**Happy Blue Baby**

Shuchismita Dutta (contact: sdutta@rcsb.rutgers.edu)

Institute of Quantitative Biomedicine, Rutgers University, Piscataway NJ 08854

**Part 1: A Special Baby Girl**

A little baby girl, born in 2008 in Toms River NJ, showed cyanosis but was otherwise healthy and happy. She became the subject of clinical and scientific research and her case was reported in the New England Journal of Medicine in 2011.

Let’s begin learning about her story by reading a news article published in Patch, the local newspaper of Toms River (<https://patch.com/new-jersey/tomsriver/genetic-mutation-named-for-toms-river-may-shed-light-49e5fd1947>).

Q1. What symptoms did the newborn baby girl have when she was brought to the Children's Hospital of Philadelphia?

Q2. What tests did the doctors do to diagnose the baby’s condition?

Q3. Based on the test results, what was the diagnosis?

**Part 2: Grandma provides a clue**

While the doctors were testing the newborn baby, her Grandma’s comment 'My son had the same thing,' gave the doctors an important clue.

Q1. What did this comment suggest to the doctors?

Q2. Draw a pedigree chart for the Newborn baby with information provided by Grandma.

Q3. Examine a figure from the New England Journal of Medicine article (<https://www.nejm.org/doi/full/10.1056/NEJMoa1013579>, Figure 1A) showing the DNA sequence seen in the newborn baby.



Based on the DNA sequence, what is the mutation seen in the newborn?

Q4. Is the mutated residue side chain similar to or different from that found in the native protein? Explain your in terms of the size and physicochemical properties.

**Part 3: Molecular Basis of Cyanosis**

a. To explore the molecular bases of the newborn’s cyanosis, search the Protein Data Bank (at [www.rcsb.org](http://www.rcsb.org)) for one or more structures of this mutant protein.

*Box 1: 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).

You can start your search using the protein name or other details that you know. (Hint: search by the name of the mutation, or the mutation itself (e.g. X##Y, where X is the original amino acid, ## is the position of that amino acid in the protein chain, and Y is the mutated amino acid). Examine the search results and refine them as necessary.

Q1. Did you find any structures in the PDB that contain the mutation that the Toms River newborn (focus of this case) has? List the PDB ID(s).

Q2. For the PDB ID that you wish to explore open the structure explorer page for the entry by entering the PDB ID in the top search box on [www.rcsb.org](http://www.rcsb.org). You can learn many things by exploring the the page that opens (called structure summary page).

Explore the box below to learn about what you can find on this page, review the contents of the page and complete the following table with information about the entry.

*Box 2: Navigating the Structure Summary Page*

1. **Title** - that tells you what the structure is about

2. **Snapshot** - of what the structure of the molecule/complex looks like.

3. **Authors** – who solved the structure

4. **Literature** –access the article that describes the structure. This section also includes links to PubMed page and the abstract of the article describing this structure, when available.

5. **Macromolecules** – All proteins and nucleic acids present in the structure are listed here. Each unique type of macromolecule or molecular chain is listed as a separate entity. There may be multiple copies of each molecule in the structure.

6. **Small molecules** – All ligands, ions, cofactors, inhibitors that are present in the structure are listed here. You can find links here to explore the interaction of this ligand with the target protein.

7. **Experimental details** – describe details about how the structure was determined

8. **Structure quality** – shows a slider that provides insights about the quality of the structure and its agreement with the experimental data and geometric standards.

See <http://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/introduction> for details

|  |  |
| --- | --- |
| PDB ID |  |
| Author(s) of entry |  |
| Year when the structure was published/released |  |
| Structure determination method |  |
| Number of protein chains in the entry |  |
| Names and number of copies of ligands (Small Molecules) present in the structure |  |

b. Visualize the structure as follows in order to examine the mutation:

* Go to the iCn3D website at <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.

Q3. How many protein chains do you see? Take and screenshot of the structure and include it below.

Q4. What is the most common secondary structural element seen in this structure?

* To examine the mutated residue and interaction with carbon monoxide - click on the button called Windows >> View Sequences & Annotations. Now click on the Details button to see the one letter code sequence of all the protein chains in the structure. Scroll down to the bottom of the “Sequences and Annotations” window and click and drag on the carbon monoxide (shown as CMO). When you release the mouse button these ligands are highlighted in yellow.
* Click on the Style button >> Chemicals >> Sphere. Now the carbon monoxide molecules in the structure should be displayed with yellow halos around them.

