**Insulin synthesis**

Cell types/tissues: Pancreas (beta cells)

From: <https://link.springer.com/article/10.1007/s00125-020-05192-7>

**Production of insulin starts with translation of pre-pro-insulin (ENSG00000254647, main transcript t: ENST00000381330.5 (INS-202)).**

A close up of text

AI-generated content may be incorrect.

<https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000254647;r=11:2159779-2161221>

A screenshot of a computer

AI-generated content may be incorrect.

**Pre-pro-insulin is then moved to the endoplasmic reticulum (ER) but as there is no step in between, we cannot make this happen in ER immediately: the pre-pro-insulin-1 -> pre-pro-insulin-2 [a fake reaction that we can give the mRNA expression of pre-pro-insulin] should happen in the [c] compartment. So the task will look like:**

**-> pre-pro-insulin-1[c]  
pre-pro-insulin-1[c] (ENSG00000254647) -> pre-pro-insulin-2[c]**

**However in** [**https://www.ncbi.nlm.nih.gov/books/NBK279029/**](https://www.ncbi.nlm.nih.gov/books/NBK279029/) **and** [**https://www.sciencedirect.com/topics/neuroscience/preproinsulin**](https://www.sciencedirect.com/topics/neuroscience/preproinsulin) **it is mentioned that translocation is facilitated by SRP ribo-nucleoprotein and the SRP-receptor. After learning that SRP is a ribo-nucleoprotein (meaning rRNA&protein), we lookup the protein(s) encoding for the protein parts [we skip the rRNA as RNA-sequencing often actively removes rRNA]:**

**A diagram of a cell block

AI-generated content may be incorrect.** [**https://pmc.ncbi.nlm.nih.gov/articles/PMC5224484/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC5224484/) **:A close up of text

AI-generated content may be incorrect.**

**Looking those up we get an SRP recognition step, and to quantify both the receptor and SRP steps, we make that 2 reactions, first:**

**SRP enzyme =** [**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000143742;r=1:225777699-225790670**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000143742;r=1:225777699-225790670) **(SRP9, ENSG00000143742, main transcript ENST00000304786.12 )**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000140319;r=15:40035690-40039181**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000140319;r=15:40035690-40039181) **(SRP14, ENSG00000140319, main transcript ENST00000267884.11)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000153037;r=5:112861188-112898371**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000153037;r=5:112861188-112898371) **(SRP19, ENSG00000153037, main transcript ENST00000505459.6)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000100883;r=14:34981957-35029686**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000100883;r=14:34981957-35029686) **(SRP54, ENSG00000100883, main transcript ENST00000216774.11**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000167881;r=17:76038775-76072517**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000167881;r=17:76038775-76072517) **(SPR68, ENSG00000167881, main transcript ENST00000307877.7)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000174780;r=4:56467595-56503681**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000174780;r=4:56467595-56503681) **(SPR72, ENSG00000174780, main transcript ENST00000642900.1)**

**SRP GPR rule = (ENSG00000143742 & ENSG00000140319 & ENSG00000153037 & ENSG00000100883 & ENSG00000167881 & ENSG00000174780)**

**pre-pro-insulin-2[c] (SRP GPR rule) -> pre-pro-insulin-2-SRP[c]**

**Then for the second reaction we get the SRP receptor and Sec61:**

[**https://www.science.org/doi/10.1126/science.aar7924**](https://www.science.org/doi/10.1126/science.aar7924)

**A close up of text

AI-generated content may be incorrect.**

**SR-alpha** [**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000182934;r=11:126262912-126269144**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000182934;r=11:126262912-126269144) **(SRPRA, ENSG00000182934, ENST00000332118.11)**

**A screenshot of a computer

AI-generated content may be incorrect.**

**SR-beta** [**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000144867;r=3:133784023-133825772**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000144867;r=3:133784023-133825772) **(SRPRB, ENSG00000144867, ENST00000466490.7)**

**SRP-receptor = (ENSG00000182934 & ENSG00000144867)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000058262;r=3:128051641-128071705**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000058262;r=3:128051641-128071705) **(Sec61A ENSG00000058262, main transcript: ENST00000243253.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000065665;r=10:12129637-12169961**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000065665;r=10:12129637-12169961) **(SEC61A2, ENSG00000065665, main transcript ENST00000298428.14)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000132432;r=7:54752180-54759974**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000132432;r=7:54752180-54759974) **(SEC61G, ENSG00000132432, main transcript ENST00000352861.9)**

**Sec61 = (ENSG00000058262 & ENSG00000065665 & ENSG00000132432)**

**pre-pro-insulin-2-SRP[c] (SRP-receptor & Sec61) -> pre-pro-insulin-2-SRP[r]**

**Then this is cleaved in the ER into pro-insulin, which is facilitated by**

**A close up of text

AI-generated content may be incorrect.**

**A close up of a text

AI-generated content may be incorrect.**

[**https://www.sciencedirect.com/science/article/pii/S1097276521006006**](https://www.sciencedirect.com/science/article/pii/S1097276521006006)

