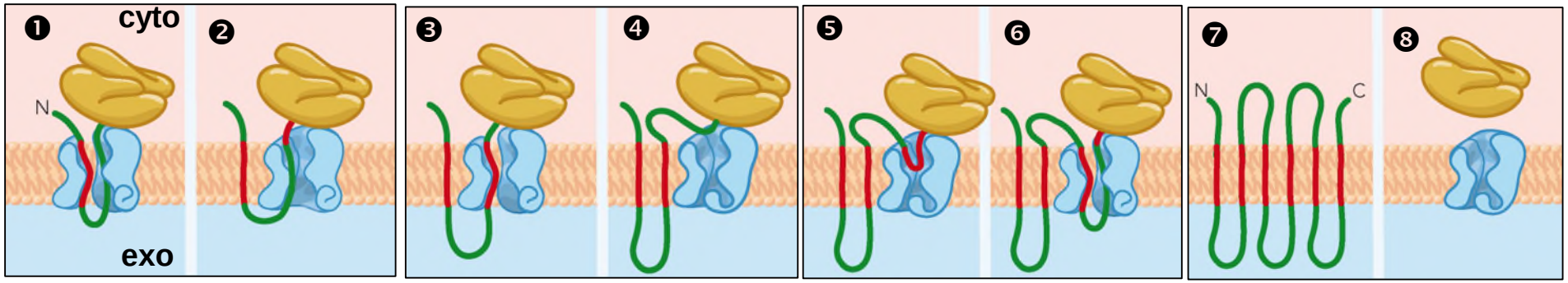
❹ The translocon channel closes temporarily and translation continues → cyto

❺ 2nd SA-II emerges in the NC and enters Sec61- Translocon reopens

❻ SA-II passes through the lateral exit to the membrane and translation continues → exo

❼ This process is repeated with the membrane insertion of alternating STA and SA-II

❽ When the C-terminus emerges the Ribosome detaches from the Translocon. If the N° of TM segments (SA-II + STA) is even, this is on the same cytoplasmic side as the N-terminus



MEMBRANE INSERTION – TYPE IV B 'MULTIPASS' PROTEINS

TYPE IV-B MP have initial SA-III followed by alternating SA-II and STA

❶ SA-III close to N-term binds the SRP; *a)* SRP guides NC to Translocon; *b)* SRP inserts SA-III into Translocon channel (N_{exo} direction due to cationic residues at C-term of SA-III)

❷ SA-III moves through lateral exit to the membrane (1st TMD) and the Translocon closes

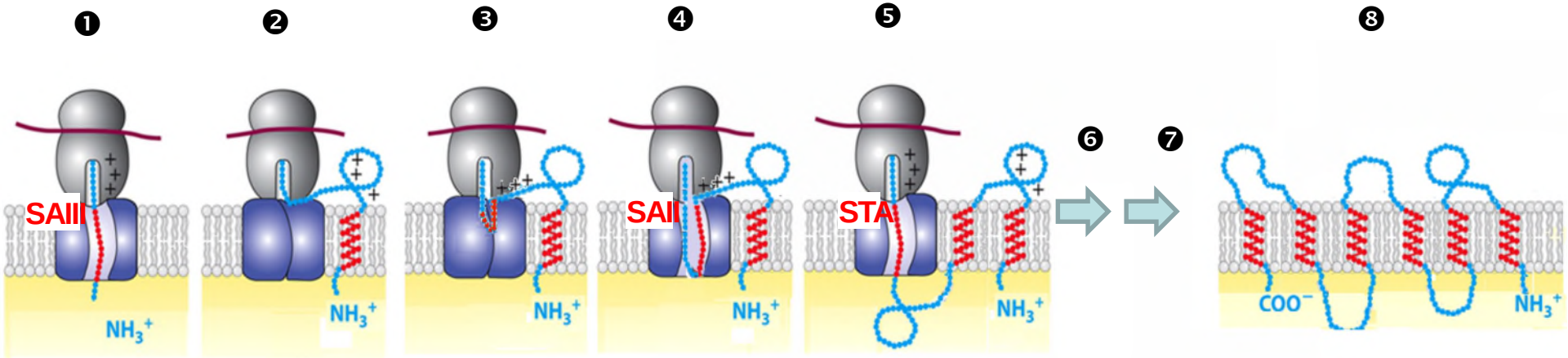
❸ Translation resumes and continues on *cyto* side until an SA-II emerges in the NC

❹ SA-II inserts into Translocon with C_{exo} topology (due to cationic residues at its N-term) *a)* Translocon re-opens *b)* SA-II moves through lateral exit becoming 2nd TMD

❺ Translation continues inserting NC into ER lumen, until STA sequence is reached - translation stops as it moves into the translocon

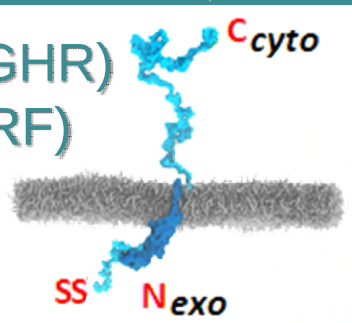
❻ STA moves through lateral exit into the membrane (3rd TMD) with opposite topology to SA-II - Translocon closes → back to step ❸

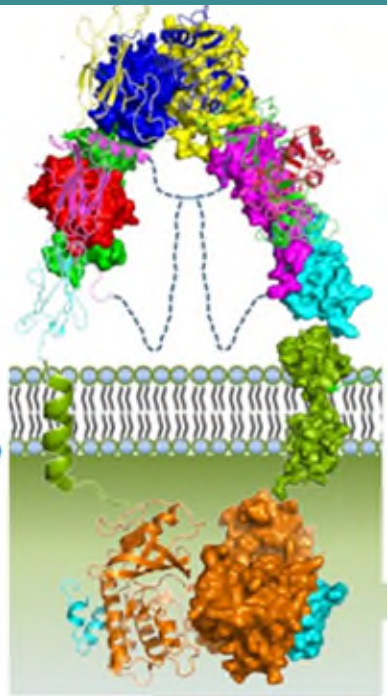
❼ More SA-II & STA enter in succession until C-term. ❽ Even N° TMD → N_{exo}/C_{exo} topology



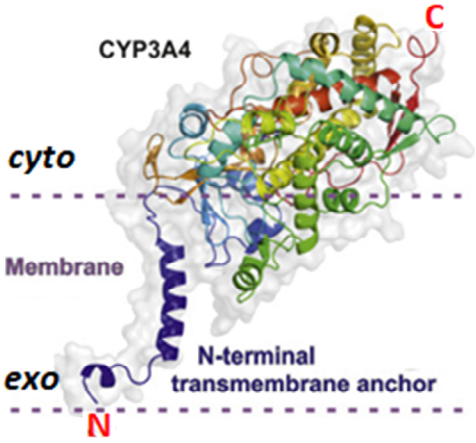
TRANSMEMBRANE PROTEINS – EXAMPLES OF DIFFERENT TYPES

- **Type I:** - Human growth hormone receptor (HGHR)
  - Epithelial growth factor receptor (EGRF)
  - Insulin receptor



