Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Proinsulin Tokyo**

Review the biosynthesis of insulin by reading the “About Insulin” section on the Insulin activity page (<http://pdb101.rcsb.org/learn/resource/insulin-activity-page>).

*Q1. Use the following words to complete the sentences written below*:

Preproinsulin

Proinsulin

Chain A

Chain B

C-peptide

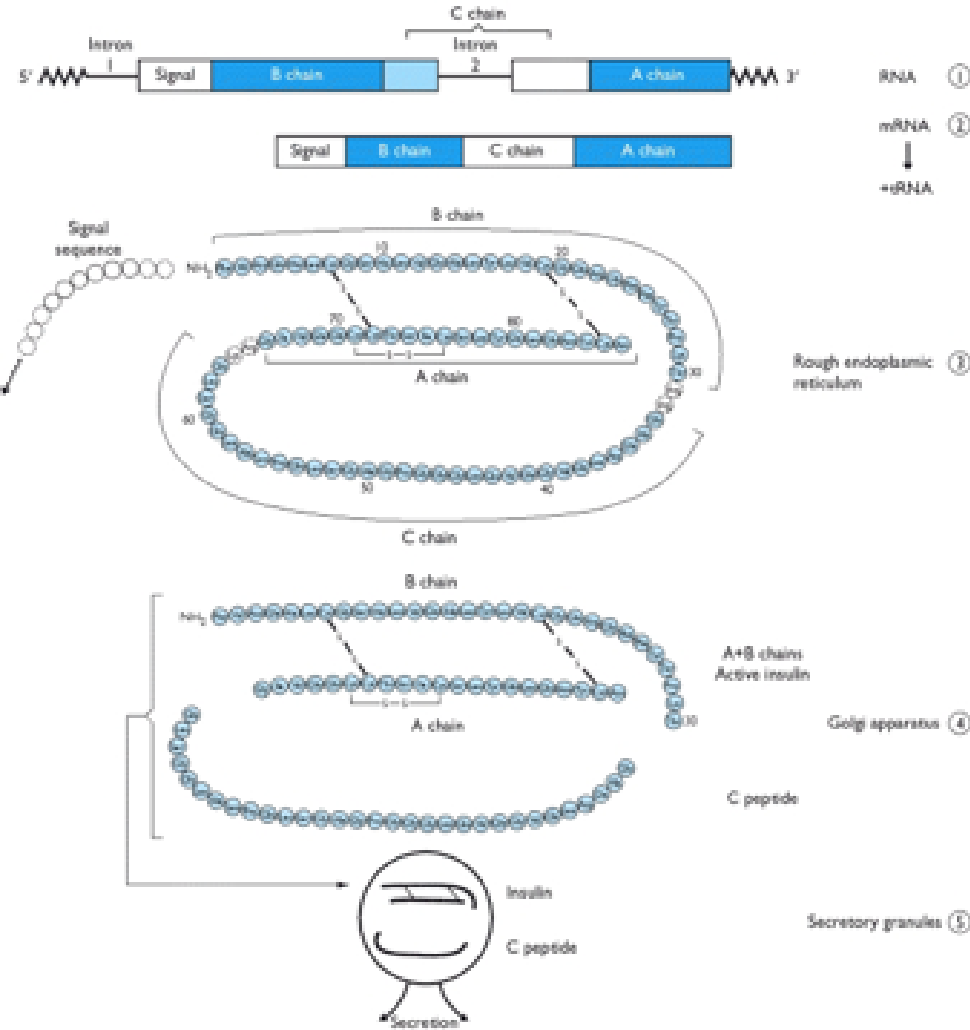
Disulfide linkages

Rough Endoplasmic Reticulum

Golgi

1. The 24 amino acid signal peptide present in \_\_\_ Preproinsulin \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ targets it to the \_\_\_\_\_ Rough Endoplasmic Reticulum \_\_\_\_\_\_\_\_.
2. After cleavage of the signal peptide the \_\_\_\_\_\_Proinsulin\_\_\_\_\_\_\_\_\_\_\_\_\_ molecule forms three \_\_\_Disulfide linkages\_\_\_\_\_\_\_\_\_.
3. In the \_\_\_\_Golgi\_\_\_\_\_\_\_, proinsulin is cleaved into three fragments by the action of the Prohormone convertases and Carboxypeptidase E. The resulting peptides are also called \_\_\_\_Chain A\_\_\_\_\_\_, \_\_\_\_\_\_\_\_Chain B\_\_\_\_\_\_\_\_, and \_\_\_\_\_\_C-peptide\_\_\_\_\_\_\_\_\_\_.

*Q2. Draw a schematic of insulin biosynthesis showing the relationship between preproinsulin, proinsulin and insulin.*



This activity will follow the following steps in order to answer a question and tell a molecular story.

1. Ask a question – this is the theme for the molecular storytelling
2. Build model based on literature review – what will you explore in the PDB
3. Investigate - Query/Browse PDB; Select PDB entries; Visualize
4. Analyze - Explore interactions; Compare Structures
5. Construct molecular explanations for original question
6. Develop argument - relate structure to bioinformatics information (go back to the literature to see if the molecular explanation makes sense)
7. Communicate - Tell a Molecular Story with Illustrations

**Overview:**

This activity focuses on a familial case of Diabetes, called Proinsulin Tokyo. It provides an opportunity to explore and understand a molecular basis for this disease and its symptoms.