**This another protein complex of either 4 or 5 proteins:**

**A text on a page

AI-generated content may be incorrect.A close-up of a text

AI-generated content may be incorrect.**

****

**Some searching (this same paper) indicates that it is two distinct complexes:**

**A close up of a text

AI-generated content may be incorrect.**

**These complexes will be core: SPCS12, SPCS22/23, SPCS25   
and (SEC11A OR SEC11C)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000114902;r=3:52704955-52711148**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000114902;r=3:52704955-52711148) **(SPSC1, ENSG00000114902, main transcript ENST00000619898.5)**

**NOTE: here the main transcript isn’t the top pick (but the one with refseq, and MANE select):   
A screenshot of a computer

AI-generated content may be incorrect.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000129128;r=4:176319966-176332245**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000129128;r=4:176319966-176332245) **(SPCS3, ENSG00000129128, main transcript: ENST00000503362.2)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000118363;r=11:74949259-74979033**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000118363;r=11:74949259-74979033) **(SPCS2, ENSG00000118363, main transcript ENST00000263672.11)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000140612;r=15:84669529-84716460**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000140612;r=15:84669529-84716460) **(SEC11A, ENSG00000140612, main transcript ENST00000268220.12)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000166562;r=18:59139475-59158837**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166562;r=18:59139475-59158837) **(SEC11C, ENSG00000166562, main transcript ENST00000587834.6)**

**GPR Signal peptidase rule = (ENSG00000114902 & ENSG00000129128 & ENSG00000118363) & (ENSG00000140612 | ENSG00000166562)**

**pre-pro-insulin-2-SRP[r] (SP GPR rule) -> pro-insulin-2-SRP[r]**

[**https://link.springer.com/article/10.1007/s00125-020-05192-7**](https://link.springer.com/article/10.1007/s00125-020-05192-7)

**A diagram of insulin

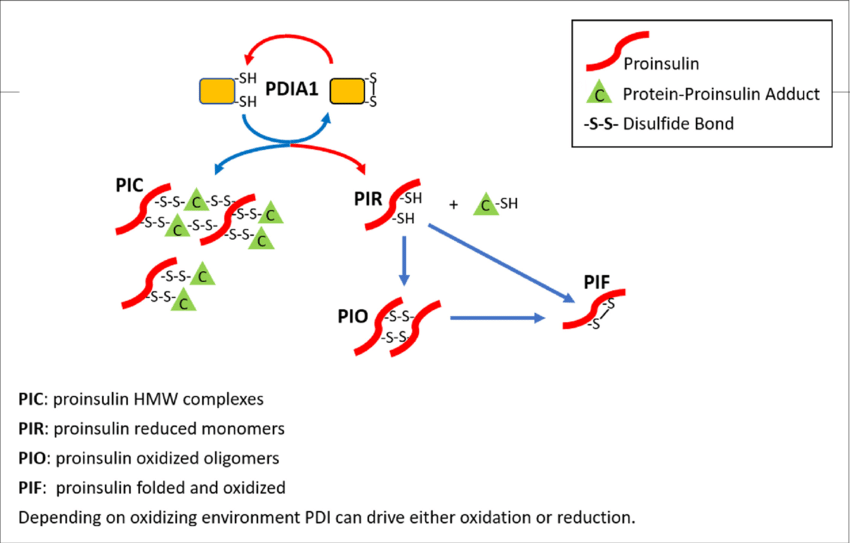
AI-generated content may be incorrect.**

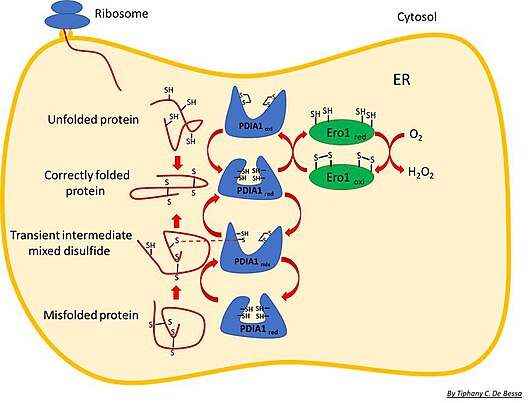
**A close-up of a text

AI-generated content may be incorrect.**

**This means that now we have pro-insulin, it needs to go through disulfide processing which is done by PDIA1 in conjunction with ERO1a/b; ERO1a/b is necessary for recuperation and involves additional metabolic reactions. In your own project I would like you to mention this in the special notes, and suggest which reaction(s) should be added to make this happen. However it will suffice to just note that both are necessary:**

[**https://www.researchgate.net/publication/333729100\_PDIA1P4HB\_is\_required\_for\_efficient\_proinsulin\_maturation\_and\_ss\_cell\_health\_in\_response\_to\_diet\_induced\_obesity/figures?lo=1&utm\_source=google&utm\_medium=organic**](https://www.researchgate.net/publication/333729100_PDIA1P4HB_is_required_for_efficient_proinsulin_maturation_and_ss_cell_health_in_response_to_diet_induced_obesity/figures?lo=1&utm_source=google&utm_medium=organic)



