"Characterization of the Pathogenicity of the MSH2 P640T Mutation in Saccharomyces cerevisiae"

Ramie Fathy

Excerpt

DNA mismatch repair (MMR) is a highly conserved mechanism whose function is to maintain genetic stability by correcting mismatches made during DNA replication (for review see Kolodner 1996; Fishel & Wilson 1997; Schofield & Hsieh 2003; Kunkel & Erie 2005). In a cell lacking this repair system, several types of mutations accumulate, some of which eliminate protein functionalities completely (deleterious) and can lead to cancer in mammals and other higher organisms (Jacob & Praz 2002; Peltomaki 2003; de la Chapelle 2004). DNA MMR begins with the recognition of and binding to mismatches in the DNA helix after replication, followed by incision and degradation of the error-containing strand (Kunkel & Erie 2005). After the error is excised, a correct DNA strand is synthesized to replace it (ibid.).

The mechanism of MMR was first studied in *E. coli*, where the related MutS protein was found to operate as a homodimer to detect mismatches and short insertion/deletion loops (Su & Modrich 1986; Jiricny et al. 1988) and to subsequently recruit MutL, which initiates and coordinates subsequent repair events by complexing with MutS (Au et al. 1992; Hall & Matson 1999; Bowers et al. 2001; Acharya et al. 2003; Junop et al. 2003). Part of the subsequent repair process involves incision of the incorrect strand in order to target it for degradation, which is performed by another protein, MutH (Welsh et al. 1987; Acharya et al. 2003). MMR in *E. coli* requires all three of the MutS, MutL, and MutH proteins for correct function (for review, see Modrich & Lahue 1996).

Homologs of MutS (*M*ut*S h*omolog or MSH genes) and MutL (*M*ut*L h*omolog or MLH genes) have been identified in eukaryotes (Kramer et al. 1989; Reenan & Kolodner 1992; Fishel & Wilson 1997; Kolodner 1996; for review, see Modrich & Lahue 1996). However, though these

homologs share similar functions with their bacterial counterparts, they function with key differences. For example, identification of mismatched bases in eukaryotes is accomplished by two separate heterodimers, each with a separate role, as opposed to the prokaryotic homodimer of MutS. The first of these heterodimers, MutSα, consists of Msh2 and Msh6 (Iaccarino et al. 1996) and is involved in the recognition of single base-pair mismatches and small insertion/deletion loops (Acharya et al. 1996; Marsischky et al. 1996). The second heterodimer involved in mismatch recognition in eukaryotes is MutSβ, which is composed of Msh2 and Msh3 (Habraken et al. 1996; Palombo et al. 1996), functions to recognize larger insertion/deletion loops up to 16 base pairs long (Acharya et al. 1996; Marsischky et al. 1996; McCulloch 2003). Both of these heterodimers proceed to localize to the nucleus to participate in DNA MMR (Hayes et al. 2009), accounting for a significant amount of the localization of MSH2 to the nucleus (ibid.).

In the nucleus, MSH2 dimers interact with a number of other binding partners. One of these binding partners is MutLα, which performs the equivalent function of MutL in *E. coli* and consists of Mlh1 and Pms2 (Pms1 in yeast) (Prolla et al. 1994; Pang et al. 1997). In eukaryotes, MutLα appears to perform the endonuclease function of MutH, which can be initiated through interaction with MutSα or MutSβ (Larrea et al. 2010). Another binding partner that interacts with MSH2 dimers in the nucleus is the exonuclease Exo1 (Tran et al. 2004; Larrea et al. 2010; Fukui 2010), whose activity has been found to be dependent on interaction with the MutSα and MutSβ complexes (Genschel & Modrich 2003; Modrich 2006). Because cleavage of the error-containing strand can occur on either side of the mismatch, the excision process requires some way of being directed toward the mismatch. In eukaryotes, this directing process is performed by Pol30/PCNA, whose function also requires interaction with MutSα, MutSβ, or MutLα (Johnson et al. 1996; Umar et al. 1996).

Hereditary Nonpolyposis Colorectal Cancer (HNPCC) is an autosomal dominant hereditary disease that results from germline mutations in MMR genes, particularly in hMSH2

and hMLH1 (Peltomaki & Vasen 2004; Fishel et al. 1993; Leach et al. 1993). Similarly, missense mutations in the yeast MSH2 gene have previously been found to lead to increased mutation rates arising from faulty MMR (Gammie et al. 2007. Previous studies have identified a number of different biochemical bases for such pathogenicity of missense mutations (Gammie et al. 2007; Lützen et al. 2008). For example, it has been found that single amino acid missense mutations can lead to decreased levels of MSH2 in the cell, leading to increased mutation rates due to the lack of MSH2 participating in DNA MMR (see Ollila et al. 2008 for review). Another basis of pathogenicity of missense mutations in MSH2 is mislocalization. MSH2 must localize to the nucleus in order to function in DNA MMR, but mutations in one of the protein's nuclear localization signals (NLSs) or in an area of interaction with its dimer partners MSH3 and MSH6 prevents the protein from localizing properly to the nucleus, as the protein reaches the nucleus either by recognition of an NLS by importin or by co-localization with a dimer partner (Hayes et al. 2009). Therefore, a missense mutation that prevents proper localization may also lead to dysfunctional MMR and, consequently, may cause increased mutation rates. Missense mutations in MSH2 have also been found to lead to increased mutation rates by means of decreasing the protein's ability to interact with its functional partners (Gammie et al. 2007). Without the ability to interact with the aforementioned binding partners, either MSH2 cannot function properly in MMR to recognize mismatched DNA or a crucial step in DNA MMR cannot occur, as interaction with MSH2 is required for the proper functioning of a number of other members of DNA MMR machinery (Gammie et al. 2007; Modrich 2005; Ollila et al. 2008).