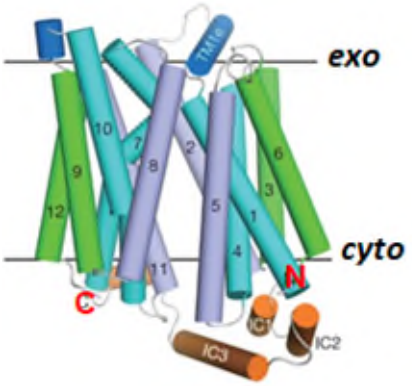


- **Type II:** - Transferrin receptor
  - Enteropeptidase

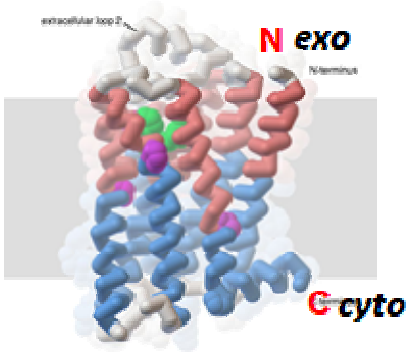


- **Type III:** - Cytochrome P450

- **Type IV-A:** - GLUT (Glucose transporters, 12 TMH)
  - Ion channels



- **Type IV-B:** - G protein coupled receptors
  - 7TMH



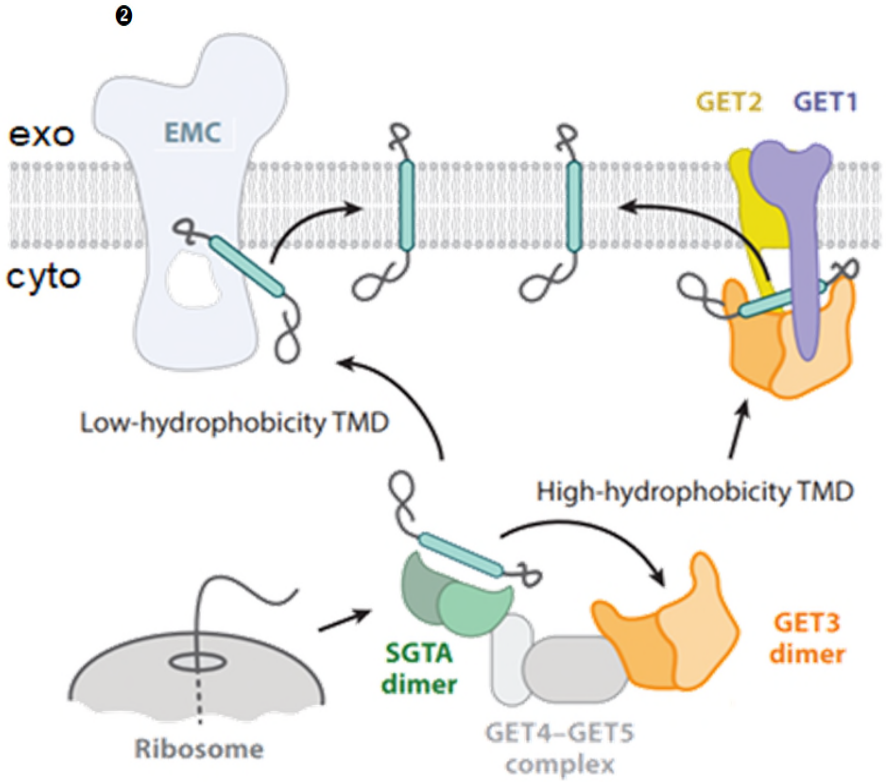
TRANSMEMBRANE PROTEINS – SUMMARY

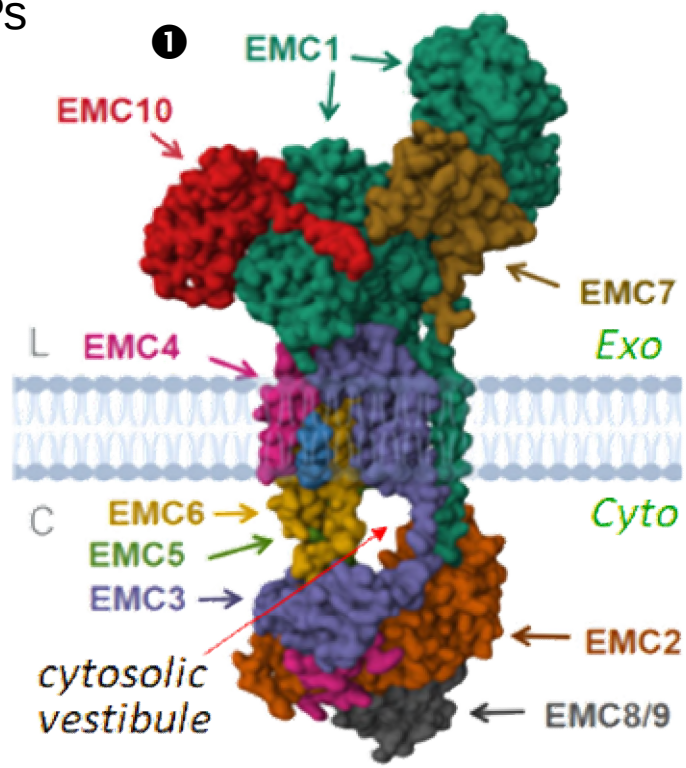
TYPE	TOPOLOGY	SIGNAL SEQUENCE	EXAMPLES	KEY NOTES
Type I	**Single-pass**; *N-term. exo*	**N-term. SS** & **STA**	Insulin receptor, EGFR	- C-term. cyto. - SS cleaved
Type II	**Single-pass**; *N-term. cyto*	Internal SA	Transferrin receptor	- No SS
Type III (Head anchored)	**Single-pass**; *N-term.exo*	Internal SA	Cytochrome P450	Orientation opposite of Type II
Tail-anchored	**Single-pass**; *C-term.exo*	C-term. SA	Cytochrome b5	Often inserted post-translationally
Type IV	**Multipass;** (multiple TM helices)	Multiple internal SA & STA	SNARE	Post-translational insertion (Get/Trc)
- Type-IVA	*N-term. cyto*	Internal SA & STA	β-adrenergic recept.	Type IV subclass
- Type-IVB	*N-term. exo*	Internal SS & STA	Na/K ATPase Some transporters	Type IV subclass

ER MEMBRANE COMPLEX (EMC) - MULTIPASS TMP BIOGENESIS

Alternative multi-protein complex involved in TMP insertion & quality control

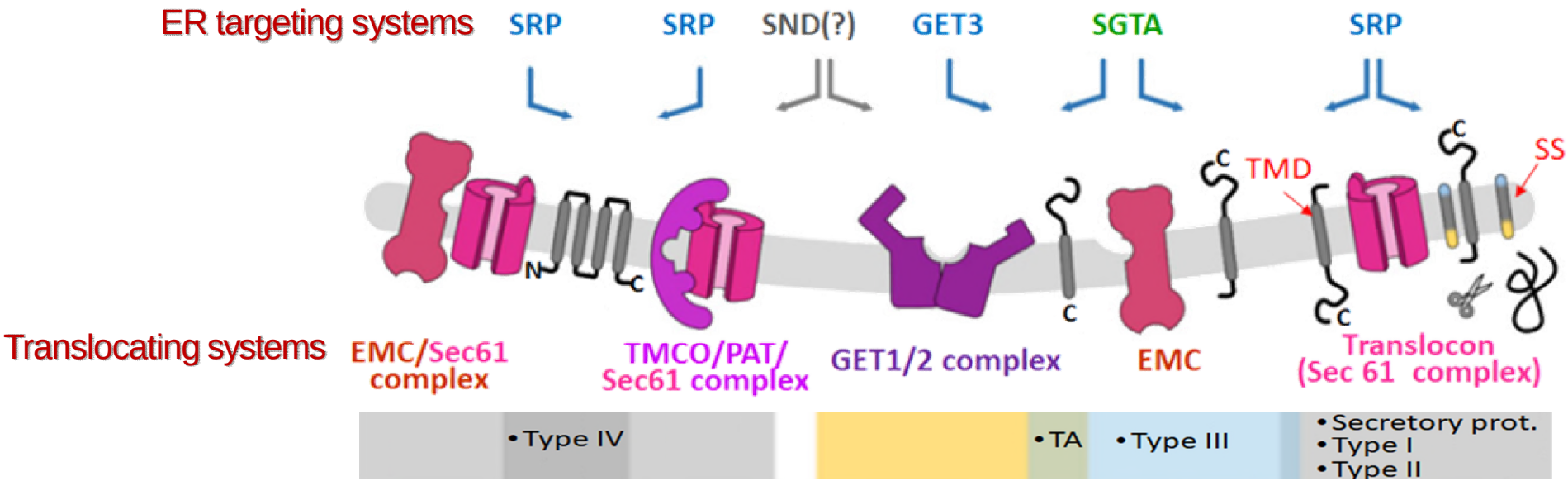
- EMC was recently identified (2009) by high-throughput genetic interaction analyses in yeast, which also identified the GET1-5 system for Tail-Anchored (TA) TMP insertion.
- Highly conserved in eukaryotes (6 subunits in yeast, 10 in vertebrates with 5 core subunits ❶)
- EMC is an *insertase* and potential chaperone for TMPs
- It replaces GET for insertion of some TA proteins ❷
- It assists the Translocon for some Type I & IV TMP





ALTERNATIVE SYSTEMS FOR TMP INSERTION INTO ER MEMBRANE

- **Several different systems ensure TMP insertion into the ER membrane**
  - Different ER targeting systems (SRP/SRPR; SGTA/GET3/GET1/2; SGTA/EMC;)
  - Different translocating machinery (Sec61; EMC; Sec61/EMC; Get1/2; TMCO/PAT/Sec61)



SRP = signal regocnition particle;

SND = SRP indpendent targetting protein

SGTA = small glutamine rich tetratricopeptide repeat co-chaperone

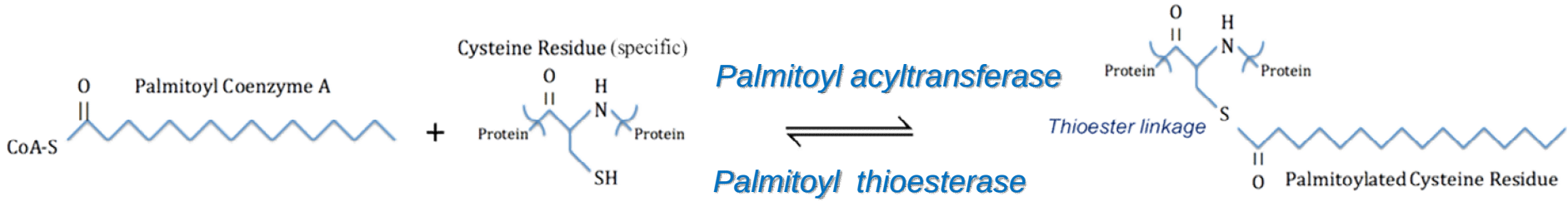
TMCO = TransMembrane Coiledcoil Protein

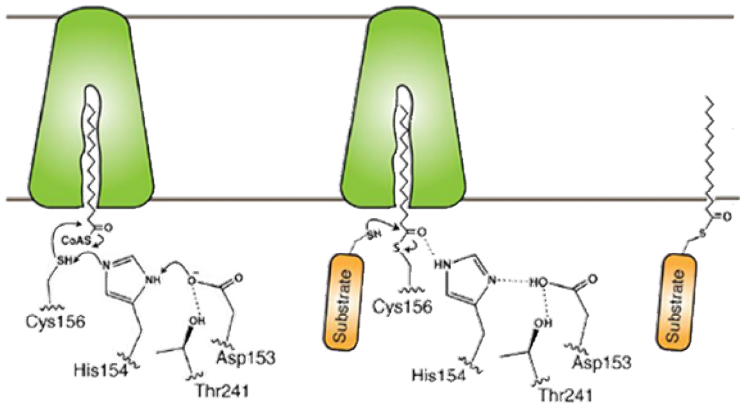
PAT = Protein associated to Translocon

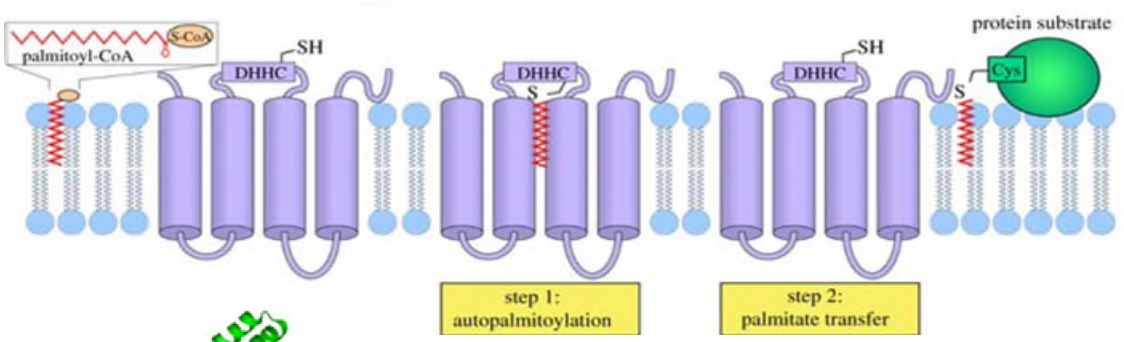
LIPID-ANCHORED PROTEINS - PALMITOYLATION

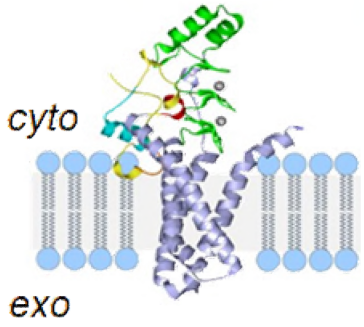
Covalent attachment of Palmitic acid (FA) to proteins

