Mutations of the MECP2 gene's effect on schizophrenia

Written by James Escobedo and Rosalee Whithaus

Abstract

Schizophrenia is a severe mental disorder affecting about 1% of the population, characterized by hallucinations, delusions, and cognitive impairment. Genetic factors play a significant role in schizophrenia, but identifying specific mutations is challenging. This study focused on the MECP2 gene, which is crucial for neural development, to understand its role in schizophrenia. Using advanced software like AlphaFold and AlphaMissense, we analyzed the impact of a specific mutation (rs61752387) in MECP2. Our results showed conflicting evidence: structural analysis suggested the mutation might be pathogenic, while another method indicated it was benign. Further research is needed to clarify the mutation's role in schizophrenia.

Introduction

Schizophrenia is a disorder characterized by hallucinations, which may be visual, tactile, or auditory. Schizophrenia affects around 1% of the general population, which equates to 800,000 people. Schizophrenia has a high genetic factor, which is said to be up to 80%.[4] Identifying specific genetic mutations contributing to schizophrenia remains challenging, complicating the development of effective treatments. This is the reason we are studying our specific gene: to help identify a factor that might cause schizophrenia, which in turn will help other researchers find potentially highly effective treatments.

Schizophrenia is defined as a chronic and severe mental disorder that affects a person's thinking, feelings, and behavior. It is characterized by symptoms such as delusions, hallucinations, disorganized speech, and impaired cognitive function, leading to significant social and occupational dysfunction. The exact cause of schizophrenia is not fully understood, but it involves a combination of genetic, biological, environmental, and psychological factors. Treatment typically includes antipsychotic medications and psychosocial therapies aimed at managing symptoms and improving quality of life.[4]

Our goal was to carefully pick a protein significant to causing schizophrenia and then try to visualize a 3D model of that protein with and without mutations to see what it does.

Methods & Results

We found the paper "Detection of Rare Methyl-CpG Binding Protein 2 Gene Missense Mutations in Patients With Schizophrenia" by Chen et al.[1] that observed an overlap of rare variant risk between *schizophrenia*, autism spectrum disorder, epilepsy, and other severe neurodevelopmental disorders. We downloaded the supplementary table for this paper along with the supplementary table from the paper "Rare coding variants in 10 genes confer substantial risk for schizophrenia" by Singh et al.[2] Two supplementary tables were used in order to find suitable proteins for the prediction and visualization process. We also downloaded all of the associations on the GWAS Catalog for *schizophrenia* for analysis.

Table 1:
schizophrenia GWAS Schizophrenia Association GWAS Associations for *schizophrenia* [3]

Table 2:
schizophrenia GWAS Schizophrenia Association Filtered 70 coding mutations were found within the data.

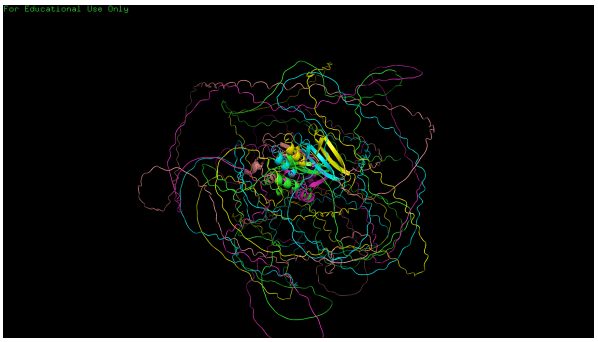
Using the association table, we tried to find mutations that affected coding regions of a gene and affected genes that had clear function. We found several suitable proteins but settled on the Methyl-CpG Binding Protein 2 (MECP2) gene. The MECP2 gene is a transcriptional regulator required for proper neural development and could be considered a risk gene of *schizophrenia*. The mutation being observed is under the rsID rs61752387.

The mutation is an amino acid change from proline (P) to serine (S) at position 376 within isoform 1 of the gene.