[**https://en.wikipedia.org/wiki/Protein\_disulfide-isomerase**](https://en.wikipedia.org/wiki/Protein_disulfide-isomerase)

**ERO1 is involved in a set of metabolic reactions which would look like this:**

**2 O2[r] + ERO1-4SH[r] (ERO1 GPR) -> 2 H2O2 + ERO1-4S[r]**

**PDIA1-4SH[r] + ERO1-4S[r] (ERO1 + PDIA1 GPR) -> PDIA1-4S[r] + ERO1-4SH[r]**

**pro-insulin-2-SRP[r] + PDIA1-4S[r] (PDIA1) -> pro-insulin-2-SRP-oxidized-folded[r] + PDIA1 4SH[r]**

**For our purposes we will only add the last with both genes involved in the gene rule, but we note it down nonetheless for future more complete projects:**

**pro-insulin-2-SRP[r] + PDIA1-4S[r] (PDIA1 & ERO1 GPR) -> pro-insulin-2-SRP-oxidized-folded[r] + PDIA1-4SH[r]**

**EDIT; on second thought, it makes more sense to include the above, as that will be used more often:**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000185624;r=17:81843159-81860856**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000185624;r=17:81843159-81860856) **(P4HB, ENSG00000185624, main transcript ENST00000331483.9)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000197930;r=14:52639915-52696025**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000197930;r=14:52639915-52696025) **(ERO1A, ENSG00000197930, main transcript ENST00000395686.8)**

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC6463291/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC6463291/)

**A close up of a text

AI-generated content may be incorrect.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000086619;r=1:236214681-236282019**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000086619;r=1:236214681-236282019) **(ERO1B, ENSG00000086619, main transcript, ENST00000354619.10)**

**PDIA1 GPR rule = ENSG00000185624**

**ERO1 GPR rule = (ENSG00000197930 OR | ENSG00000086619)**

[**https://link.springer.com/article/10.1007/s00125-020-05192-7**](https://link.springer.com/article/10.1007/s00125-020-05192-7) **A close up of a text

AI-generated content may be incorrect.  
Now pro-insulin is moved to the Golgi, the paper doesn’t mention anything else but some googling seems this appears to be a somewhat more difficult part to settle, but it seems that WFS1 is important as its deficiency leads to accumulation of pro-insulin in the ER, almost all Golgi transport is performed by COPII vesicles so those are also involved:**[**https://pmc.ncbi.nlm.nih.gov/articles/PMC8632972/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC8632972/) **A close up of a text

AI-generated content may be incorrect.**[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000109501;r=4:6269839-6303265**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000109501;r=4:6269839-6303265) **(WFS1 ENSG00000109501, main transcript ENST00000226760.5)**

**WFS1 GPR = ENSG00000109501**

[**https://www.nature.com/articles/s41467-021-22110-6**](https://www.nature.com/articles/s41467-021-22110-6)

**A white background with black text

AI-generated content may be incorrect.**

**COPII**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000079332;r=10:70147289-70170905**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000079332;r=10:70147289-70170905) **(SAR1A, ENSG00000079332, main transcript ENST00000373241.9)**

<https://pubmed.ncbi.nlm.nih.gov/30065114/>

A close up of a text

Description automatically generated

**Appear to be paralogs, so they will be in OR GPR**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000101310;r=20:18507520-18561415**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000101310;r=20:18507520-18561415) **(SEC23B, ENSG00000101310, main transcript ENST00000650089.1)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000100934;r=14:39031918-39109646**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000100934;r=14:39031918-39109646) **(SEC23A, ENSG00000100934, main transcript ENST00000307712.11)**

**SEC23 GPR rule = (ENSG00000101310 | ENSG00000100934)**

**SEC24 has 4 versions, so we need to check their distinctions**[**https://onlinelibrary.wiley.com/doi/10.1111/tra.12817**](https://onlinelibrary.wiley.com/doi/10.1111/tra.12817)

A close-up of a text

Description automatically generated

**It seems we missed one for Sar1. [As this is an example case, I keep it here to show how it came about, but in the final version you produce, this should be together with Sar1 and logically structured].**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000152700;r=5:134601149-134649271**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000152700;r=5:134601149-134649271) **(SAR1B, ENSG00000152700, main transcript: ENST00000402673.7)**

**Sar1 is also activated by Sec12, and appears to be required for it (from:** [**https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2017.00075/full**](https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2017.00075/full)**)**

A screenshot of a computer

Description automatically generated

**Similarly, from the same article, Sec16 is also apparently necessary, so we need to look into that one too:**

A close up of a screen

Description automatically generated

**First Sec12, looking that up there’s a gene PREB which is also called SEC12.**A screenshot of a computer

Description automatically generated

[**https://string-db.org/cgi/network?taskId=bsMK6E79MCps&sessionId=bVlHxYghUAmt**](https://string-db.org/cgi/network?taskId=bsMK6E79MCps&sessionId=bVlHxYghUAmt)