- S-palmitoylation of a Cys residue – a reversible modification
- Mainly for proteins anchored to *cyto* side of cytoplasmic membrane (e.g. G_{as} *subunit*, *Ras)*
- Reversible reaction catalyzed by *Palmitoyl acyltransferase* (with a DHHC catalytic domain)
- Catalytic triad (Asp/His/Cys) and mechanism similar to Ser proteases





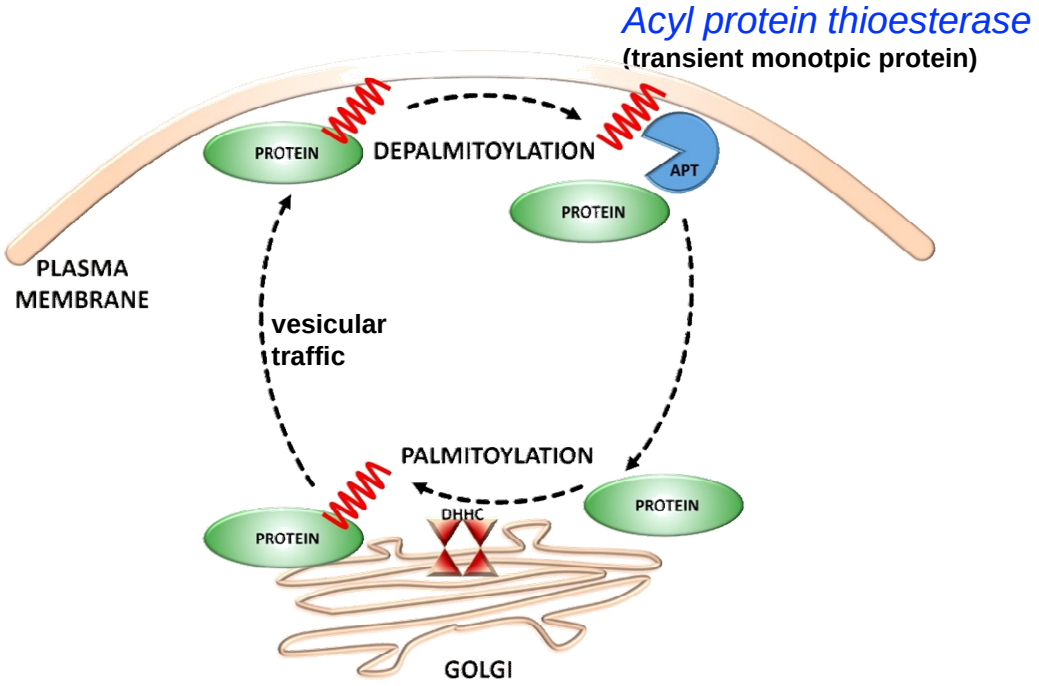




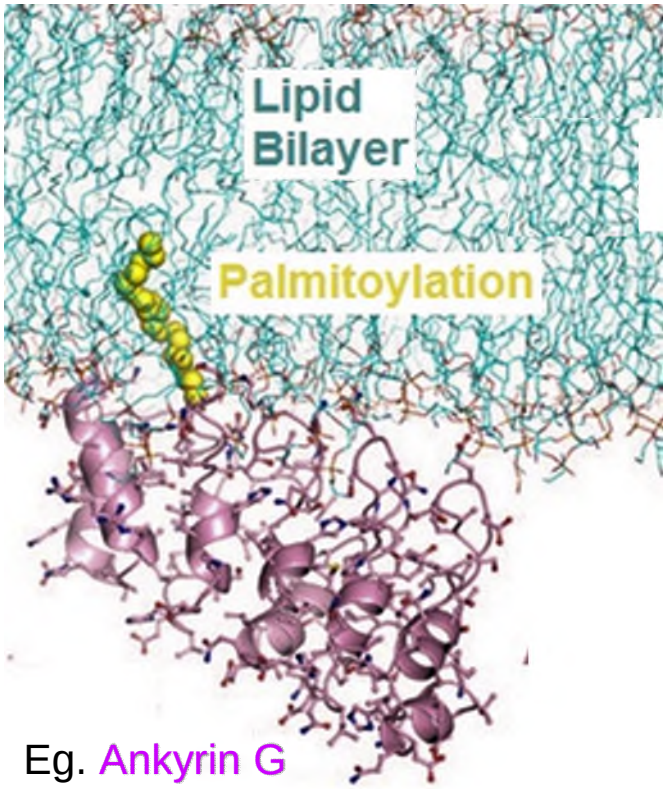
LIPID-ANCHORED PROTEINS – PALMITOYLATION (cont.)

S-Palmitoylation is a *reversible and dynamic* protein modification

- Mediated by zinc-finger *zDHHC-palmitoyltransferases* with a Asp-His-His-Cys motif and *Acyl protein thioesterases,* that respectively add and remove the acyl group
- These enzymes are membrane-associated, usually residing in the Golgi complex or cytoplasmic membrane
- They modify proteins involved in signaling, trafficking, or with structural roles



zDHHC-Palmitoyltransferase

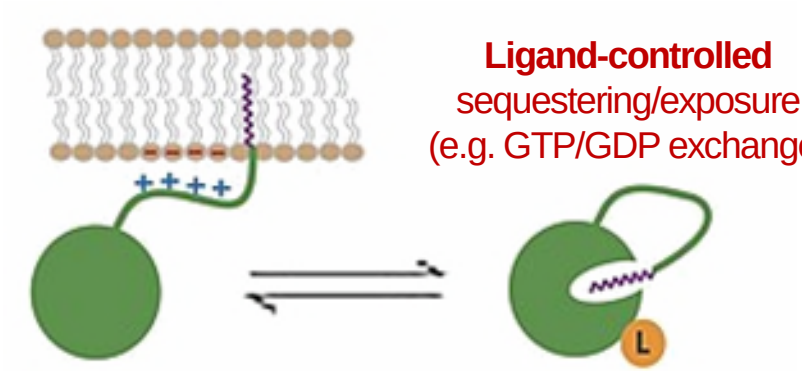


Eg. Ankyrin G

LIPID-ANCHORED PROTEINS – N-MYRISTOYLATION

An irreversible lipidation

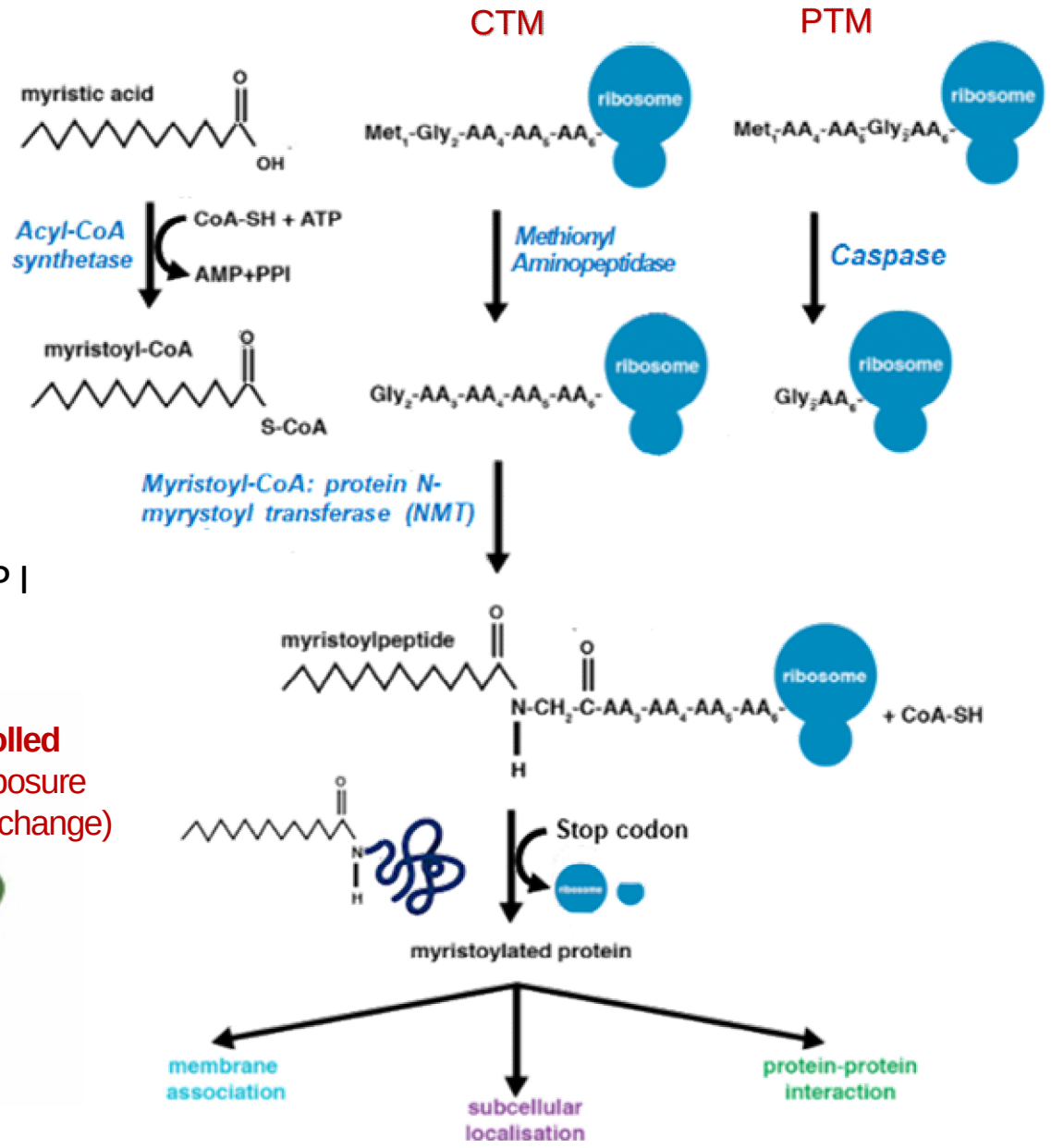
- myristoyl group covalently linked by amide bond the α-NH_2 of N-term Gly
- Occurs co-translationally after *Met aminopeptidase* removes start Met
- Occurs post-translationally - *protease* (e.g. *Caspase*) exposes internal Gly
- Catalysed by *Myristoyl-CoA:protein N-myristoyltransferase* (*NMT*)
- Molecular switch for membrane interaction e.g. *ARF kinase* acting on COP I vesicles (see module 4)



Membrane anchoring
(hydrophobic & electrostatic non-covalent interactions)

Ligand-controlled
sequestering/exposure
(e.g. GTP/GDP exchange)

