**Piwi Matters**

By Laurel Lorenz1 and Shuchismita Dutta2\*

1Department of Molecular Biology, Princeton University, Princeton NJ 08544

2Institute of Quantitative Biomedicine, Rutgers University, Piscataway NJ 08854

\*contact author: sdutta@rcsb.rutgers.edu

**Preparation:**

As homework and prior to the case discussion in class, get acquainted with the case.

* Watch the video Piwi Matters (<https://youtu.be/ccbMPvyqpko>).
* Review the following concepts, learn/review some techniques commonly used in this field of study, and answer questions 1-3 in preparation for the case discussion.

In order for Drosophila ovaries to continue producing eggs, they need specialized cells, called germline stem cells. Stem cells have two unique properties – (a) self-renewal – i.e., they are able to divide and (b) differentiation – i.e., they are able to become different types of cells such as cystoblasts, nurse cells, and oocytes (see Box 1).

*Box 1: Concept*

**Stem cells** possess two fundamental properties:

* ability to self-renew and
* ability to produce numerous differentiated progenies.

Stem cells that form gametes (eggs and sperms) are called **Germline Stem Cells** (or GSCs).

Sperm and egg production require a balance between self-renewal and cell differentiation. Self-renewal at the expense of differentiation can cause tumorigenesis, whereas differentiation at the expense of self-renewal can cause germ cell depletion and infertility.

**In GSCS, there are two general types of cell division**:

* asymmetric – that produce a daughter GSC and a differentiated daughter cell (e.g., in *Drosophila* ovary), and
* symmetric – that produce two daughter cells each of which has an equal probability of differentiating (e.g., in *C. elegans* and several *Hydra* species).

Q1. According to the Piwi Matters video, what happens to flies lacking the Piwi protein?

To understand the role of Piwi in *Drosophila* stem cell self-renewal*,* scientists have begun to delete or replace specific amino acid residues in the Piwi protein to produce mutant flies. By observing ovary development in these flies, they can conclude whether the deleted or changed amino acids of Piwi are required for stem cell maintenance.

When Le Thomas et al., compared the fruit fly ovaries with wild-type Piwi (panel A) to that with a specific Piwi mutant (called the YK mutant, panel B) they saw the following morphology. Observe the images included below and answer the following questions.



Q2. Based on the results shown, are the roles of the mutated residues (in the YK mutant of Piwi) required or dispensable for stem cell self-renewal? Explain your answer.

The Le Thomas et al., manuscript explains that in the YK mutations there were 2 point mutations - Y551 was mutated to L and K555 was mutated to E.

Q3. What kinds of interactions may be changed in a structure by:

a. replacing a Tyr (or Y) with Leu (or L) and

b. replacing a Lys (or K) with Glu (or E)?

This case will explore the structure and function of Piwi to understand the molecular basis of the YK mutant phenotype.

**Part 1: Getting to know Piwi (Sequence and Domains)**

Piwi appears to be a complex protein that interacts with many proteins and RNA to play an important role in gametogenesis. In this section, we will learn about the architecture and interactions of Piwi through a series of bioinformatics explorations.

To learn more about the functions and overall organization of the Piwi protein, we will explore a biological data resource called UniProt.

*Box 2: Resource*

**UniProt** (<https://www.uniprot.org/>) is a bioinformatics data resource that provides comprehensive, high-quality, freely accessible protein sequences, and their functional information. This information comes from research that has been published by others. For eukaryotic proteins, it also lists information about specific domains, post-translational processing and modifications, and pathology resulting from mutations in the protein. UniProt provides links to other biological data resources to access other relevant information about the protein, such as gene sequence, protein structures, functional annotations etc.

Search for Piwi in UniProt by typing the protein name in the top search box. From the results returned, select the entry for *Drosophila melanogaster* (Fruit fly) Piwi. Open the UniProt page and review the various types of information listed there. Refer to this page to answer the following questions:

Q1. What is the identifier of the UniProt entry that you have selected? (Hint: this is the set of alpha numerical string, listed at the top of the page, after UniProtKB -, and before the parenthesis.)

Q2. List three types of information that you found out about Piwi from this data resource. (Hint: click on the various tabs in the left hand menu. List the name of the tab and what interesting information you found about the protein).

Piwi domains: Like many other eukaryotic proteins, Piwi is composed of different domains.

*Box 3: Vocab*

A **domain** is a conserved part of a protein that can evolve, function, and exist independently of the rest of the protein chain. It usually has a stably-folded, three dimensional structure.

Explore the primary structure and domain organization of Piwi as listed in the “Family & Domains” section of the UniProt page for Piwi.

Q3. Draw a linear diagram of the Piwi protein. Mark all the positions (amino acid residue numbers) of the PAZ and Piwi domains and color them yellow and green respectively.

To learn more about the functions of the protein domains present in the Piwi protein, explore the Family and domain databases available from InterPro, Pfam, and PROSITE.

Q4. Describe in 2-3 sentences what you learned about the function(s) of each domain.