A screenshot of a computer

Description automatically generated

**So sec12:**[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000138073;r=2:27130752-27134671**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000138073;r=2:27130752-27134671) **(PREB, ENSG00000138073, main transcript ENST00000260643.7)**

**This picture also shows that SEC16, has SEC16A, so potentially a SEC16B version as well:**[**https://string-db.org/cgi/network?taskId=bpkF9P6PwLr5&sessionId=bVlHxYghUAmt**](https://string-db.org/cgi/network?taskId=bpkF9P6PwLr5&sessionId=bVlHxYghUAmt)

A computer screen shot of a network

Description automatically generated

**This is however not as well connected as SEC16A. Looking into it appears SEC16B is involved in insulin secretion, but not in pro-insulin transport.**[**https://pubmed.ncbi.nlm.nih.gov/40705078/**](https://pubmed.ncbi.nlm.nih.gov/40705078/)

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000148396;r=9:136440096-136483843**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000148396;r=9:136440096-136483843) **(SEC16A, ENSG00000148396, main transcript ENST00000684901.1)**

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC6613978/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC6613978/)

A screenshot of a document

Description automatically generated

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC12598545/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC12598545/)

A close up of a text

Description automatically generated

**So Sec12 is necessary for SAR1a and B, and both are indepdently capable of creating COPII coats, as single knockout didn’t lead to reduction of pro-insulin folding. Note that while the article discusses pro-insulin forming, this is due to changes in the ER as result of impaired COPII functionality and we thus include it here for the COPII step, and not the disulfide formation that we included earlier with PDIA1.** A screenshot of a computer

Description automatically generated

**So Sar1a/b/Sec12 GPR rule = ((ENSG00000138073 and (ENSG00000152700 |** **ENSG00000079332))**

**That leaves Sec16A (not B, see above):**

A screenshot of a computer

Description automatically generated

**So Surf4 is also important it appears.**

**Sec16A GPR rule will go together with Sar1 = (ENSG00000148396 & Sar1A/B/Sec12)**

[**https://www.nature.com/articles/s42003-022-03417-6**](https://www.nature.com/articles/s42003-022-03417-6)

A white text with black text

Description automatically generated

**So now Surf4 adds to that rule too.**

A screenshot of a computer

Description automatically generated

**We find two the same ones, but only is correct:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**The second one is located on a chromosome, the other on a scaffold, we want the one on a chromosome.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000148248;r=9:133361450-133376166**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000148248;r=9:133361450-133376166) **(SURF4, ENSG00000148248, main transcript: ENST00000371989.8)**

**Surf4/Sec16A/Sar1A/B/Sec12 = (ENSG00000148248 &(ENSG00000148396 & Sar1A/B/Sec12))**

**Back to Sec24 and its 4 versions. Googling for Sec24A, B, C, and D, together with pro-insulin seems to indicate that A/C, and D have something to do with proinsulin. For SEC24B we don’t see this. We put in the notes this apparent odd one out, but if there is no evidence then there is no evidence.**

[**https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0233502**](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0233502)

A close up of a text

Description automatically generated

A diagram of a structure

Description automatically generated

**So we found that Sec24D is associating with proinsulin, but also that ERp29 plays a role here. Both will go into the AND rule**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000089248;r=12:112012815-112025268**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000089248;r=12:112012815-112025268) **(ERP29, ENSG00000089248, main transcript ENST00000261735.4)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000150961;r=4:118722822-118838683**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000150961;r=4:118722822-118838683) **(SEC24D, ENSG00000150961, main transcript ENST00000280551.11)**

[**https://elifesciences.org/articles/71642**](https://elifesciences.org/articles/71642)

A close up of a text

Description automatically generated

[**https://www.nature.com/articles/s41467-021-27344-y**](https://www.nature.com/articles/s41467-021-27344-y)

A screenshot of a computer screen

Description automatically generated

**And this is important in proinsulin transport, so we can conclude that both SEC24B and C are important.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000138802;r=4:109433602-109540896**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000138802;r=4:109433602-109540896) **(SEC24B, ENSG00000138802, ENST00000265175.5)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000176986;r=10:73744366-73772173**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000176986;r=10:73744366-73772173) **(SEC24C, ENSG00000176986, ENST00000345254.9)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000113615;r=5:134648588-134727909**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000113615;r=5:134648588-134727909) **(SEC24A, ENSG00000113615, ENST00000398844.7)**

**The SEC24 rule will thus be a large OR rule = (ENSG00000138802 | ENSG00000176986 | ENSG00000113615 | ENSG00000150961)**

**And since ERP29 is an AND rule we get: ERP29 & SEC24A/B/C/D & Surf4/Sec16A/Sar1A/B/Sec12**

**In google I noticed the following reference when searching for the previous proteins’ involvement in proinsulin:**[**https://pmc.ncbi.nlm.nih.gov/articles/PMC5716288/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC5716288/)