In order to visualize the 3D structures of the MECP2 gene with and without mutation, we utilized AlphaFold. AlphaFold is an artificial intelligence software that predicts the structure of a protein. AlphaFold structures are noted to be vastly more accurate than competing methods of predicting protein structures.[5]

pyMOL Figures:

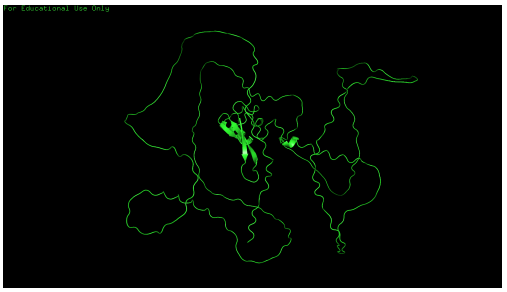
a.



i. Isoform 1 all 5 structures aligned without mutation

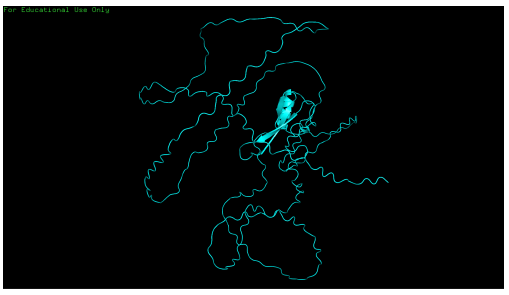
ii. RMSD - 40.711

iii.



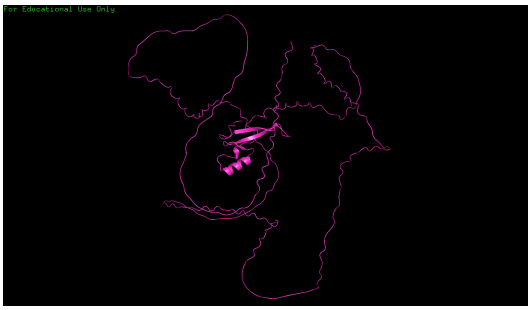
1. Structure 1 without mutation

iv.

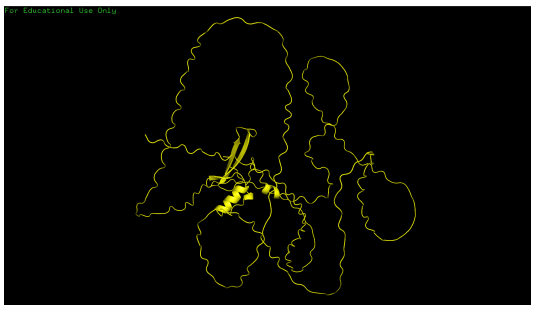


1. Structure 2 without mutation

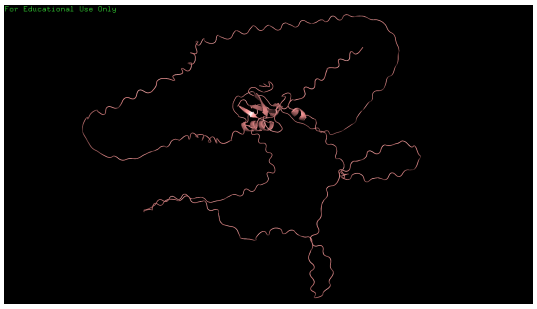
v.



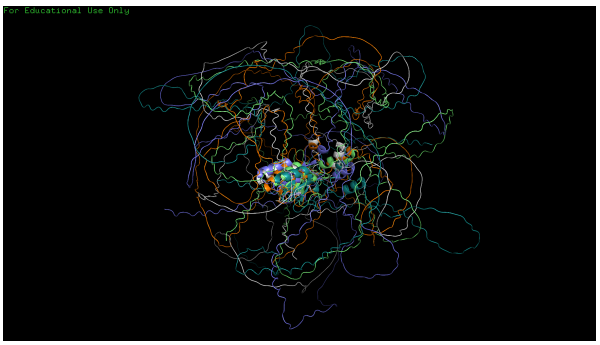
1. Structure 3 without mutation

vi. 