**Learning Goals:**

1. Define a question/topic for exploration at a molecular level
2. Explore the literature to find out about the topic
3. Query the RCSB PDB website to find specific structure(s) for exploration.
4. Explore relevant molecular structures to develop a molecular story explaining the topic.

**The Curricular Standards met include**:

**NGSS**

Disciplinary Core Ideas

* LS1.A: Structure and Function
  + All cells contain genetic information in the form of DNA molecules. Genes are regions in the DNA that contain the instructions that code for the formation of proteins.

Crosscutting Concepts

* Patterns.
* Cause and effect: Mechanism and explanation
* Systems and system models
* Structure and function.

Science and Engineering Practices

* Asking questions (for science) and defining problems (for engineering)
* Developing and using models
* Planning and carrying out investigations
* Analyzing and interpreting data
* Constructing explanations (for science) and designing solutions (for engineering)
* Engaging in argument from evidence
* Obtaining, evaluating, and communicating information

**Common Core State Standards Connections**:

ELA/Literacy -

* RST.11-12.1 Cite specific textual evidence to support analysis of science and technical texts, attending to important distinctions the author makes and to any gaps or inconsistencies in the account.
* RST.11-12.7 Integrate and evaluate multiple sources of information presented in diverse formats and media (e.g., quantitative data, video, multimedia) in order to address a question or solve a problem. (HS-LS2-6),(HS-LS2-7),(HS-LS2-8)
* WHST.9-12.2 Write informative/explanatory texts, including the narration of historical events, scientific procedures/ experiments, or technical processes. (HS-LS2-1),(HS-LS2-2),(HS-LS2-3)
* WHST.9-12.7 Conduct short as well as more sustained research projects to answer a question (including a self-generated question) or solve a problem; narrow or broaden the inquiry when appropriate; synthesize multiple sources on the subject, demonstrating understanding of the subject under investigation. (HS-LS2-7)

**AP Biology Curriculum**

Essential Knowledge:

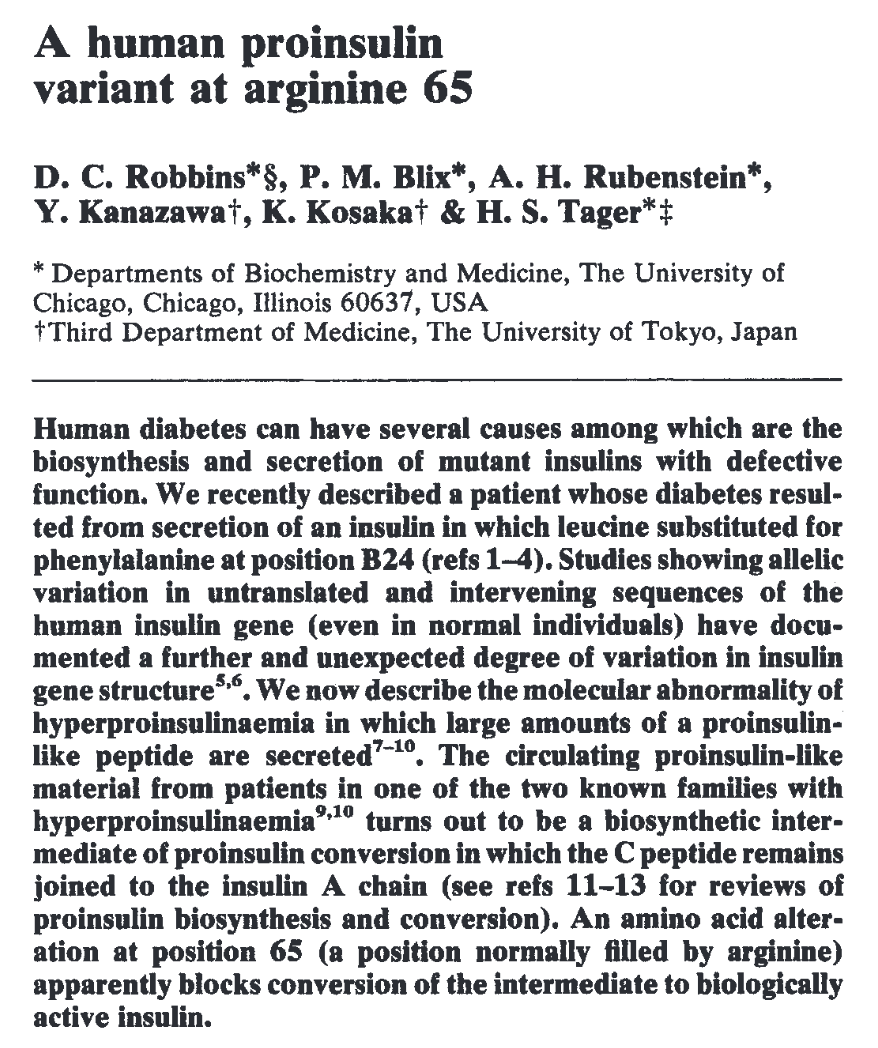
* EK 2.E.1: Timing and coordination of specific events are necessary for the normal development of an organism, and these events are regulated by a variety of mechanisms.
* EK 2.E.2: Timing and coordination of physiological events are regulated by multiple mechanisms.
* EK 3.C.1: Changes in genotype can result in changes in phenotype.
* EK 4.A.1: The subcomponents of biological molecules and their sequence determine the properties of that molecule.
* EK 4.A.2: The structure and function of subcellular components, and their interactions, provide essential cellular processes.