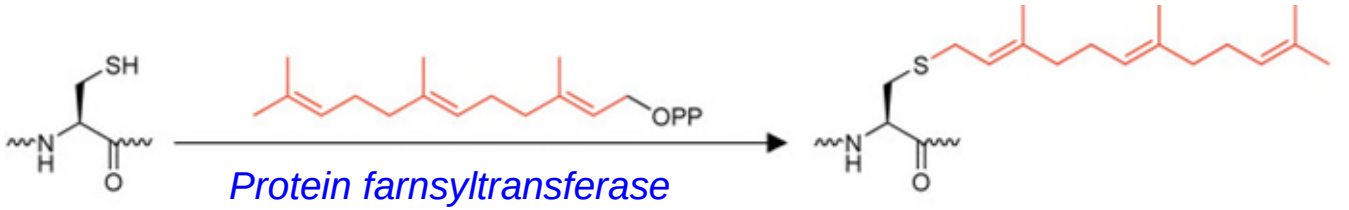
Co- & Post-translational myristoylation



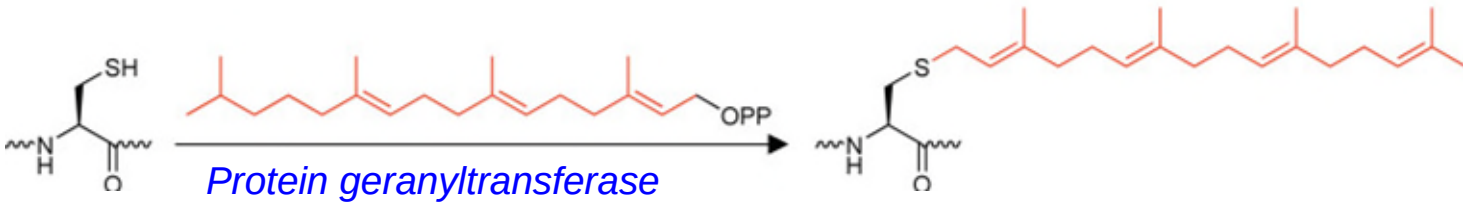
LIPID-ANCHORED PROTEINS – PRENYLATION

A prenyl anchor is a polyisoprene hydrocarbon chain

- Activated farnesyl or geranylgeranyl groups used in cholestrol metabolism (see Module 2B)
- A thioether bond is formed with internal Cys:

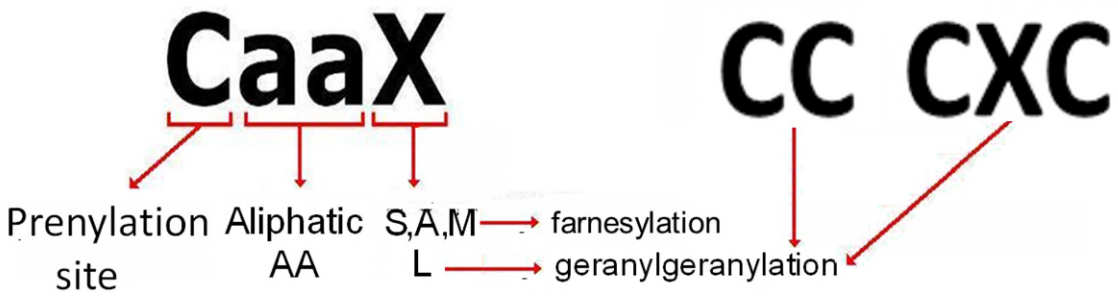


farnesyl
15C



geranylgeranyl
20C

- Enzymes require a prenylation consensus sequences at C-terminus of target proteins



Act as Molecular switch
for membrane interaction

- Examples of prenylated proteins: Ras (farn), Rab (gerger), G-protein Gγ subunit (gerger)

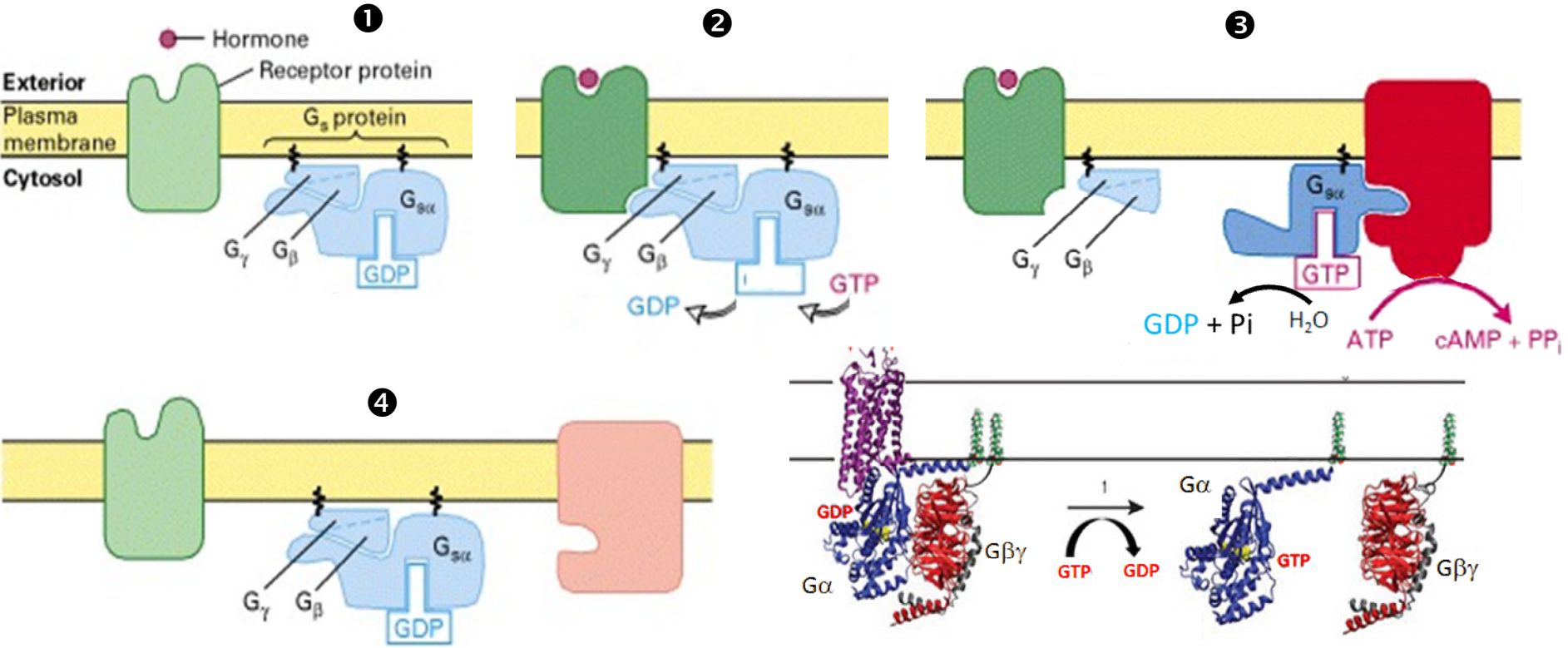
LIPID-ANCHORED PROTEINS – SIGNAL TRANSDUCTION

❶ G proteins are anchored to *cyto* side of membrane by palmitoylation (G_α) or prenylation (G_γ) and are only loosely associated with GPCR when GDP is bound

❷ Hormone binding *i)* tightens association and *ii)* induces GTP exchange

❸ This results in complete release of G_α and $G_{\beta\gamma}$ from GPRC, and both are lipid anchored to the membrane surface - they efficiently seek out cognate protein effectors by a 2D movement (e.g. *Adenylate cyclase* for G_α , *Phospholipase* for G_q ,ion channels for some G_α and $G_{\beta\gamma}$)

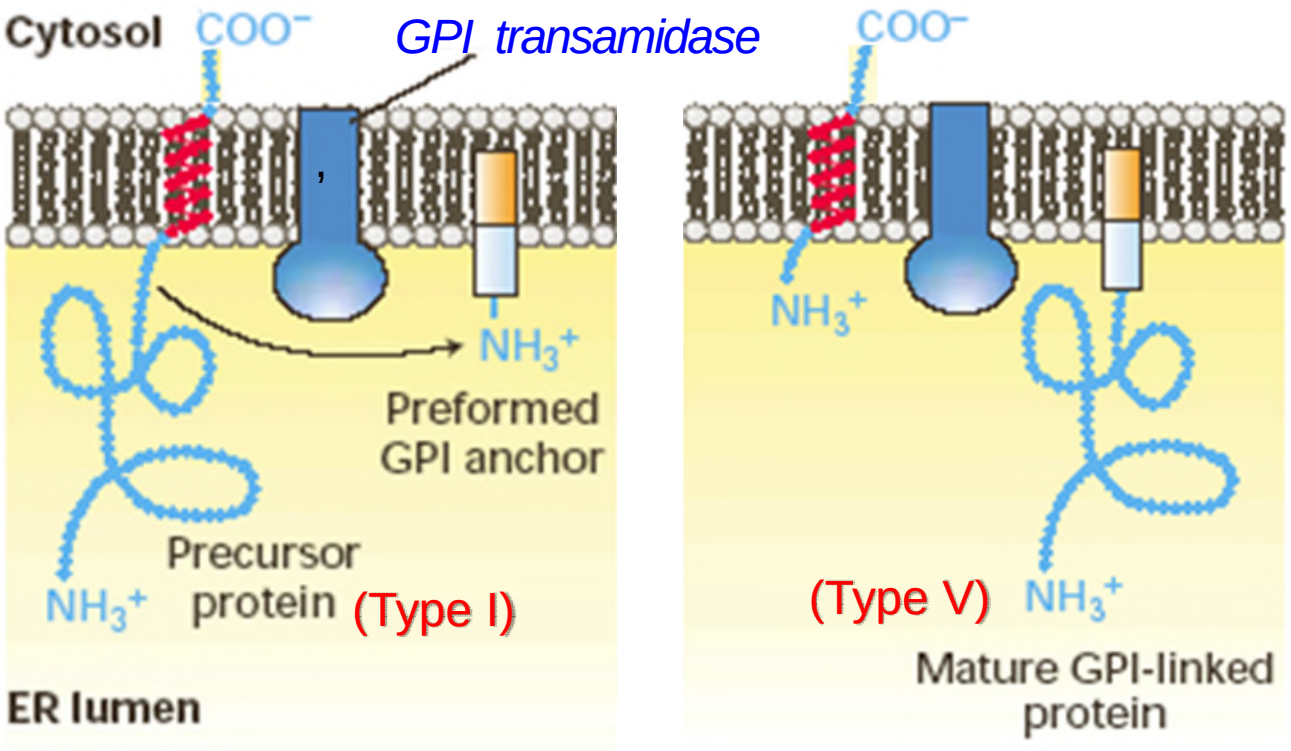
❹ G_α is a *GTPase* acting as timer for *AC / PL* activation; GDP-bound G_α releases *AC* and returns to the $G_{\beta\gamma}$ subunits, so the reassembled trimeric G-protein can return to the GPCR