A screenshot of a computer

Description automatically generated

**We have to be careful as this is in mice.**

A close-up of text

Description automatically generated

**While not conclusive, it seems pretty good. We can go into google scholar to see all articles that have referenced it since its publication:**

A screenshot of a computer

Description automatically generated

**However nothing new in human seems to be indicated. We put in then notes that we aren’t entirely sure, but we will include it.**

**Turns out it now is called MIA2, and we saw that one in the SEC12 (PREB) string network.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000150527;r=14:39230231-39388513**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000150527;r=14:39230231-39388513) **(MIA2, ENSG00000150527, ENST00000640607.2)**

**MIA2 GPR = ENSG00000150527**

**GPR Rule = ERP29 & SEC24A/B/C/D & Surf4/Sec16A/Sar1A/B/Sec12 & MIA2**

**Now that we have done the inner COPII coat, we have the outer coat consisting of Sec31 (which has two paralogs) and Sec13 (only one paralog it appears). Sec31A and B don’t seem to be separately implicated in proinsulin trafficking, but both are expressed in pancreatic tissue:**

[**https://www.proteinatlas.org/ENSG00000138674-SEC31A/tissue**](https://www.proteinatlas.org/ENSG00000138674-SEC31A/tissue)

[**https://www.proteinatlas.org/ENSG00000075826-SEC31B/tissue**](https://www.proteinatlas.org/ENSG00000075826-SEC31B/tissue)

**Although B at much lower RNA values (it appears that the ratio of protein to RNA for SEC31B is higher though):**

**A screenshot of a computer

AI-generated content may be incorrect.**

**As for A protein and mRNA appear to be more similar: A screenshot of a computer

AI-generated content may be incorrect.**

**Zooming in on pancreas:**

[**https://www.proteinatlas.org/ENSG00000138674-SEC31A/tissue/pancreas**](https://www.proteinatlas.org/ENSG00000138674-SEC31A/tissue/pancreas)

**A screenshot of a computer

AI-generated content may be incorrect.**

[**https://www.proteinatlas.org/ENSG00000075826-SEC31B/tissue/pancreas**](https://www.proteinatlas.org/ENSG00000075826-SEC31B/tissue/pancreas)

**A screenshot of a computer

AI-generated content may be incorrect.**

**So while we don’t have evidence that both play a role, they both show expression in pancreatic islet cells, and their string networks don’t seem different. Thus we go with this for now, but note down the lack of confidence in our notes:**

[**https://string-db.org/network/9606.ENSP00000359370**](https://string-db.org/network/9606.ENSP00000359370)

[**https://string-db.org/network/9606.ENSP00000378721**](https://string-db.org/network/9606.ENSP00000378721)

**A network of colored circles with black text

AI-generated content may be incorrect.A network of colored circles with text

AI-generated content may be incorrect.**

**Here we do see PDCD6 which might be worth having a quick look at in relation to proinsulin, however that appears to not be involved in proinsulin directly (data not shown) and in STRING itself, we only find relations to Ca2+ ALG-2 colocalising with SEC31A. ALG2 is involved in misfolded response, which is relevant for misfolded insulin, but not for us directly.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000075826;r=10:100486639-100519864**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000075826;r=10:100486639-100519864) **(SEC31B, ENSG00000075826, main transcript: ENST00000370345.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000138674;r=4:82818081-82901166**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000138674;r=4:82818081-82901166) **(SEC31A**, **ENSG00000138674,** **main transcript: ENST00000395310.7)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000157020;r=3:10293131-10321123**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000157020;r=3:10293131-10321123) **(SEC13, ENSG00000157020, main transcript ENST00000350697.8)**

**So Sec31A/B/SEC13 GPR rule = (ENSG00000138674 | ENSG00000075826) & ENSG00000157020**

[**https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2025.1554717/full**](https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2025.1554717/full) **A screenshot of a computer screen

AI-generated content may be incorrect.**

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC9659391/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC9659391/) **A screenshot of a computer screen

AI-generated content may be incorrect.**

**Last is IER3IP1 which appears to be also involved in this trafficking**[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000134049;r=18:47152834-47176381**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000134049;r=18:47152834-47176381) **(IER3IP1, ENSG00000134049, main transcript ENST00000256433.6)**

**This means the full transport becomes:**

**ERP29 & SEC24A/B/C/D & Surf4/Sec16A/Sar1A/B/Sec12 & MIA2 & Sec31A/B/SEC13 & IER3IP1**

**And WFS1 also is part of it so**

**ENSG00000089248 & [ERP29]**

**(ENSG00000138802 | ENSG00000176986 | ENSG00000113615 | ENSG00000150961) &**

**ENSG00000148248 & [Surf4]**

**ENSG00000148396) & [Sec16A]**

**ENSG00000138073 & [Sec12]**

**(ENSG00000152700 |** **ENSG00000079332) & [Sar1A/B]**

**ENSG00000150527 & [MIA2]**

**(ENSG00000138674 | ENSG00000075826) & [Sec31A/B]**

**ENSG00000157020 & [SEC13]**

**ENSG00000134049 & [IER3IP1]**

**ENSG00000109501 [WFS1]**

**Let’s call that “insulin-specific-COPII” GPR rule**

**pro-insulin-2-SRP-oxidized-folded-dimerized[r] (insulin-specific-COPII GPR) -> pro-insulin-2-oxidized-folded[g]**

**EDIT: the dimerized is added as there is another step before this, see later in document**

**Now it needs to go to the (stress) granules, and here too are a large amount of proteins involved [fortunately this, just like COPII, will be similar for many proteins].**