1. Structure 4 without mutation

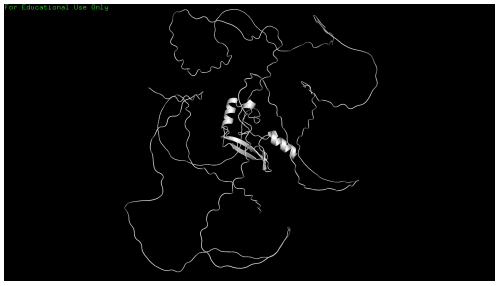
vii. 

1. Structure 5 without mutation

b. 

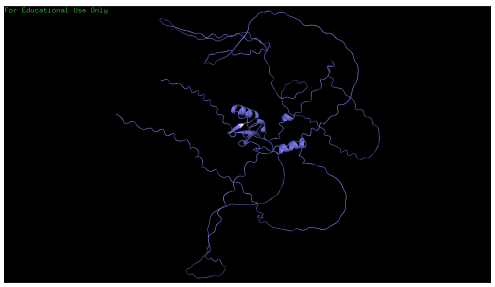
i. Isoform 1 all 5 structures aligned with mutation

ii. RMSD - 37.907

iii. 

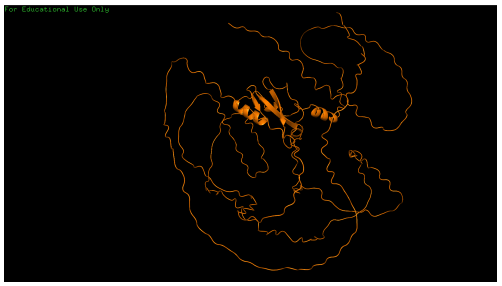
1. Structure 1 with mutation

iv.



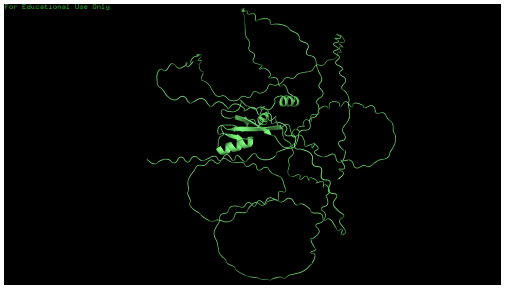
1. Structure 2 with mutation

v.



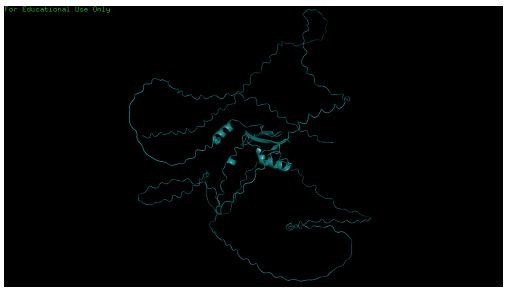
1. Structure 3 with mutation

vi.



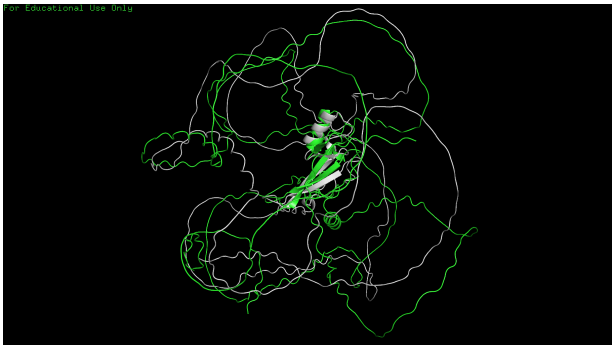
1. Structure 4 with mutation

vii.