Science Practices:

* SP1: The student can use representations and models to communicate scientific phenomena and solve scientific problems.
* SP3: The student can engage in scientific questioning to extend thinking or to guide investigations within the context of the AP course.
* SP4: The student can plan and implement data collection strategies appropriate to a particular scientific question.
* SP5: The student can perform data analysis and evaluation of evidence.
* SP6: The student can work with scientific explanations and theories.
* SP7: The student is able to connect and relate knowledge across various scales, concepts and representations in and across domains

**Guided Molecular Storytelling**:

In 1981, an article published in the Nature magazine, reported two patient families with Diabetes resulting from defective biosynthesis and function of insulin (Robbins et al., Nature. 1981 Jun 25; 291(5817): 679-81; doi:10.1172/JCI111973.). The article abstract is included below:



1. Ask a question:

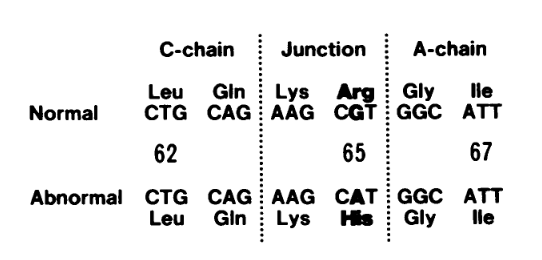
The question: Why does the mutation R65 lead to Hyperproinsulinemia? Explain this observation at a molecular level.

1. Build a model:

Search for information about proinsulin, insulin, and the mutant reported as Proinsulin Tokyo. Try to learn about the process of insulin synthesis, its key steps, and players. For your initial explorations you may search:

1. Online – using your favorite search engine (Google, Bing etc.)
2. In text books – that you or your library owns; or check the NCBI bookshelf (online) at <http://www.ncbi.nlm.nih.gov/books>
3. Review articles – you can search for these using the NCBI PubMed (online) at <http://www.ncbi.nlm.nih.gov/pubmed>

Literature research reveals that the Proinsulin Tokyo mutation in the case discussed here is changing an Arg to His. The article titled “Posttranslational Cleavage of Proinsulin Is Blocked by a Point Mutation in Familial Hyperproinsulinemia” by Shibasaki et. Al., 1985 (<http://dx.doi.org/10.1172/JCI111973>). A key result reported in this paper shows the changes at the gene and protein level in this mutation – see below:



*Q3. Based on your knowledge of insulin biosynthesis, which molecule(s) would you like to explore in the PDB? Name at least 2.*

Ans 3: The 2 molecules to explore at a structural level are proinsulin and insulin.

1. Investigate:

Search for your molecule(s) of interest in the PDB using known properties – e.g. molecule name, mutation, presence of ligand etc.

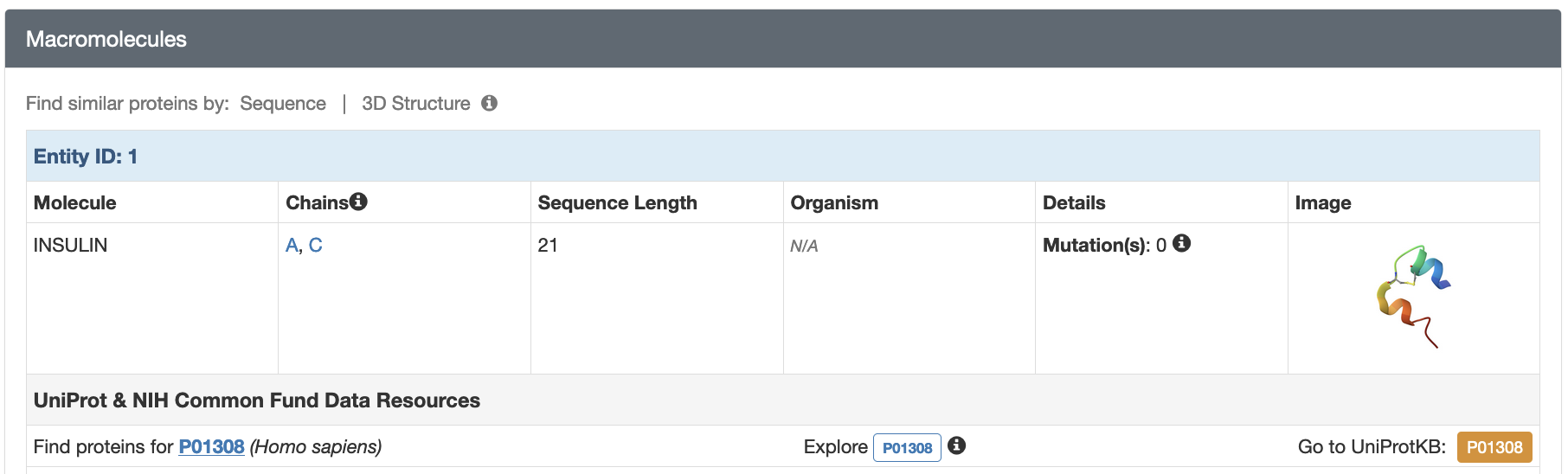
*Q4. List the PDB ID and Structure titles for at least 1 structure in the PDB for the larger of the two molecules that you have identified in Ans 3. (Hint: You may begin your search from the native human insulin PDB ID 1trz.*