[**https://link.springer.com/article/10.1007/s00018-020-03688-4**](https://link.springer.com/article/10.1007/s00018-020-03688-4)

**A screenshot of a text

AI-generated content may be incorrect.  
ISG = insulin secretory granule  
TGN = trans-Golgi-network**

**A diagram of a cell

AI-generated content may be incorrect.**

**Here we also find that we missed (potentially) another protein involved in earlier folding: A close up of text

AI-generated content may be incorrect.**

[**https://pmc.ncbi.nlm.nih.gov/articles/PMC6425875/**](https://pmc.ncbi.nlm.nih.gov/articles/PMC6425875/)

**A close up of text

AI-generated content may be incorrect.**

**A close up of text

AI-generated content may be incorrect.**

**So this would indicate that after folding, but before transport to the Golgi, there’s a dimerization step that is dependent on GPR94**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000166598;r=12:103930107-103953931**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166598;r=12:103930107-103953931) **(HSP90B1, ENSG00000166598, main transcript ENST00000299767.10)**

**GPR94 GPR rule = ENSG00000166598**

**So we get**

**pro-insulin-2-SRP-oxidized-folded[r] (GPR94 GPR) -> pro-insulin-2-oxidized-folded-dimerized[r]**

**and remake the golgi transport reaction to**

**pro-insulin-2-SRP-oxidized-folded-dimerized[r] (insulin-specific-COPII GPR) -> pro-insulin-2-oxidized-folded-dimerized[g]**

**Potential proteins for transport to stress granules: chromogranin A, B, VGF, Phogrin, Secretogranin 3, Arf1, AP-1, PICK1 and ICA69.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000128564;r=7:101162509-101167227**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000128564;r=7:101162509-101167227) **(VGF, ENSG00000128564, ENST00000249330.3)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000276781;r=HSCHR14\_7\_CTG1:46239-47817;t=ENST00000631274**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000276781;r=HSCHR14_7_CTG1:46239-47817;t=ENST00000631274) **(CHGA, ENSG00000276781, main transcript ENST00000613166.3)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000089199;r=20:5911394-5925361**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000089199;r=20:5911394-5925361) **(CHGB, ENSG00000089199, ENST00000378961.9)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000155093;r=7:157539056-158587823**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000155093;r=7:157539056-158587823) **(PTPRN2 (phogrin), ENSG00000155093, ENST00000389418.9)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000104112;r=15:51681492-51721026**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000104112;r=15:51681492-51721026) **(SCG3, ENSG00000104112, main transcript ENST00000220478.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000143761;r=1:228081841-228099212**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000143761;r=1:228081841-228099212) **(ARF1, ENSG00000143761, main transcript ENST00000272102.10)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000070371;r=22:19179320-19291774**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000070371;r=22:19179320-19291774) **(CLTCL1 ENSG00000070371, main transcript ENST00000427926.6)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000141367;r=17:59619680-59696956**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000141367;r=17:59619680-59696956) **(CLTC, ENSG00000141367, main transcript ENST00000269122.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000122705;r=9:36190856-36304781**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000122705;r=9:36190856-36304781) **(CLTA, ENSG00000122705, main transcript ENST00000345519.10)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000175416;r=5:176392494-176416557**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000175416;r=5:176392494-176416557) **(CLTB, ENSG00000175416, main transcript ENST00000310418.9)**

**Clathrin GPR rule = (ENSG00000070371 & ENSG00000141367 & ENSG00000122705 & ENSG00000175416)**

**There appear to be several different AP1 proteins that build the complex, so some searches for images on google yields:**

[**https://www.mdpi.com/2077-0375/4/4/747**](https://www.mdpi.com/2077-0375/4/4/747)

**A screenshot of a computer

AI-generated content may be incorrect.**

**So we are looking for AP-1A and not AP-1B**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000072958;r=19:16197854-16245906**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000072958;r=19:16197854-16245906) **(AP1M1 ENSG00000072958, main transcript: ENST00000291439.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000166747;r=16:71729000-71809201**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166747;r=16:71729000-71809201) **(AP1G1, ENSG00000166747, main transcript ENST00000299980.9)**