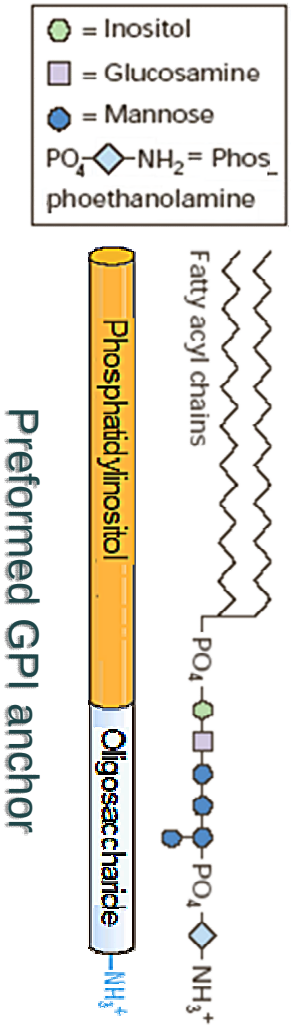


LIPID-ANCHORED PROTEINS – GPI

Glycosylphosphatidylinositol (GPI) anchors proteins to the *exo* surface

- GPI targets proteins to: *i)* the apical surface of polarized cells; *ii) caveole*; *iii) lipid rafts*
- Examples of GPI-anchored proteins are *Plasminogen activator* and *Acetylcholinesterase*
- GPI anchored proteins are formed by transferring a Type I MP from the STA TMD to the preformed GPI anchor in a reaction catalyzed by *GPI transamidase*





LIPID-ANCHORED PROTEINS – GPI PROTEIN SYNTHESIS

❶ GPI anchor synthesis begins on the *cyto* side of the ER membrane linking UDP-GlcNac to PI

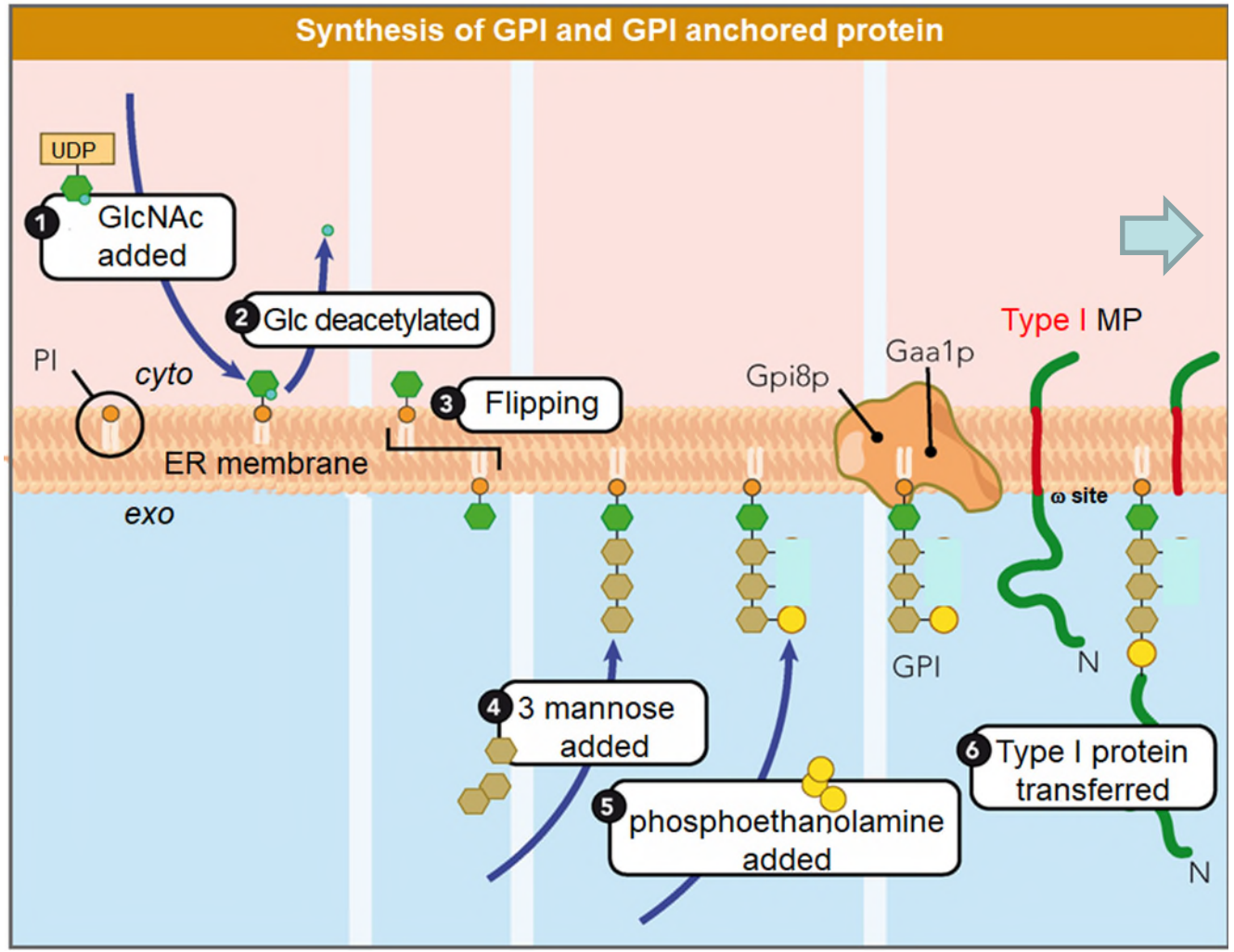
❷ The sugar is deacetylated to Glc-PI; ❸ A putative *flippase* translocates GlcPI to the ***exo*** side

❹ 3 mannose units are linked in succession to Glc (see Module 3-3).

❺ Phosphoethanolammine is added and GPI anchor is ready to receive the protein.

❻ Transfer of the protein from a Type I IMP depends on a SS N-t to the STA with an ω site where the protein is cut and the N-t exoplasmic fragment transferred to GPI.

It requires the action of *GPI transamidase*, an IMP composed of at least two subunits, Gaa1p and Gpi8p



LIPID-ANCHORED PROTEINS – GPI PROTEIN SYNTHESIS (cont.)

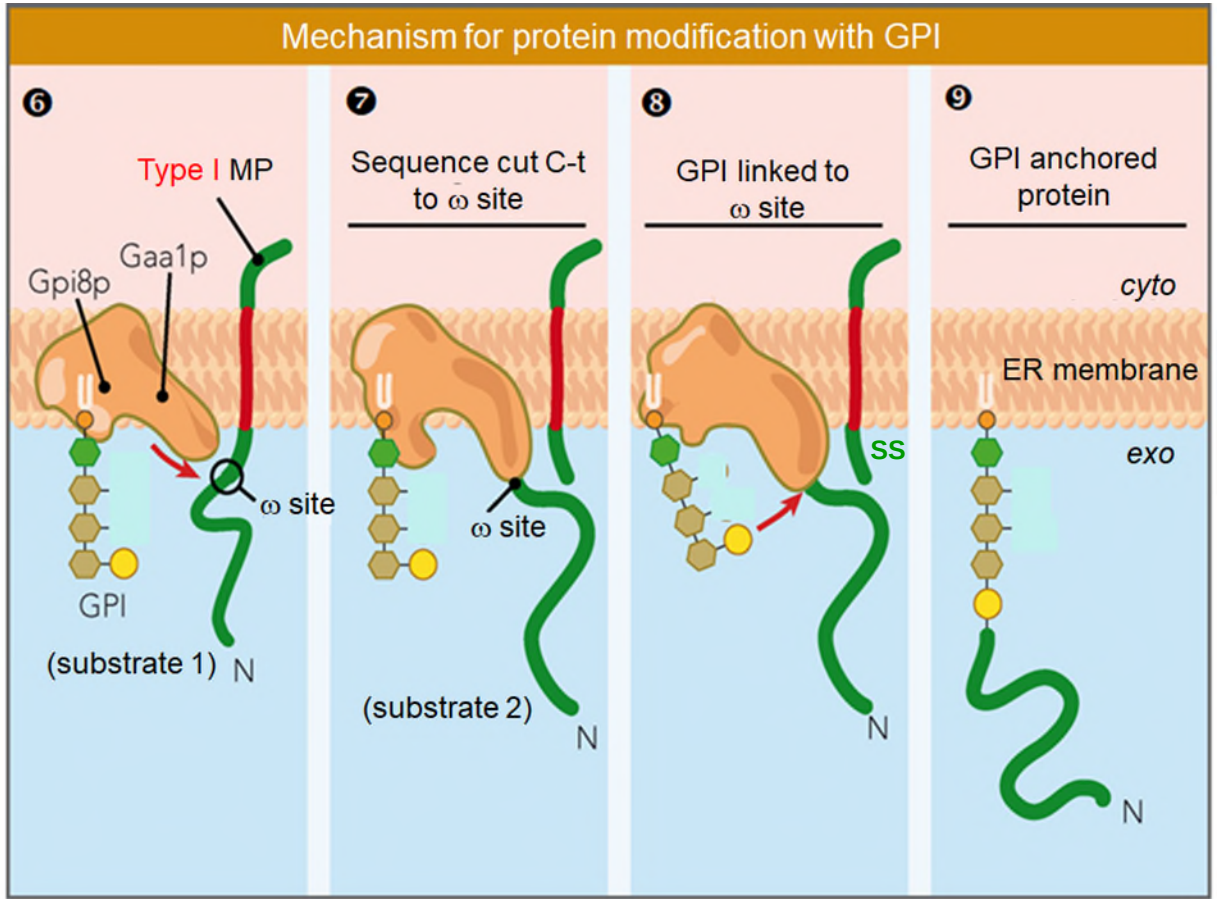
GPI transamidation from TypeI MP is believed to proceed in two steps.

❻ GPI is bound to a site on the Gpi8p subunit of the *GPI transamidase* (substrate 1)

❼ Type I protein precursor is cleaved at the ω site and the N-terminal fragment forms a covalent intermediate with the *GPI transamidase* Gaa1p subunit (substrate 2). A short AA hydrophylic sequence in the exo side of the TMD acts as signal sequence (SS not hydrophobic)

❽ The phosphoetanolammine moiety of the GPI anchor is then transferred to the ω site of the protein fragment to form the GPI-anchored protein.