Here, we attempt to determine the biochemical basis of the pathogenicity of the P640T mutation in yeast, which we have previously determined to cause significantly increased mutation rates. We study the mutation's effects on the MSH2 protein by testing several possible modes of pathogenicity with specific assays, namely an immunoblot to measure steady state levels of the P640T protein product, a yeast-two hybrid assay to determine the level of interaction between the protein and its aforementioned functional partners, and an

immunofluorescence assay to determine whether or not the protein was localizing properly to the nucleus to participate in MMR. We determined that MSH2-P640T is expressed at .97% of wildtype protein steady state levels, has a nuclear to cytoplasmic concentration ratio (N/C) of 1.05±.025, significantly less than the 1.25±.02 N/C of the wildtype, and that it does not interact with any of the aforementioned binding partners. Additionally, we mapped the location of the Pro640 residue in order to elucidate its structural relevance and to gain insight on the effect of the P640T mutation on the protein's folding and structure. Furthermore, we propose a model of the nature of the P640T mutation wherein *msh2*-P640T is unstable and is therefore degraded by the cell, preventing it from participating in MMR. Finally, we suggest future projects to test our model and apply this information in a clinical setting.

Author Commentary

Ramie Fathy

Part of the Molecular Biology curriculum includes Core Lab, where students learn laboratory techniques in the context of a single, 12-week-long set of experiments that study the role of the MSH2 protein in DNA mismatch repair (MMR). After each half-semester, we were required to compose a 25-page lab report that followed the format of scientific articles published in research journals today. One of the greatest challenges I faced while writing my final Core Lab paper was identifying and incorporating only relevant background information—excluding extraneous information. To overcome this, I first studied the relevant literature, using PubMed and Google Scholar to find papers on proteins and biological mechanisms similar to those I studied in Core Lab; I then used those and the papers cited therein to find more examples of good scientific writing. This served several purposes: I gained an idea of the format of scientific writing to model my paper after; I became familiar with the DNA mismatch repair process that I was studying; and I learned where my work and its findings fit in the context of the current field, allowing me to frame the information accordingly. This part of the process was very time-consuming but extremely vital to the production of my paper.

Because much of the lab work was performed directly as outlined in the lab manual, without much chance to reflect on its relevance, I was unsure of how its results fit into the larger research field. So sifting through the relevant literature allowed me to determine exactly what questions the literature had not yet answered that I could elucidate with my work. Considering the fact that much work had been done to elucidate the process of DNA mismatch repair (MMR) and related proteins, as well as the fact that my work dealt with the effects of mutating one of these MMR proteins, I decided to begin with the entire process of MMR and "zoom in" on the proteins involved in the process, MSH2 in particular, from there. This way, I could provide functional context on the protein in question while avoiding irrelevant information.

While writing this background, I imagined a model of informing the reader by assuming the reader lacks most background knowledge and that my role was to build a foundation, or "pyramid," of information for them to use to understand the rest of the paper. I developed this model after reading previous works on the subject—while reading I identified the reasons they included specific information in their Introductions. Using this model prevented me from writing at length about specific processes that were not crucial to understanding the topic of the paper, such as the specific functions of the proteins that bind with MSH2. Instead, I simply mention the function that those proteins play when bound to MSH2.

As a rule to determine what background information was necessary to include in the Introduction, when writing each sentence I asked myself how it prepared the reader for later discussion in the paper. If I found that it did not provide support for later information, I either adjusted it or got rid of the sentence entirely. Similarly, as I wrote the Discussion section and as I began to make sense of each experiment we performed and its significance, I made sure I covered that information in the Introduction at a level that would allow the reader to understand the conclusions I reached. For example, learning about the interactions between MMR proteins from previous works motivated me to include the relevant relationships and interactions between these proteins in my Introduction so that the later discussed Yeast 2-Hybrid Assay that tests these relationships has a logical basis for the reader. In this way, analysis of my lab work and its results provided an additional criterion for determining the

relevant material and how to frame it in order to motivate and contextualize my work.

Ultimately, all of this background information provides the foundation for my final introductory paragraph, where I share my findings, what they reveal about the MMR protein being studied, and how this greater understanding benefits scientific and clinical efforts moving forward.

Fellow Commentary

Brian Lax

Ramie's introduction to his Core Lab final paper is remarkable for its careful handling of evidence. He suggests above that writing his introduction was all about providing only the amount of information deemed absolutely necessary—no more, no less. Otherwise, the 30-page report might have doubled or tripled in size. So in that way it is almost as if he was curating an exhibit with limited floor space, his introduction analogous to a museum plaque: he provides just enough context for his reader to eventually appreciate the significance of his claims even as he avoids leading the reader to premature conclusions. This can be a tricky balance to strike; above he describes a helpful litmus test to achieve it. If a piece of information did not directly or explicitly lead to a future conclusion, he axed it from the introduction. To put this litmus test to work, we include later in this issue an excerpt from Ramie's Discussion section, accompanied by commentaries from Ramie and Professor Aleksandra Snyder of the Molecular Biology Department. Its inclusion allows our readers to come to their own conclusions about the concision of Ramie's introduction.