Ans 4: Proinsulin (PDB ID 2kqp)

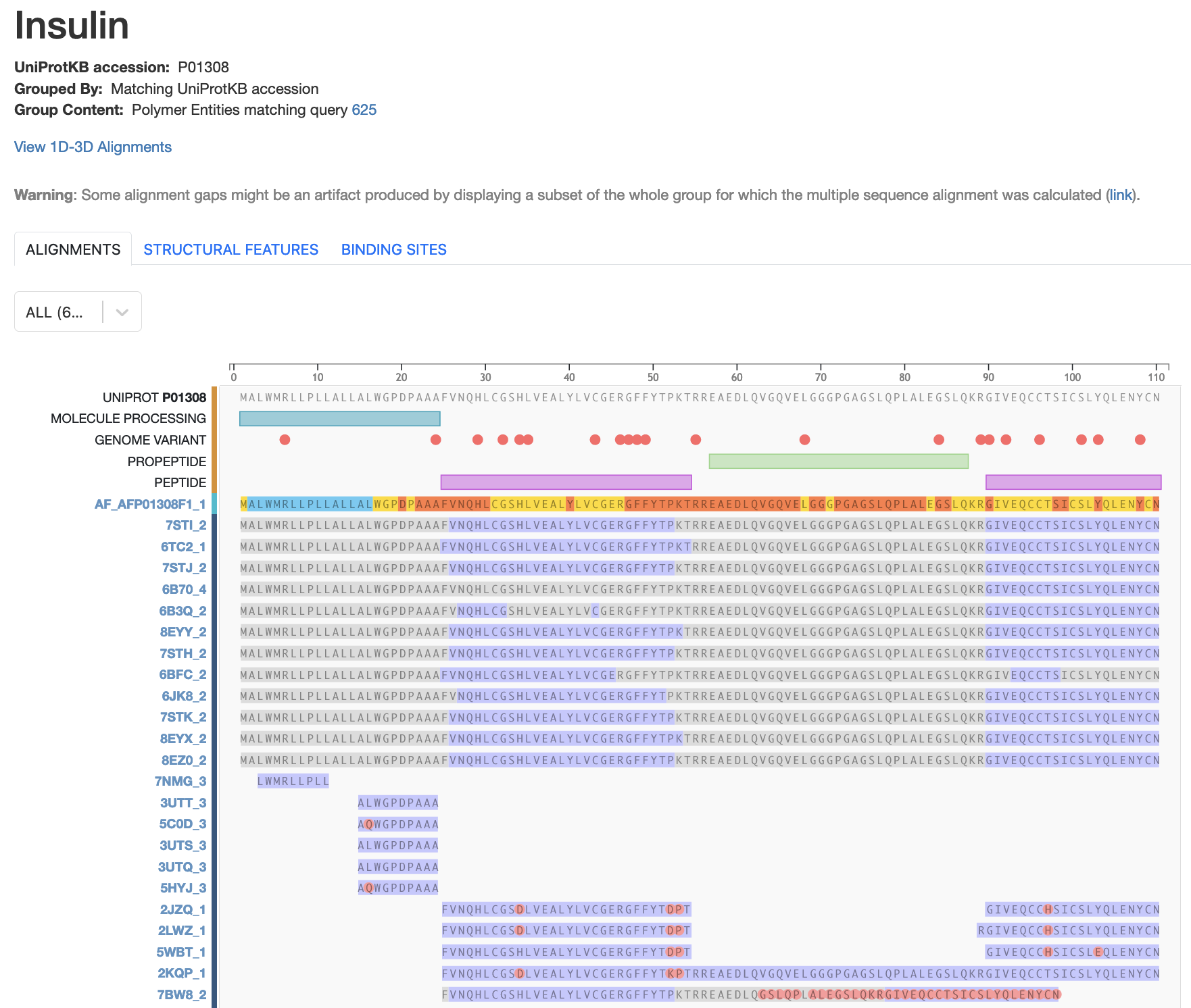
*Q5. How did you perform the search on the RCSB PDB website? List your search options and any logic that you used to refine your search results.*

Ans 5: You can try to search the archive by typing the term Proinsulin in the top search box on the home page ([www.rcsb.org](http://www.rcsb.org)). Note that depending on the autocomplete option you select, you may get different results, including a large number of incorrect matches.