❾ The GPI-anchored protein is then transferred into a secretory vesicle, which eventually fuses with the plasma membrane, and it will find itself on the extracellular side of the membrane.



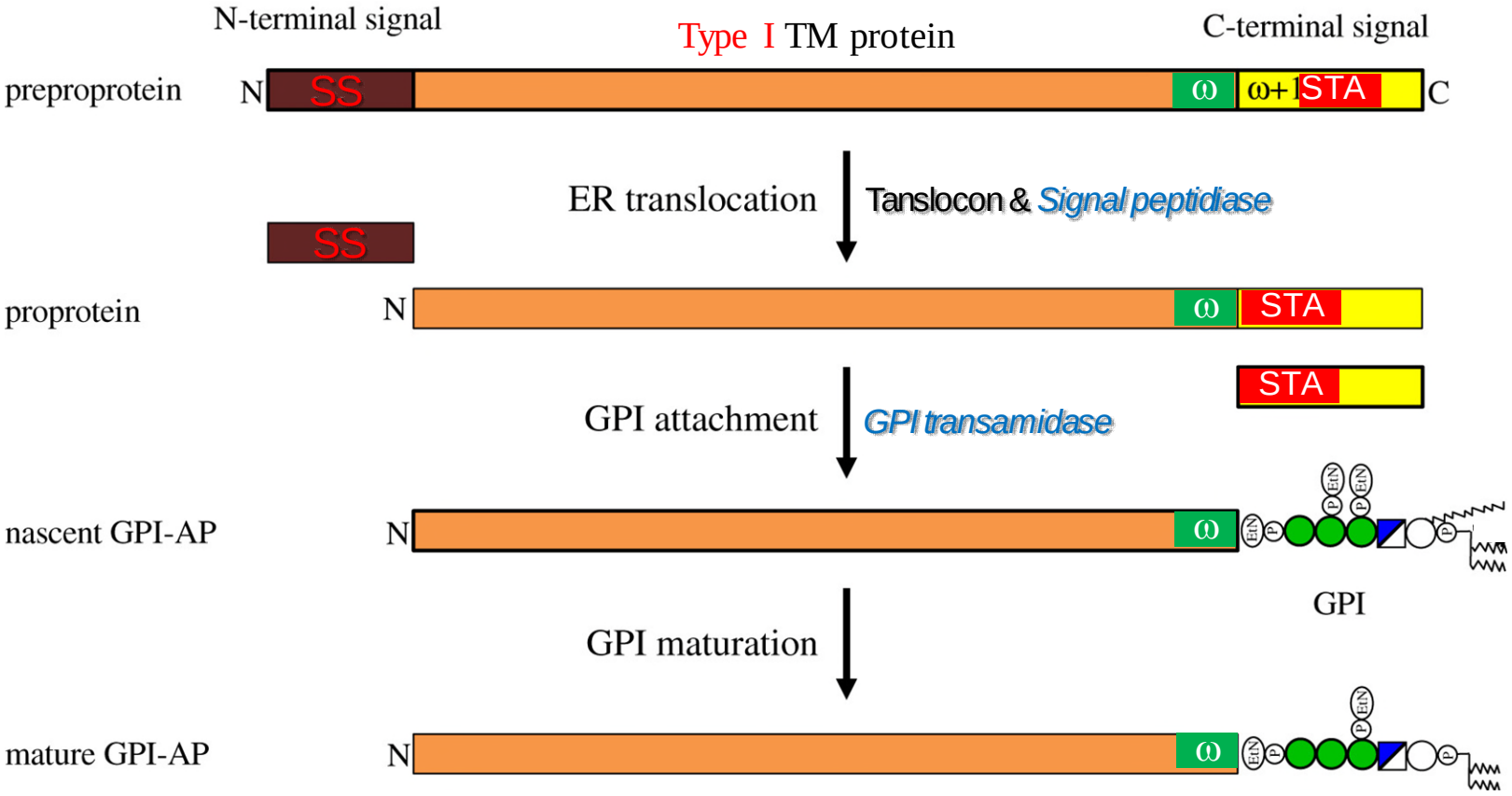
Function of GPI anchor

1) Efficient 2D mobility
2) Targets the protein to apical membranes of polarized cells

LIPID-ANCHORED PROTEINS – GPI PROTEIN SIGNAL SEQUENCE

GPI anchoring requires different types of signals:

- N-terminal SS + STA for Type I insertion
- A short hydrophylic SS on the *exo* side of STA as GPI transfer signal, which contains the ω cleavage site
- During GPI preformation, the Glc sugar can be acylated; this is then remove from nascent GPI-anchored protein to produce the mature GPI-AP



PREDICTING TM DOMAINS – AA HYDROPATHY INDICES

- **Hydropathy indices (H_i) for amino acid residues quantify the hydrophobic or hydrophilic properties of the side-chains (a.k.a. hydropathy scales)**

– They can be predicted, estimated or measured in different ways

A) Predicted/calculated H_i – based on energy or statistical calculations such as

e.g. *i)* the energy required for transfer from an organic to an aqueous environment

ii) the molecular surface area exposed to H_2O (hydrophilic environment)

B) Empirical – based on characteristics that depend on hydrophobicity

e.g. *i)* statistical assessment of the antigenicity

ii) more or less esoteric characteristics such as bitterness (inaccurate)

C) Experimentally measured – based on experimentally determined parameters that depend on hydrophobicity/hydrophilicity

e.g. *i)* octanol/water partition coefficients (requires AcNH--AA-OaAm)

ii) retention times in RP-HPLC for host guest peptides

Ac−WL−X−LL-NH_2

NB. Hydoropathy = H_2O affinity of a molecule (low or high)

Hydrophilicity : high – strong tendency to interact with water

low – low tendency to interact with water (hydrophobic)

$$P_{o/w} = \frac{[I]^o}{[I]^w}$$

HYDROPHOBICITY INDEX SCALES

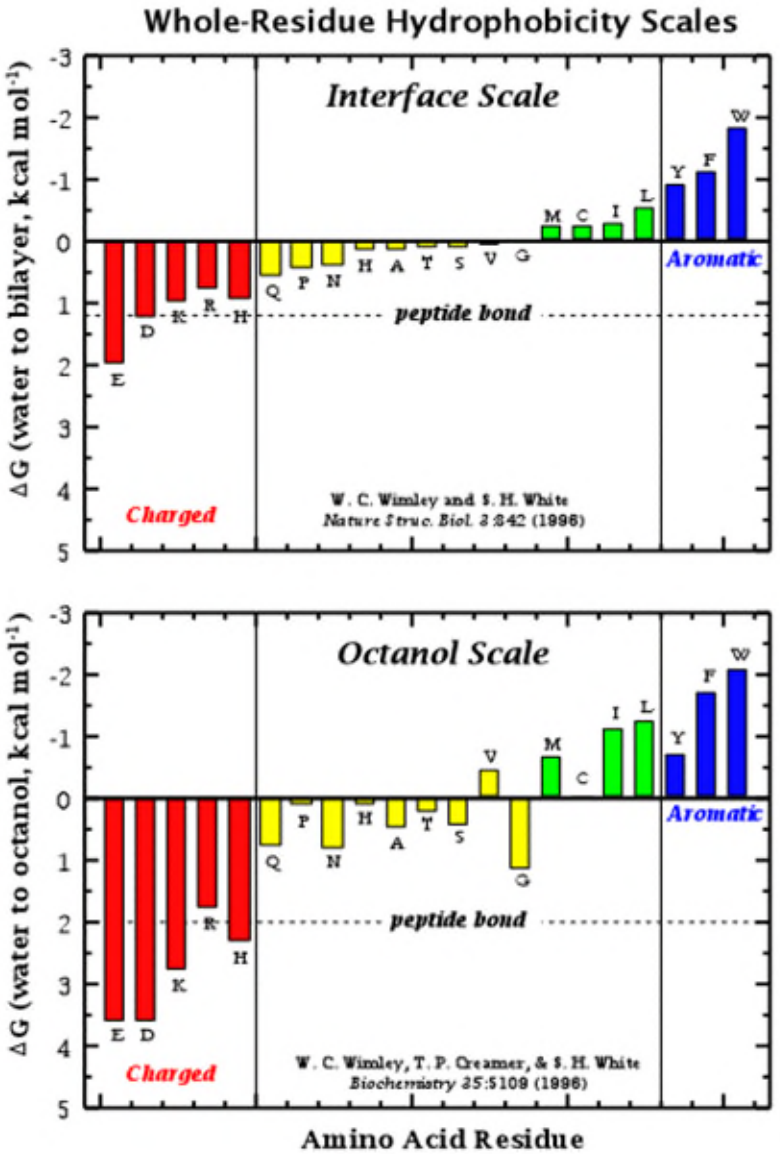
Calculated whole residue *H* scales (e,g, Wimley & White scale)

1) Transfer of unfolded chains from H_2O to the bilayer interface of LUVs
(interfacial hydrophobicity scale).

NB. Negative index = hydrophylic
Positive index = hydrophobic

2) Transfer of unfolded chains into octanol, (similar to hydrocarbon core of a lipid bilayer)

These scales include the contributions of the peptide bonds as well as the sidechains, providing absolute values, and are based on direct, experimentally determined values for transfer free energies of polypeptides



HYDROPHOBICITY INDEX SCALES (cont.)

Hydropathy scales have widely different ranges

- Hundreds of amino acid hydrophobicity scales have been developed.
- The five shown were determined using different theoretical or empirical methods (free energy of displacement from interior to surface of protein; partitioning between aqueous and membrane environment; experimental observations such as water solubility of amino acids, statistical evaluation of residues inside or on the surface of a protein
- Generally, hydrophilic resdiues have +ve indices and hydrophobic residues have – ve indices (but not always)
- The absolute values of H_i also vary onsiderably
- Difficult to campare hydropathy based on the different scales

Amino Acid	Group	Eisenberg and Weiss	Engleman et al.	Kyte and Doolittle	Hoop and Woods	Janin
Ile	Nonpolar	0.73	3.1	4.5	-1.8	0.7
Phe	Nonpolar	0.61	3.7	2.8	-2.5	0.5
Val	Nonpolar	0.54	2.6	4.2	-1.5	0.6
Leu	Nonpolar	0.53	2.8	3.8	-1.8	0.5
Trp	Nonpolar	0.37	1.9	-0.9	-3.4	0.3
Met	Nonpolar	0.26	3.4	1.9	-1.3	0.4
Ala	Nonpolar	0.25	1.6	1.8	-0.5	0.3
Gly	Nonpolar	0.16	1.0	-0.4	0.0	0.3
Cys	Unch/Polar	0.04	2.0	2.5	-1.0	0.9
Tyr	Unch/Polar	0.02	-0.7	-1.3	-2.3	-0.4
Pro	Nonpolar	-0.07	-0.2	-1.6	0.0	-0.3
Thr	Unch/Polar	-0.18	1.2	-0.7	-0.4	-0.2
Ser	Unch/Polar	-0.26	0.6	-0.8	0.3	-0.1
His	Charged	-0.40	-3.0	-3.2	-0.5	-0.1
Glu	Charged	-0.62	-8.2	-3.5	3.0	-0.7
Asn	Unch/Polar	-0.64	-4.8	-3.5	0.2	-0.5
Gln	Unch/Polar	-0.69	-4.1	-3.5	0.2	-0.7
Asp	Charged	-0.72	-9.2	-3.5	3.0	-0.6
Lys	Charged	-1.10	-8.8	-3.9	3.0	-1.8
Arg	Charged	-1.80	-12.3	-4.5	3.0	-1.4

CONSENSUS HYDROPHOBICITY INDEX SCALES

- **Normalized and averaged Hi values**

CCS – combined H_i consensus (exp./emp./calc.)