**Part 2: Exploring Piwi’s Structure and Function**

Learning about Piwi’s structure can help in understanding how it mediates its various functions.

To find an experimentally determined structure of Piwi, we will explore the Protein Data Bank (PDB) at [www.rcsb.org](http://www.rcsb.org).

*Box 4: Resource*

RCSB Protein Data Bank (**RCSB PDB**, [www.rcsb.org](http://www.rcsb.org)) provides access to 3D structural data of biological macromolecules (proteins, nucleic acids, carbohydrates and their various complexes). In addition, it provides information about the experiment used to derive the data, details about the molecules included in the experiment, and links to various bioinformatics resources that can provide additional information about the protein/molecule of interest. Each structure in the PDB is identified by a unique identifier (called PDB ID). Atomic coordinates form the PDB can be visualized and analyzed using various visualization software (some available from RCSB PDB).

Search for *Drosophila* Piwi structure(s) in the PDB ([www.rcsb.org](http://www.rcsb.org)) to see what structure(s) is/are available.

* Type the protein name “Piwi” in the top search box and hit on the search button.
* Once the results are returned, click on Drosophila melanogaster in the left hand menu to view only the structures that contain the fruit fly Piwi protein

Q1. List the PDB identifiers and titles for 2-5 structures that contain Drosophila Piwi.

The PDB structures may represent a single protein or a complex structure – i.e., the file may contain multiple polymer chains (protein, DNA, and RNA) that form a complex structure and are co-crystalized. In the list of Piwi containing structures, you should find a structure of Piwi-piRNA complex.

Select the PDB structure of Piwi-piRNA complex from the above list and open the structure summary page for this entry to explore the Piwi protein residues, peptides, and chains that are present in these structures.

*Box 5: Vocab*

**Residues**: Building blocks of biological macromolecules are sometimes referred to as residues. Depending on the context, this may refer to amino acids (frequent use) or a nucleotide (less common use).

**Protein** vs **Peptide**: while both proteins and peptides are composed of covalently linked amino acids, peptides are short (composed of 2-25 amino acids)

**Chains**: A chain (such as protein/ peptide chain) refers to covalently linked amino acids. To help locate amino acids in a protein structure or complex a chain is given an identifier (called Chain ID) and each amino acid in the chain is assigned a number. Note that large proteins with multiple domains may be cut into smaller pieces – a short peptide or a single domain and included in the experiment. For this experiment the polymer chain formed by the peptide or domain is assigned a chain ID. Nucleic acids or polymers of nucleotide residues (DNA/RNA) may also be referred to as chains and assigned chains IDs.

Q2. Fill in the following table for the PDB structure.

Ans:

|  |  |
| --- | --- |
| PDB ID |  |
| Author(s) of entry |  |
| Year when the structure was published/ released |  |
| Structure determination method |  |
| # of macromolecular chains  (# Protein + # nucleic acid) |  |
| Names of proteins in these chains (chain ID) |  |
| Name of any other macromolecular chain |  |

Click on the 3D view tab on the top of the structure summary page on the RCSB PDB website to visualize the molecule. Take a screen-shot of the graphical window and insert the image here to show the overall shape of the protein complex.

Q3. What color is the RNA chain in this structure?

*Box 6: Concept*

**Missing residues**: Due to specific conditions in the structure determination methods, most PDB entries do not have the coordinates of all atoms present in the structure. For example, flexible regions of the structure cannot be “seen” in the experiment, so these coordinates are missing. Also, all H atoms in most structures determined by X-ray crystallography are not included in the coordinate files. To learn more read <http://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/missing-coordinates-and-biological-assemblies>.

In order to visualize and analyze the protein structure(s) in greater detail we will use an online molecular structure visualization tool, called iCn3D.

*Box 7: Resource*

**iCn3D** is a web-based visualization tool that allows users to directly open any structure from the Protein Data Bank (PDB) and visualize the structure. Users can interactively rotate the molecule/complex, select specific regions and represent them in different ways, compare structures, analyze interactions and make simple distance measurements.

* Go to iCn3D (<https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>)
* Click on the button called File >> Retrieve by ID >> PDB ID so that a new window opens. Input the PDB ID of the structure you wish to visualize and click on Load.
* The structure opens in a new tab – rotate the molecule and examine the overall structure.
* Spend a few minutes playing around with the different pull-down menus to see some of the different ways this protein view can be adjusted.
* Save an image of the structure by clicking on Files >> Save Files >> iCn3D png image and upload it in the space provided below.

A close up of a flower

Description automatically generated

Figure out where the Piwi and PAZ domains are located in the 3D structure.

* Click on the button called Windows >> View Sequences & Annotations.
* Click on the Details button to see the one letter code sequence of all the protein chains in the structure.
* Use the residue number limits identified for the Piwi and Paz domains of Piwi (from your explorations in Part 1, Q3, above). Click and drag on these residues in the sequence and annotation window. The residues will be highlighted in yellow, both in the sequence and graphics windows.
* Click on Color >> Unicolor >> Green (for the PAZ domain). Now click on Select >> Clear Selection.
* Similarly color the Piwi domain Yellow.