Alternatively, open the structure summary page for the PDB ID 1trz. Scroll down to the Macromolecules section of the page and click on the Explore <UniProt ID> button



From the aligned list of all PDB entries that match all or part of this UniProt sequence select the structure that includes the C-peptide sequence.



Note the polymer 2kqp spans the sequence of insulin chains B, C-peptide, and A.

Visualize the PDB entry that you identified and the PDB ID 1trz. Save suitable images and include them in your answers below.

* Open the structure summary page for PDB ID 1trz.
* Click on the 3D structure tab at the top of the page to open the structure in the visualization tool Mol\*.
* To save images, click on the camera (iris) icon , Download and save a \*.jpg file.
* Repeat the steps for the PDB entry for Proinsulin that you identified above.
* Import the images in any image manipulation software of your choice (e.g., PowerPoint/ Photoshop) to add labels and additional text describing the images.

*Q6. Show the images of the insulin and proinsulin proteins. List two differences between the structures you see.*

Ans 6: The insulin and proinsulin structures are shown below.

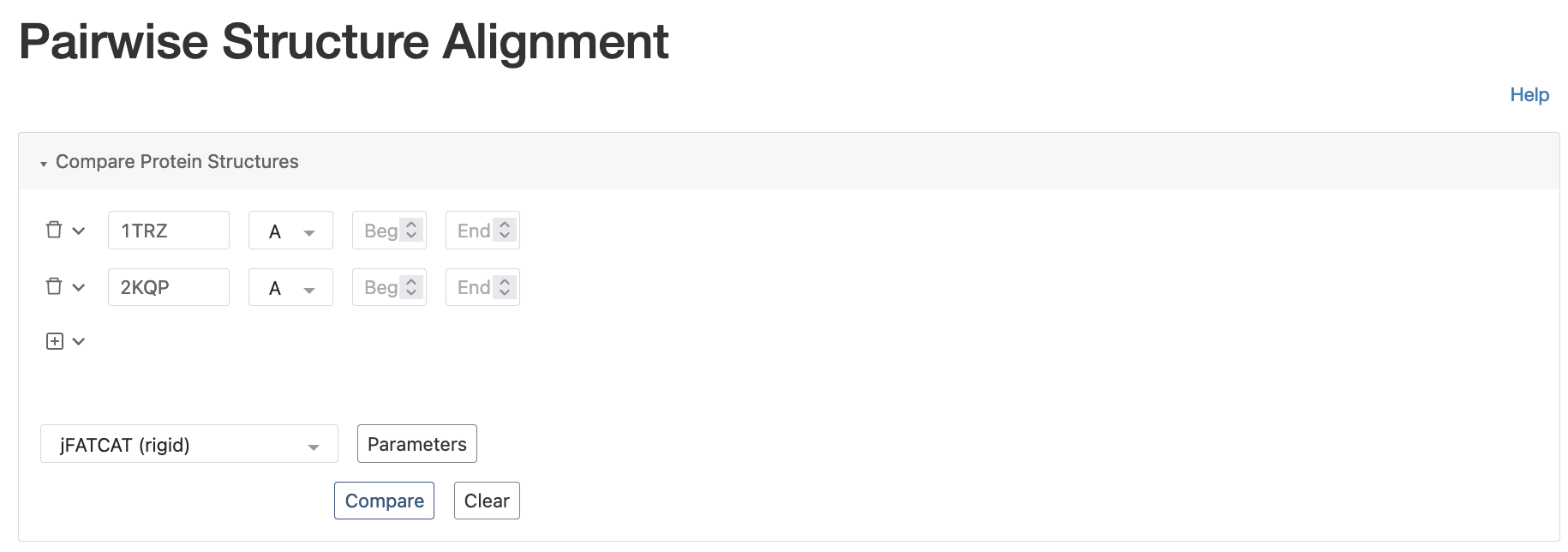
| Insulin (PDB ID 1trz) | Proinsulin (PDB ID 2kqp) |
| --- | --- |
|  |  |
| Insulin is composed of 2 polymer chains | Proinsulin has a single polymer chain |
| The polymers form a globular structure with well ordered secondary structural elements | The polymer forms a globular structure on one side of the protein while one part has a long less-ordered coiled region. |

Compare the structures of insulin and proinsulin as follows.