- Collect available H_i scales
- Eliminate clearly unsuitable ones
- Normalize the scales to a similar range

  e.g. Closed-ended (-10 to + 10)

  Open-ended (~ -10 to ~ +10)
- Obtain a mean value and SD for each AA
- Filter the H_i values in the original scales and remove anomalous values (e.g. outside an acceptable SD range)
- Recalculate mean value iteratively

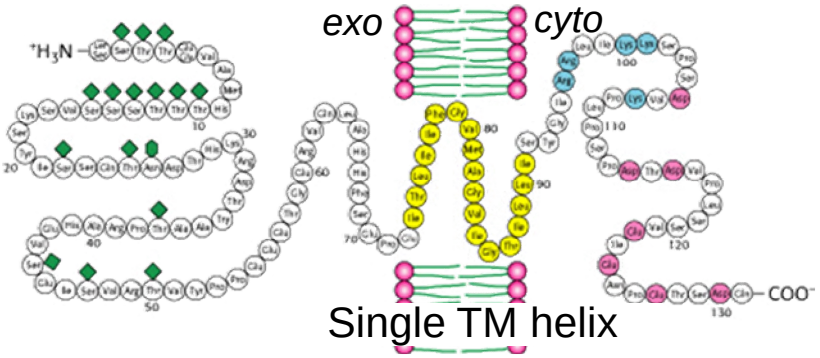
Residue	Symbol	CCS	Kyte-Doolittle	Eisenberg
		Hi	Hi	Hi
Ile	I	8.7	4.5	0.73
Leu	L	9.7	3.8	0.53
Trp	W	9.7	-0.9	0.37
Phe	F	10.0	2.8	0.61
Val	V	4.1	4.2	0.54
Met	M	4.6	1.9	0.26
Tyr	Y	2.5	-1.3	0.02
Ala	A	-1.1	1.8	0.25
Pro	P	-0.2	-1.6	-0.07
Thr	T	-3.8	-0.7	-0.18
Ser	S	-4.3	-0.8	-0.26
Cys	C	-2.3	2.5	0.04
Gly	G	-2.4	-0.4	0.16
Asn	N	-7.1	-3.5	-0.64
Asp	D	-8.3	-3.5	-0.72
Gln	Q	-6.0	-3.5	-0.69
Glu	E	-8.3	-3.5	-0.62
His	H	-3.8	-3.2	-0.40
Lys	K	-9.9	-3.9	-1.10
Arg	R	-10	-4.5	-1.80

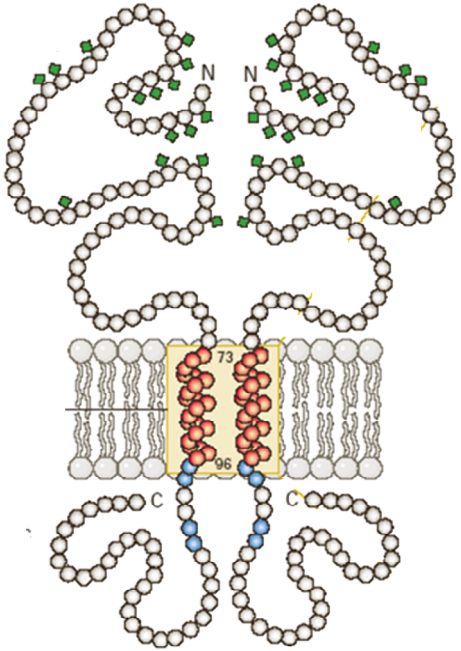
hydrophobic hydrophilic

neutral very hydrophilic

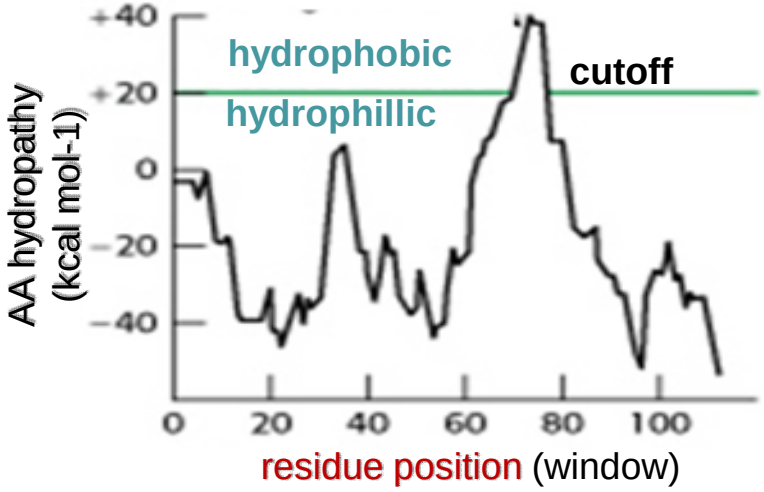
USING HYDROPATHY TO PREDICY TM DOMAINS

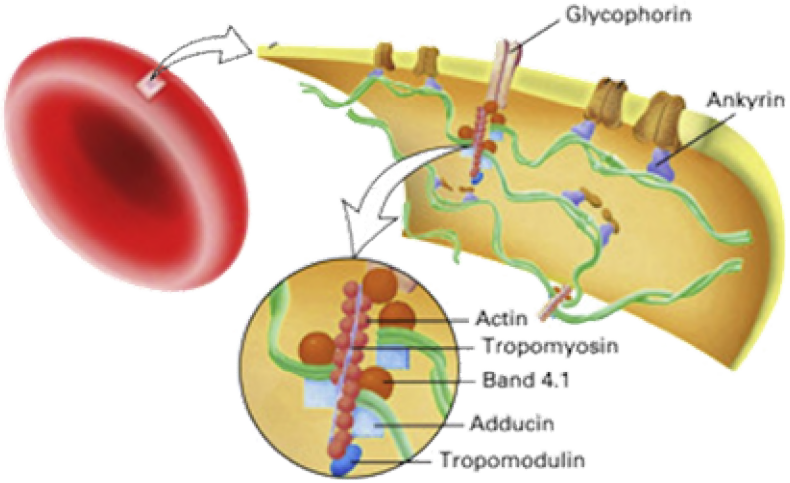
- **Predicting TMD of Type I Glycophorin (131 AA residues)**
  - Strcturl protein in RBC eith 15 O-linked oligosaccharide N-linked "





AA hydropathy = energy of transfer to aqueous environment

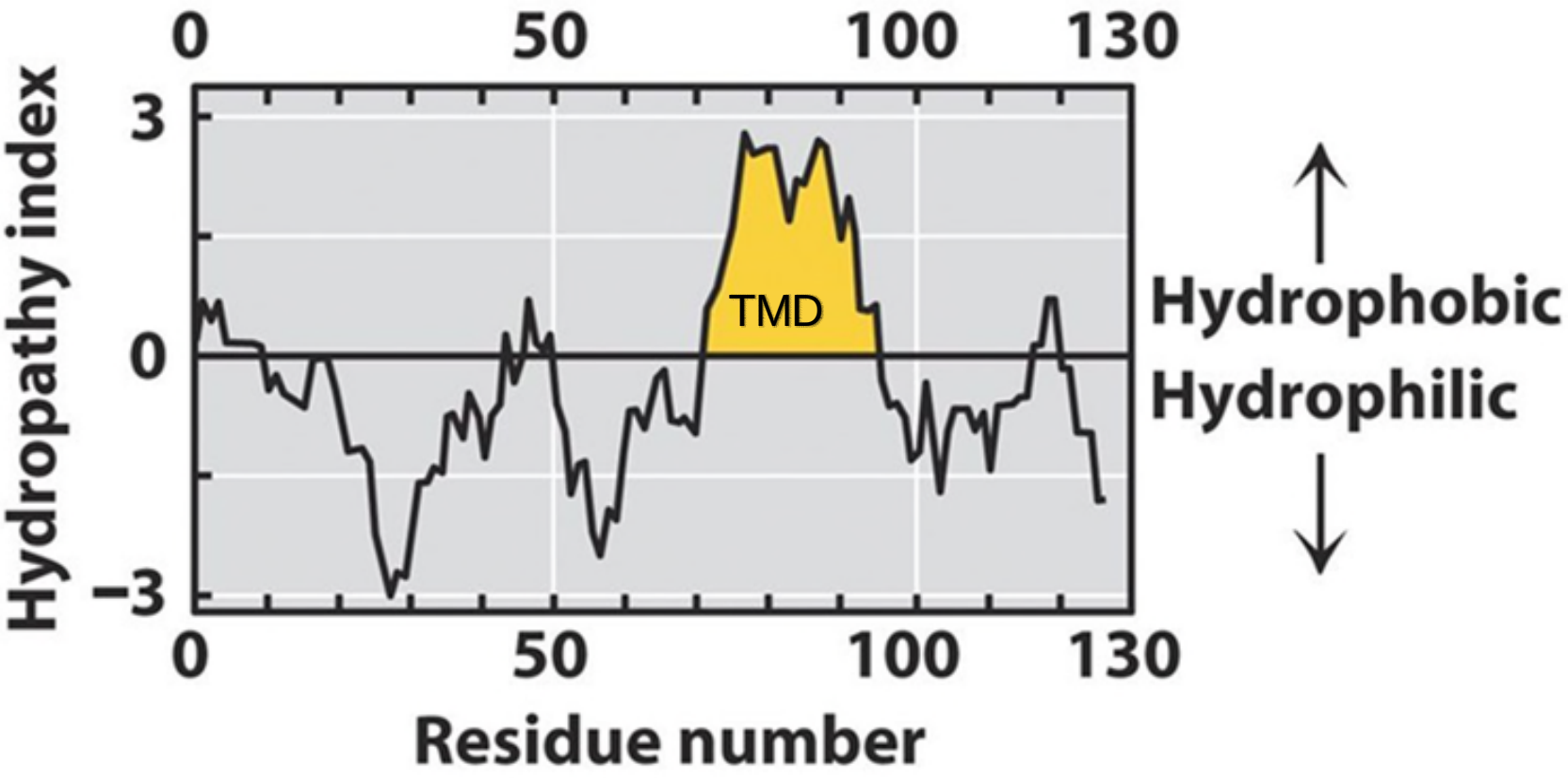




USING HYDROPATHY TO PREDICT TM DOMAINS (cont.)