Q4. Save a picture of the Piwi structure after the domain coloring and include it here.

Q5. In relation to the Piwi and PAZ domains, where is the piRNA located in the structure?

Examine the structure and explore the amino acids that interact with and stabilize the piRNA binding. In a fresh iCn3D session upload the PDB ID 6kr6.

* Select the RNA chain by clicking on Select >> Defined sets >> in the new window that opens >> click on 6KRK\_B (the blue chain should now be highlighted in yellow)
* Show all atoms in the RNA in stick representation by clicking on Style >> Nucleotide >> Stick
* Now select residues within 4 angstroms of the RNA by clicking on View >> H-bonds & Interactions; Turn off the Contacts/Interactions selection and then click on 3D Display interactions.
* Show the side chains of all the residues selected by clicking on Style >> Side chains >> Stick
* Focus on selected residues by clicking on View >> Zoom in Selection
* Color the selected side chains by CPK color scheme by clicking on Color >> Atom
* Now select the RNA chain and re color it blue for contrast (Select >> Defined sets >> in the new window that opens >> click on 6KRK\_B and Color >> Unicolor >> Blue

Examine the interactions through which pi-RNA is bound to the Piwi protein. Review the various types of non-covalent interactions in Box 8 and note the types of interactions between the piRNA fragment and Piwi.

*Box 8: Concept*

Biomolecular structural stability, interactions and functions are dependent on various non-covalent interactions. Some key interactions in molecular structures are

**Hydrogen bonds** - formed between two partially negatively charged atoms with a hydrogen atom between and covalently linked to one of them. e.g. in structures look for examples of O/N … H\_\_O/N, where … denotes hydrogen bond and \_\_ denotes a covalent bond

**Salt bridges** or **ionic interactions** - formed between oppositely charged amino acid side chains and/or charged ligands/ions. e.g. in structures look for interactions between Lys/Arg/His and Glu/Asp. These interactions may also involve phosphate groups and ions such as K+, Na+, Cl- etc.

**Hydrophobic interactions** - formed between hydrophobic amino acid side chains positioned away from the aqueous environment. e.g. look for regions with large numbers of carbon and hydrogen atoms in close proximity. Aliphatic amino acids such as Ala, Leu, Val, Ile participate in hydrophobic interactions.

**Pi stacking** - seen between amino acids with aromatic side chains. Pi clouds of aromatic rings interact with each other in staggered stacks, face to edge interactions. Alternatively, they may also interact with positively charged amino acid side chains (pi-cation interaction).

Learn more at <https://earth.callutheran.edu/Academic_Programs/Departments/BioDev/omm/jsmolnew/bonding/chymo.html#Topic2>

Q6. List any 3 amino acid residues with charged side chains that facilitate Piwi-piRNA binding. Support your answer with suitable images.

(Note: Mouse over any residue in the graphics window to see the residue number. Convert that NCBI reference number to the PDB/UniProt number by reading off the corresponding number from the Sequences and Annotations window).

**Part 3: Molecular Basis of the YK mutant Phenotype**

In the Piwi mutant studies the Y551L and K555E mutant could not rescue the morphological changes caused by the piwi-null mutation. Explore the Piwi-piRNA structure in iCn3D to locate the amino acid residues Y551 and K555 in Piwi. Select these amino acids and display the neighboring residues using steps described in Part 2.

Q1. Where are these residues located in the Piwi-piRNA structure?

Q2. Describe the interactions of the residues Y551 and K555, as seen in the structure you are exploring. Support your answer with suitable images.

Q3. Explain what would happen if Y551 and K555 were mutated to L and E, respectively.

**Part 4:** **Structure Based Hypothesis Development**

When piRNA binding is disrupted (as seen in the Piwi Y551 and K555 mutant) the ovaries are rudimentary in size. It is ideal to test this hypothesis in multiple ways – such as using multiple mutants. Exploring the structure of the Piwi-piRNA complex can help us identify additional amino acids that are critical for Piwi-piRNA binding.

Using the Piwi-piRNA structure, identify additional amino acids that are critical for Piwi’s function (e.g., piRNA binding) and discuss how to mutate them and test the hypothesis.

Q1. Where would you look for residues critical to Piwi’s self-renewal function in the piwi domain? Explain why.

Q2. Using the rationale described in the above answer, identify any 2 other amino acids that may be critical for Piwi’s function. List the amino acid residue number and type and explain why you think they are important. Support your answers with figures.

*Note: The residue numbers that you read by hovering your mouse over residues in the iCn3D graphics display are the NCBI numbers, you have to find the corresponding the PDB/UniProt numbers from the Sequences and Annotation window.*

Q3. What would you mutate the amino acid residues in the answer 2 to in order to test if they are functionally important? Predict what phenotype you would observe if you could make these mutants in flies. Note: Plan the mutations so that there are minimal substitutions of nucleotides (and in the codons) to introduce these mutations See Codon table attached.