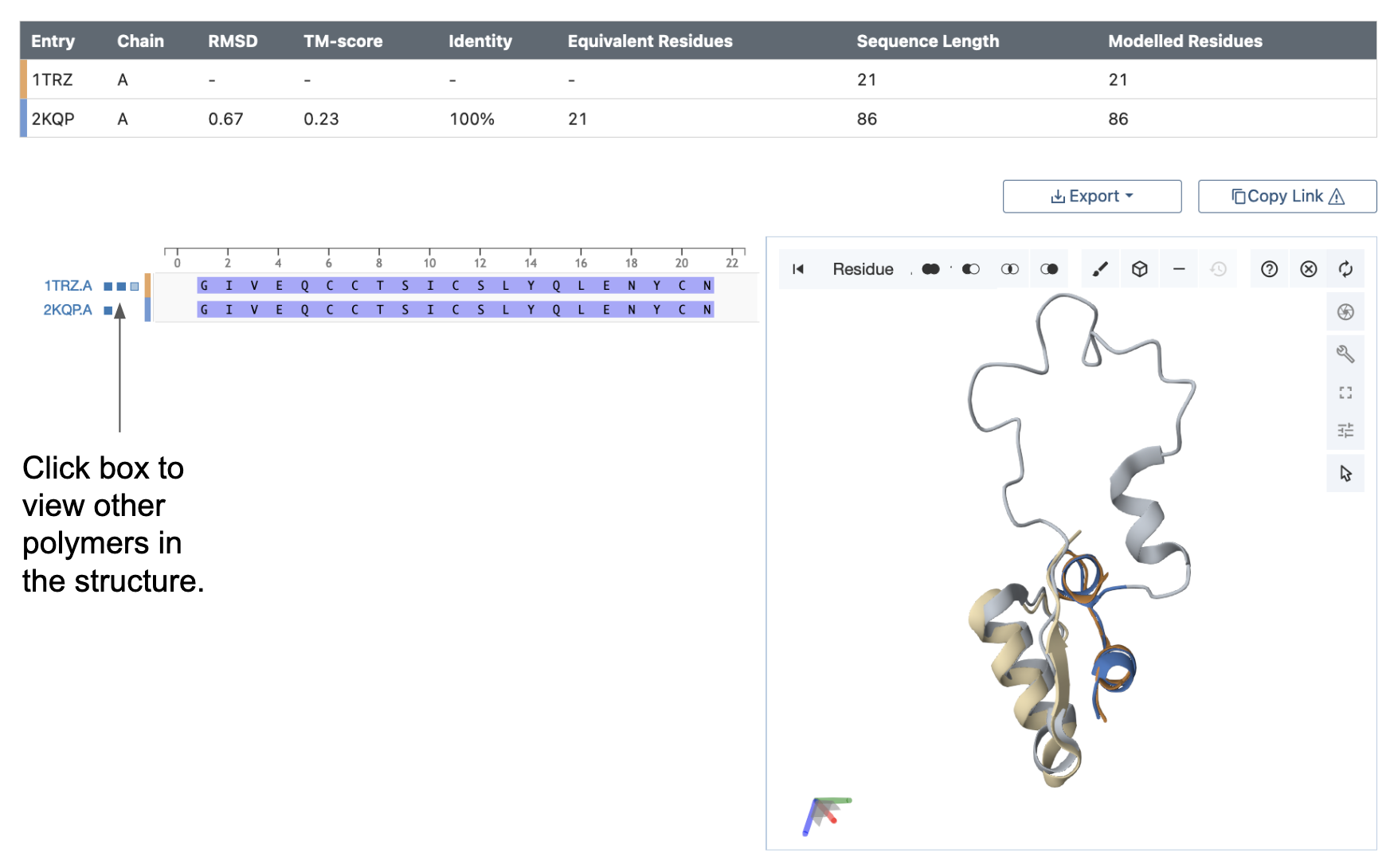
* Open the Pairwise Structure Comparison tool at <https://www.rcsb.org/alignment>
* In the boxes fill in the PDB IDs 1trz and the proinsulin PDB ID
* Select chains D in the boxes provided for the comparison.
* Use the default search options to launch the comparison by clicking on the Compare button.

Teaching Note:

The structure comparison tool should look like this before you launch the search.

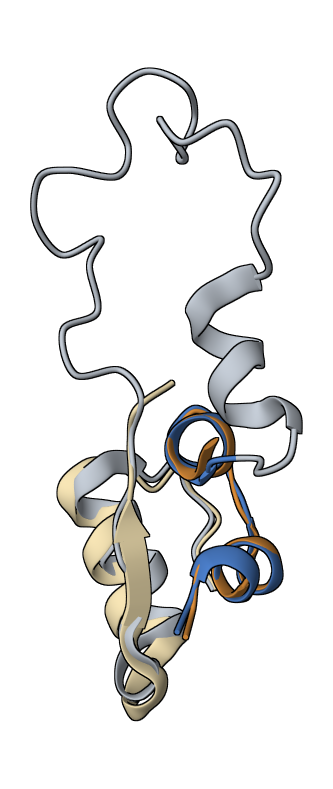


The results of the alignment look as follows.



*Q7. Which parts of the insulin and proinsulin structures match and which do not? Include a suitably labeled figure of the superposed structures here.*

Ans 7. Superposition of the insulin (PDB entry 1trz) and proinsulin (PDB entry 2kqp) structures. Notice that the loopy region in the middle of the proinsulin protein is missing from the mature insulin structure – prohormone convertase enzymes cleave it.