Q5. Where are these ligands (CMO) located in the structure? Do all the protein chains have a CMO associated with it?

* Scroll up to see the protein sequences shown in the “Sequences and Annotations” window. Identify the mutated residue in any one of the mutated protein chains – click on it and drag the mouse to select that amino acid residue in the sequence. Simultaneously the same residue is selected in the graphics window and highlighted with a yellow halo.
* Click on the Style button >> Side chains >> Stick. Now the side chain of the mutated residue is visible. In order to make it more prominent color it in a different color by clicking on the button called Color >> Unicolor >> Magenta (or select any other color of your choice).

Q6. What secondary structural element is this mutated amino acid located on?

* Examine the neighborhood of the mutated amino acid to explore its interactions.
	+ Click on the Select button >> by Distance >> a new window opens up >> input distance 4 angstrom and select the chain ID >> click on Display. This should highlight the neighboring residues in yellow. Close the new window.
	+ Show the side chains of these amino acid residues (click on Style button >> Side chains >> Ball and Stick.
	+ Color the select amino acids and other ligands by clicking on the Color button >> Atom. This will make it easier to see the nature of atoms in the neighborhood of the mutated residue and figure out the types of interactions it participates in.
	+ Focus in on the selected residues by clicking on View >> Zoom in Selection.
* Take a screenshot of these residues and include the image below (with labels showing key amino acids and key interactions).

Q7. Are there any small molecules/ligands in the neighborhood of M67? If so, what is/are it/they?

Review the contents of Box 3 regarding intermolecular interactions.

*Box 3: Concepts*

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 (e.g. Tyr, Trp, Phe). Pi clouds of aromatic rings interact with each other in staggered stacks, face to edge interactions, or interactions with positively charged amino acid side chains (pi-action interaction).

Q8. List the names and positions of two amino acid residues located in the neighborhood of the mutated residue. What type of intermolecular interactions exist between the mutated residue and these residues? If necessary, click on the View button and use any appropriate options to view specific intramolecular interactions.

Q9. In a separate window view the structure of the native protein (PDB ID 4mqj). In the native protein, focus in on the same residues (mutated residue and its neighbors). Compare the intramolecular interactions. with the neighboring residues listed in the above answer.

Q10. Explain how the mutation in the Toms River baby girl (subject of this case) may interfere with normal function of the protein?

**Part 4: Binding and Release**

The authors engineered the V67M mutation in the hemoglobin F gamma chain and used it for structural studies. In addition, they examined the biochemical consequences of the hemoglobin Toms River mutation. They expressed the recombinant hemoglobin F (α2γ2) protein in an *Escherichia coli* expression system. Here are some things they noted:

* The mutant hemoglobin F was produced at yields similar to those for wild-type hemoglobin F.
* Initial studies indicated that the oxygenated hemoglobin tetramer (α2γV67M2) was not excessively prone to oxidation, heme loss, or denaturation, as compared with wild-type hemoglobin F.

The authors used partial laser photolysis and rapid mixing methods to measure the association (k′o2) and dissociation (ko2) rate constants for the last step of oxygen binding to individual globin subunits in wild-type and V67M γ-hemoglobin F. Data from the experiments are included in the table below:



Q1. From the data provided above what can you say about the impact of the mutation on protein folding?

Q2. What do results from the binding studies tell you about the V67M mutant’s oxygen binding and release ability, as compared to the wild type protein? Quote relevant binding constants to substantiate your answer.

Q3. Relate the oxygen binding behavior of the Toms River mutant to your structural explorations. Explain in 2-3 sentences the structural bases of the binding properties.

**Part 5: Happy Ending**

The Toms River baby diagnosed with the cyanosis causing mutation grew up to be a healthy girl. In fact, by the time the doctors had completed all her tests, she was cured.

Q1. How was the newborn girl cured? (Hint: feel free to refer to the NEJM article at <https://www.nejm.org/doi/full/10.1056/NEJMoa1013579>)

Q2. The NEJM article summary mentions a condition that may arise in the mutant proteins leading to denaturation and anemia. What is that condition? Explain your answer based on the structure that you have visualized. If possible, include a figure to support your explanation.