[**https://www.ensembl.org/Homo\_sapiens/Transcript/Summary?db=core;g=ENSG00000100280;r=22:29327680-29388612;t=ENST00000357586**](https://www.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG00000100280;r=22:29327680-29388612;t=ENST00000357586) **(AP1B1, ENSG00000100280 main transcript ENST00000357586.7)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000106367;r=7:101154352-101161596**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000106367;r=7:101154352-101161596) **(AP1S1, ENSG00000106367, main transript ENST00000337619.11)**

**AP1 GPR rule = (ENSG00000072958 & ENSG00000166747 & ENSG00000100280 & ENSG00000106367)**

**When checking STRING we find that there’s S1 and G2 that also interact with AP1M1, however this is not mentioned in the other article. For now we will just note that this might not be complete.**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000100151;r=22:38056311-38075719**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000100151;r=22:38056311-38075719) **(PICK1, ENSG00000100151, main transcript ENST00000356976.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000003147;r=7:8113181-8262687**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000003147;r=7:8113181-8262687) **(ICA1, ENSG00000003147, main transcript ENST00000402384.8)**

**So the rule for moving to stress granules [s] will be:**

**(ENSG00000128564 & ENSG00000276781 & ENSG00000089199) & [CHGA, CHGB, VGF]**

**ENSG00000155093 & [PTPRN2]**

**ENSG00000104112 & [SGC3]**

**((ENSG00000143761 & [ARF1]**

**(ENSG00000070371 & ENSG00000141367 & ENSG00000122705 & ENSG00000175416) & [Clathrin]**

**(ENSG00000072958 & ENSG00000166747 & ENSG00000100280 & ENSG00000106367)) & [AP1\_m1 complex]**

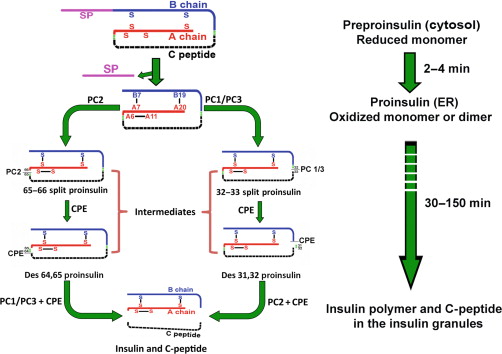
**ENSG00000100151 & [PICK1]**

**ENSG00000003147 [ICA1]**

**pro-insulin-2-oxidized-folded-dimerized[g] (ISG clathrin GPR rule) -> pro-insulin-2-oxidized-folded-dimerized[s]**

**This means that now we just need to get to the last cleavage/processing step;**

[**https://www.sciencedirect.com/topics/neuroscience/proinsulin**](https://www.sciencedirect.com/topics/neuroscience/proinsulin)



**There appear to be two separate paths to go from proinsulin to insulin. However they both use PC1/PC3, PC2, and CPE and both create C-peptide and insulin, while for the current implementation this would thus be seen as the same, it will be good to separate these already as to not have to go back in a year or 2.**

[**https://www.ncbi.nlm.nih.gov/books/NBK279029/**](https://www.ncbi.nlm.nih.gov/books/NBK279029/)

**A close up of a text

AI-generated content may be incorrect.**

**So in the PC2 part we also need 7B2, however looking further this seems to be based on chaperoning. We only include it since the book indicates that without 7B2 PC2 is also not functioning, and thus we also include proSAS)**

[**https://www.sciencedirect.com/topics/neuroscience/proinsulin**](https://www.sciencedirect.com/topics/neuroscience/proinsulin)

**A screenshot of a computer screen

AI-generated content may be incorrect.**

[**https://archive.hshsl.umaryland.edu/entities/publication/2e4db571-ce2d-4ead-a481-17f38d1ce8e8**](https://archive.hshsl.umaryland.edu/entities/publication/2e4db571-ce2d-4ead-a481-17f38d1ce8e8)

**A screenshot of a computer screen

AI-generated content may be incorrect.**

**So we don’t include proSAAS (it also inhibits activity, so not having it isn’t an issue for activity as would be the case if it was in an AND rule)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000175426;r=5:96390333-96434143**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000175426;r=5:96390333-96434143) **(PCSK1 (PC1/3)), ENSG00000175426, main transcript ENST00000311106.8)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000125851;r=20:17226107-17484578**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000125851;r=20:17226107-17484578) **(PCSK2 (PC2)), ENSG00000125851, main transcript ENST00000262545.7)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000109472;r=4:165361194-165498547**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000109472;r=4:165361194-165498547) **(CPE, ENSG00000109472, main transcript ENST00000402744.9)**

[**https://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000166922;r=15:32641385-32697098**](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166922;r=15:32641385-32697098) **(SCG5 (7B2), ENSG00000166922, ENST00000300175.9)**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC2 + 7B2) -> 65-55-pro-insulin[s]**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC1/3) -> 32-33-pro-insulin[s]**