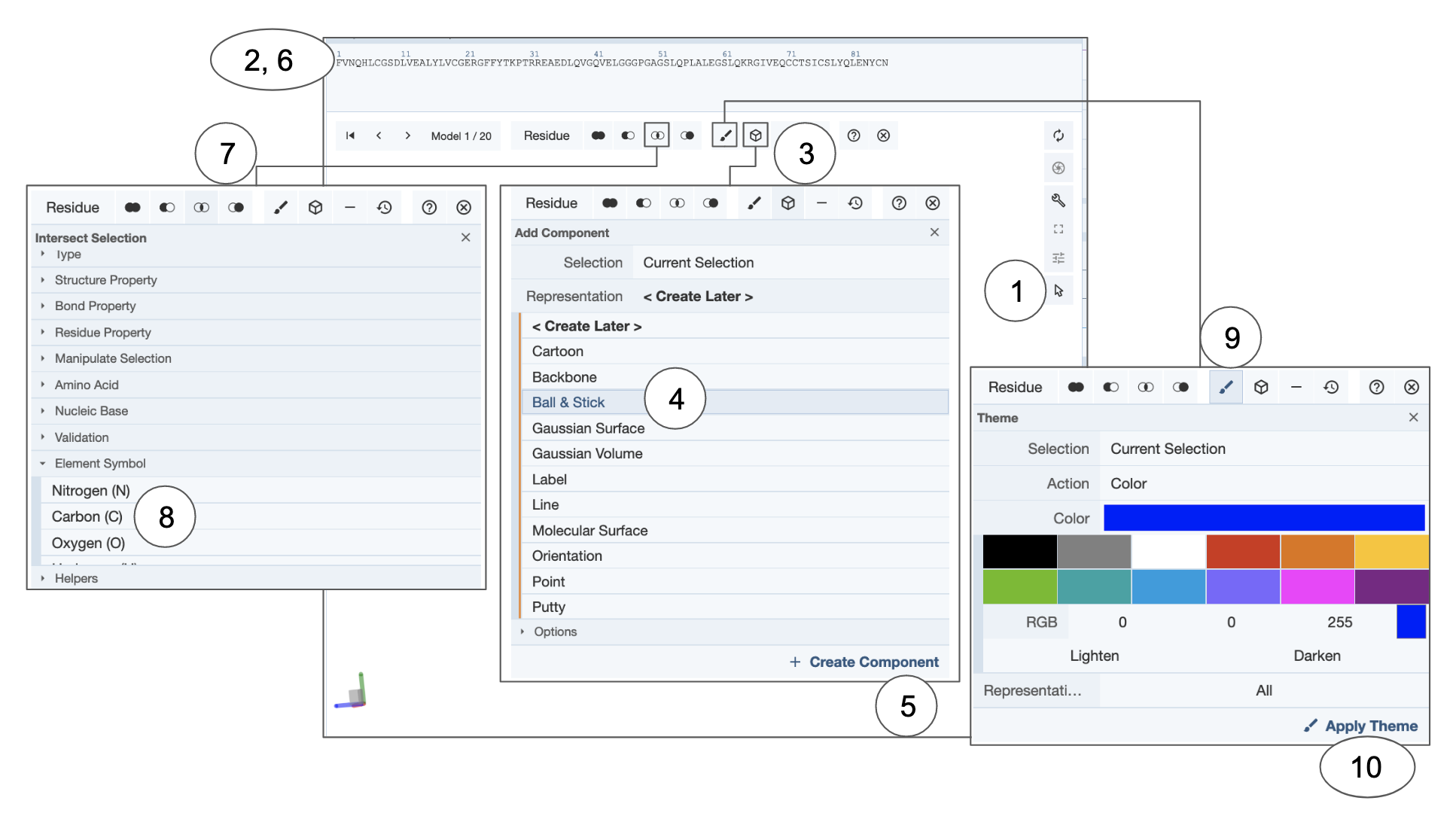


1. Analyze:

Reviewing the literature will show that the Prohormone convertase 1/3 prefers to cleave the protein chain immediately after two consecutive basic amino acids (i.e. Lys and Arg)

Proinsulin Tokyo is caused by a specific mutation of Arginine 65 being replaced by Histidine (R65H). In the proinsulin structure that you have selected, focus on the areas in the vicinity of the mutation to explore the significance of the residue Arg65.

* Open the proinsulin structure in Mol\*
* Activate the selection mode by clicking on the arrow icon in the vertical icons menu on the 3D canvas.
* Use the sequence panel at the top of the page to select consecutive RR or KR residues in the sequence.
* Once the residues are highlighted, click on the cube like components icon at the top of the page to select Representation >> Ball and Stick.
* To change the color of carbons in a specific residue - select the residue.
* Now click on the intersect selection options in the horizontal icons menu on the 3D canvas displayed when the selection model is active.
* Select Element >> Carbon (C) and then click on the paintbrush icon to select the magenta color and click on the Apply theme button.