Hydropathy values for a protein segment

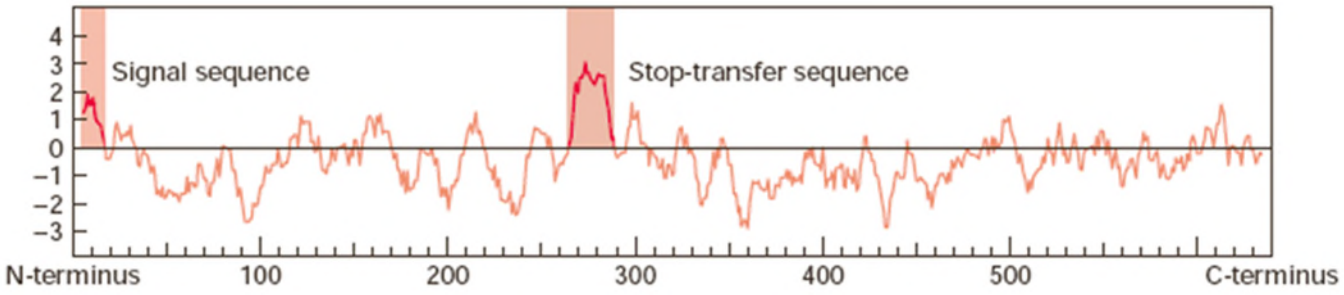
- Calculated as the mean H_i value for nAA in a sliding window (e.g. 7-20 residues)
- A region consistently showing a +ve mean H_i is hydrophobic stretch
- Hydrophobic segments likely indicate a TMD



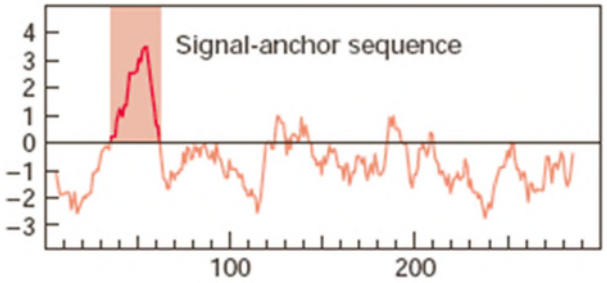
HYDROPATHY PROFILES FOR IMP

- **Examples of hydropathy profiles for integral membrane proteins**

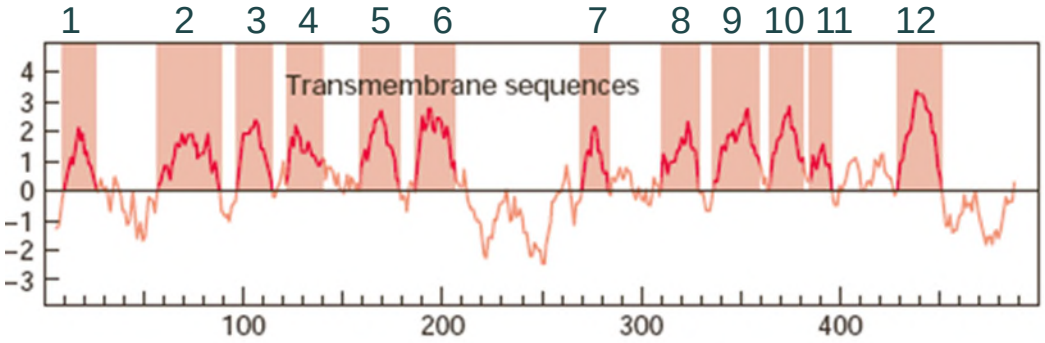
- **Type I**
  (Human growth hormone receptor)



- **Type II**
  (Glycoprotein receptor)

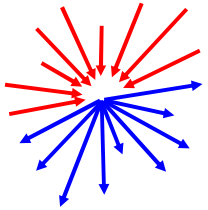


- **Type IV**
  (GLUT 1 glucose trasnorter)



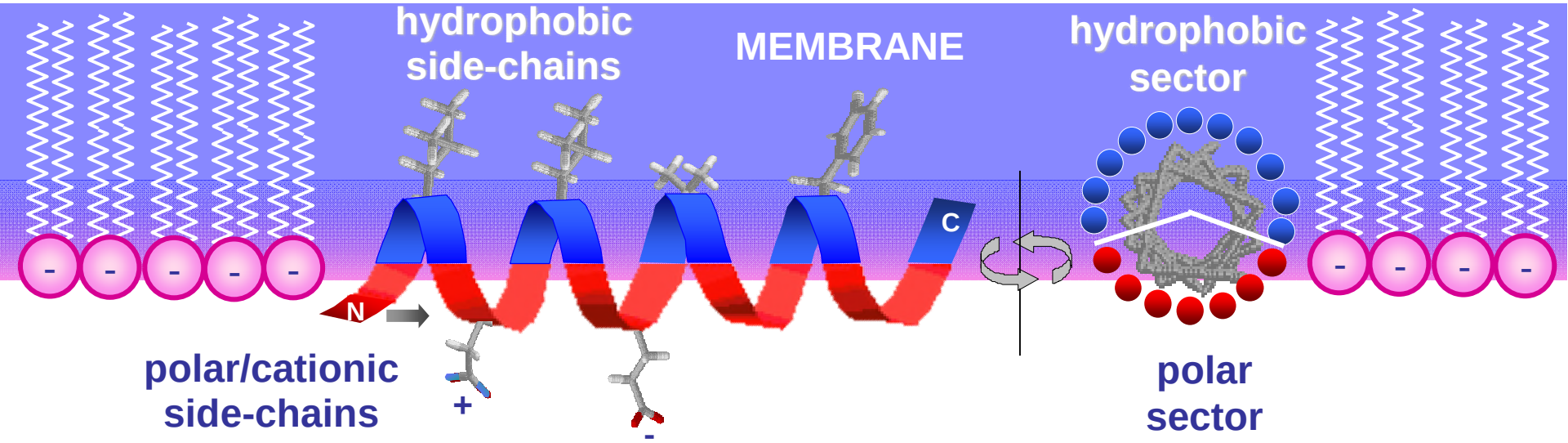
AMPHIPATHICITY

- Antimicrobial peptides (AMPs) often have amphipathic helical structures
- Hydrophobic residues are on one side of the helix and polar ones on the other
- AMPs lie on the membrane surface and insert the hydrophobic sector into the bilayer
- The more amphipathic helix → better the interaction → higher potency
- Amphipathicity can be quantified by calculating the Hydrophobic moment (μH)
- μH takes into account that AA residue side chains have a hydrophobicity and also a spatial orientation



Hydrophilic residues have –ve H_i index values

Hydrophobic residue have +ve H_i index values



ESTIMATING AMPHIPATHICITY FROM H_i SCALES

- **Hydrophobicity, hydrophobic moment (amphipathicity) and relative hydrophobic moment**

H_i

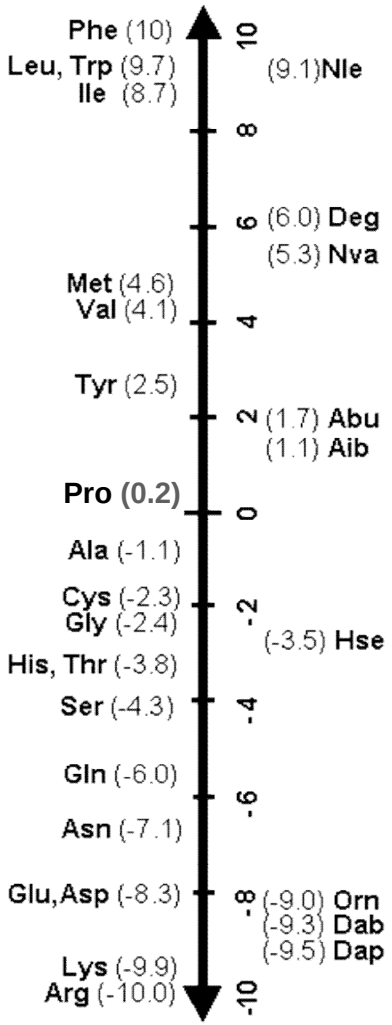
CCS Hi Scale

good correlation with

log($P_{o/w}$)

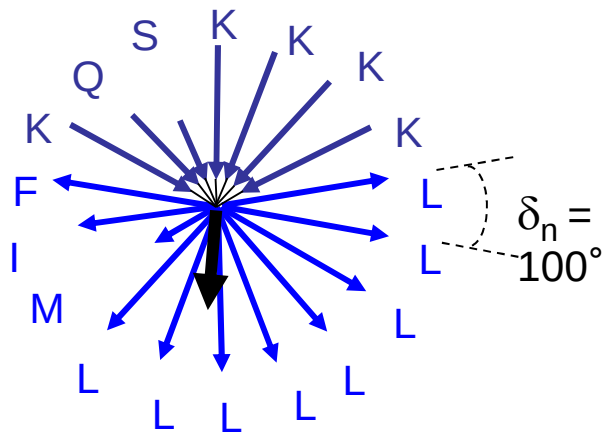
&

RP-HPLC R_t

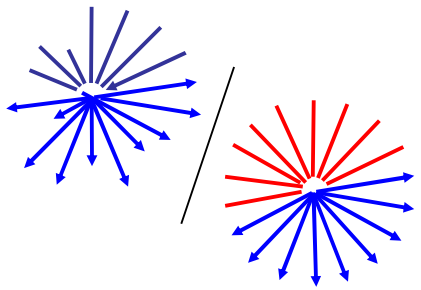


μ_H = vectorial sum of H_i

hydrophobic moment → a measure of amphipathicity



$\mu_H^{rel} = \mu_H / \mu_H^{max}$



μ_H^{max}= value for a perefectly amphipathic helix (18 residue: 9 Phe, 9 Arg)