**65-55-pro-insulin[s] (CPE) -> des-65-55-pro-insulin[s]**

**32-33-pro-insulin[s] (CPE) -> des-32-33-pro-insulin[s]**

**des-65-55-pro-insulin[s] (PC1/3 + CPE) -> insulin[s] + C-peptide[s]**

**des-32-33-pro-insulin[s] (PC2 + CPE) -> insulin[s] + C-peptide[s]**

**However now we have CPE in 2 subsequent reactions, so we can make this somewhat simpler:**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC2 + 7B2) -> 65-55-pro-insulin[s]**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC1/3) -> 32-33-pro-insulin[s]**

**65-55-pro-insulin[s] (PC1/3 + CPE) -> insulin[s] + C-peptide[s]**

**32-33-pro-insulin[s] (PC2 + CPE) -> insulin[s] + C-peptide[s]**

**And we note that this latter one is a lumped reaction. Looking at it this way I also notice that since we now have either PC1 or PC2 with CPE in one rule, this means that the paths can have different reaction activities; AND rules become the minimum of the two, so if PC1 = 60, PC2 = 15 and CPE = 50, then one path will be 15 + min(60, 50) and the other will be 60 + min(15, 50) which are slightly different. It thus already is useful to have both paths!**

**The task will thus be:**

**In:**

**-> pre-pro-insulin-1[c]**

**-> 2 O2[r]**

**Out:**

**2 H202[r] ->**

**C-peptide[s] ->**

**insulin[s] ->**

**The O2 and H202 are necessary for the disulfide bonds (PDIA1 and ERO1 part)**

**Reactions to add:**

**pre-pro-insulin-1[c] (ENSG00000254647) -> pre-pro-insulin-2[c]**

**pre-pro-insulin-2[c] (SRP GPR rule) -> pre-pro-insulin-2-SRP[c]**

**pre-pro-insulin-2-SRP [c] (SRP-receptor & Sec61) -> pre-pro-insulin-2-SRP[r]**

**2 O2[r] + ERO1-4SH[r] (ERO1 GPR) -> 2 H2O2 + ERO1-4S[r]**

**PDIA1-4SH[r] + ERO1-4S[r] (ERO1 + PDIA1 GPR) -> PDIA1-4S[r] + ERO1-4SH[r]**

**pro-insulin-2-SRP[r] + PDIA1-4S[r] (PDIA1) -> pro-insulin-2-SRP-oxidized-folded[r] + PDIA1 4SH[r]**

**pro-insulin-2-SRP-oxidized-folded[r] (GPR94 GPR) -> pro-insulin-2-oxidized-folded-dimerized[r]**

**pro-insulin-2-SRP-oxidized-folded- dimerized[r] (insulin-specific-COPII GPR) -> pro-insulin-2-oxidized-folded[g]**

**pro-insulin-2-oxidized-folded-dimerized[g] (ISG clathrin GPR rule) -> pro-insulin-2-oxidized-folded-dimerized[s]**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC2 + 7B2) -> 65-55-pro-insulin[s]**

**pro-insulin-2-oxidized-folded-dimerized[s] (PC1/3) -> 32-33-pro-insulin[s]**

**65-55-pro-insulin[s] (PC1/3 + CPE) -> insulin[s] + C-peptide[s]**

**32-33-pro-insulin[s] (PC2 + CPE) -> insulin[s] + C-peptide[s]**

**Gene rules:**

**Pre-pro-insulin = ENSG00000254647**

**SRP GPR rule = (ENSG00000143742 & ENSG00000140319 & ENSG00000153037 & ENSG00000100883 & ENSG00000167881 & ENSG00000174780)**

**PDIA1 GPR rule = ENSG00000185624**

**ERO1 GPR rule = (ENSG00000197930 OR | ENSG00000086619)**

**GPR94 GPR rule = ENSG00000166598**

**Insulin-specific-COPII = (ENSG00000089248 & [ERP29] (ENSG00000138802 | ENSG00000176986 | ENSG00000113615 | ENSG00000150961) & ENSG00000148248 & [Surf4] ENSG00000148396) & [Sec16A] ENSG00000138073 & [Sec12] (ENSG00000152700 |** **ENSG00000079332) & [Sar1A/B] ENSG00000150527 & [MIA2] (ENSG00000138674 | ENSG00000075826) & [Sec31A/B] ENSG00000157020 & [SEC13] ENSG00000134049 & [IER3IP1] ENSG00000109501 [WFS1])**

**ISG clathrin GPR rule = ((ENSG00000128564 & ENSG00000276781 & ENSG00000089199) & [CHGA, CHGB, VGF] ENSG00000155093 & [PTPRN2] ENSG00000104112 & [SGC3] ((ENSG00000143761 & [ARF1] (ENSG00000070371 & ENSG00000141367 & ENSG00000122705 & ENSG00000175416) & [Clathrin] (ENSG00000072958 & ENSG00000166747 & ENSG00000100280 & NSG00000106367)) & [AP1\_m1 complex] ENSG00000100151 & [PICK1] ENSG00000003147 [ICA1] )**

**PC1/3 GPR rule = ENSG00000175426**

**PC2 GPR rule = ENSG00000125851**

**CPE GPR rule = ENSG00000109472**

**7B2 GPR rule= ENSG00000166922**