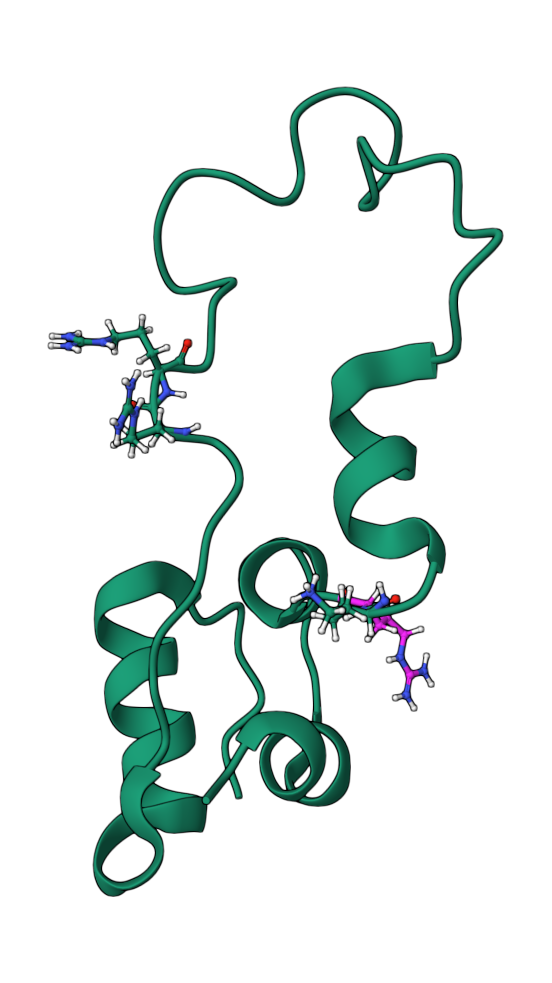


*Q8. Show the side chains of all basic amino acids (Lys and Arg) in the Proinsulin protein. What is special about the residues involved in the molecular interactions involving R65. Color this Arg (R65) so that the carbons are colored magenta. Include a suitable illustration here.*

Ans 8:

The only 2 positions where there are 2 consecutive basic amino acids are positions where the Prohormone convertase cleaves proinsulin at the chain B and C-peptide and C-peptide and chain A junctions.

The amino acid Arg65 (highlighted in magenta) is one of the residues at the C-peptide and Chain A junction. Loss of Arg from this location prevents the protease from recognizing this site for proteolysis.



Note that the 2 sets of consecutive basic residues are located at either end of the loop-like region (C-peptide). During post-translational processing of proinsulin these residues are recognized by the prohormone convertases as the sites of cleavage. The Arg65 is colored magenta in the figure above.

1. Molecular explanations:

Based on the images included above and your explorations of the molecular structures and interactions think about the R65H mutant.

*Q9. What are the main differences between the molecule(s) that you are exploring in healthy individuals compared to those in individuals with hyperproinsulinemia related to the R65H mutation?*

Ans 9: In the native protein (with Arg 65) the prohormone convertases recognize the cleavage site and completely process proinsulin to release the C-peptide. The presence of Arg 65 and 3 other Arg/Lys residues guide the processing of proinsulin to cleave the C-peptide. The mature insulin can bind to the insulin receptor efficiently and function in glucose homeostasis.

In the absence of Arg65 the prohormone convertase is unable to recognize the specific site and cleave it. This means that while the “chain B”-“C-peptide” cleavage is completed the “C-peptide”-“chain A” junction site is not cut. The disulfide linkages hold the whole protein together but this molecule is neither insulin, nor proinsulin. The proinsulin-like molecule does not bind the insulin receptor efficiently so it cannot function in glucose homeostasis. At the same time the non-functional proinsulin-like molecules accumulates in the blood causing hyperproinsulinemia

1. Argument:

*Q10. What changes at the molecular level lead to hyperproinsulinemia? Substantiate your answer with at least one additional fact or observation about this condition.*

Ans 10.

As discussed above, the presence of a mutation at Arg 65 in the proinsulin molecule prevents the prohormone convertase enzyme from specifically cleaving C-terminal to 2 consecutive basic amino acids. The incomplete processing of proinsulin in turn affects its binding to the insulin receptor and subsequent signaling. The proinsulin-like molecule collects in the blood, leading to hyperproinsulinemia.

Another resource for exploring the insulin sequence is the UniProt database. Point your browser to the website <http://www.uniprot.org/> and search for human insulin or type in the UniProt ID P01308 in the top search box (<http://www.uniprot.org/uniprot/P01308>).

Read through the page to see the various details included here about the insulin molecule. Especially examine the “Pathology and Biotech” section as well as the “PTM”/Processing or Post-translational modification sections. Note that the amino acid residue ranges for signal peptide, chains B and A, and C-peptide are included. Note that the residue numbers listed here start from 1 for preproinsulin. Thus the Arg65 in discussion above is Arg 89 (24+65) in this numbering scheme.

Another mutation at the same position Arg89Leu has the same effect as Arg89His has suggested that the presence of Arg89 or Proinsulin Arg65 is critical to its post-translational processing.

**Extension and Enrichment:**

*Q11. Based on your understanding of the molecular basis of Proinsulin Tokyo – propose a treatment strategy for individuals with the R65H mutation.*

Ans 11. Assuming that the individual is not insulin resistant, he/she could be given insulin injections for glucose homeostasis. The injected insulin would bind to the insulin receptors and function normally, unlike the proinsulin-like-material that does not bind effectively to the insulin receptors.