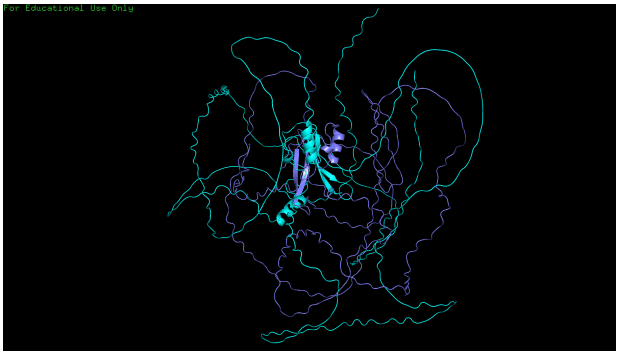
1. Structure 5 with mutation

c.

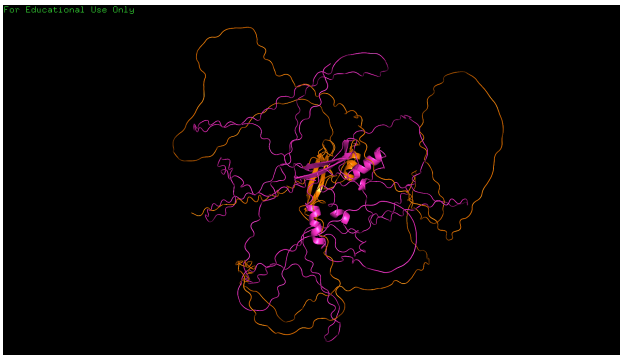


i. Structure 1, with and without mutation, aligned

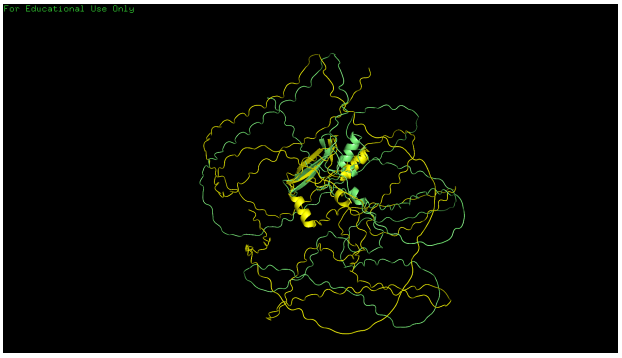
ii. RMSD - 9.490

d. 

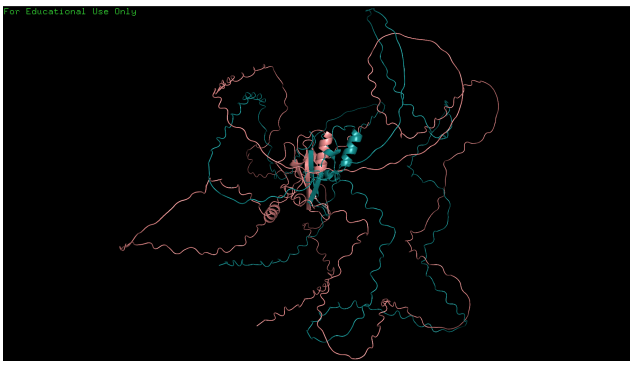
i. Structure 2, with and without mutation, aligned
ii. RMSD - 40.489

e. 

i. Structure 3, with and without mutation, aligned
ii. RMSD - 45.801
1. **RMSD suggests mutation alters global structure**

f. 

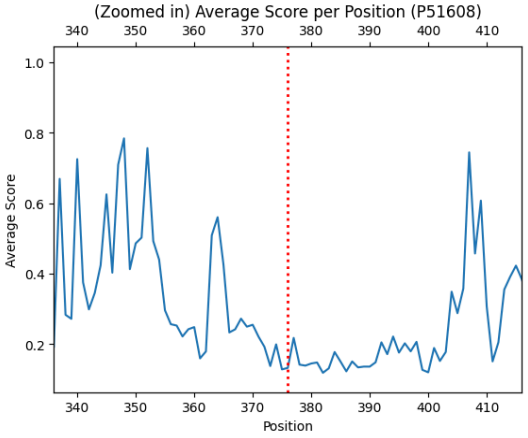
i. Structure 4, with and without mutation, aligned
ii. RMSD - 12.403

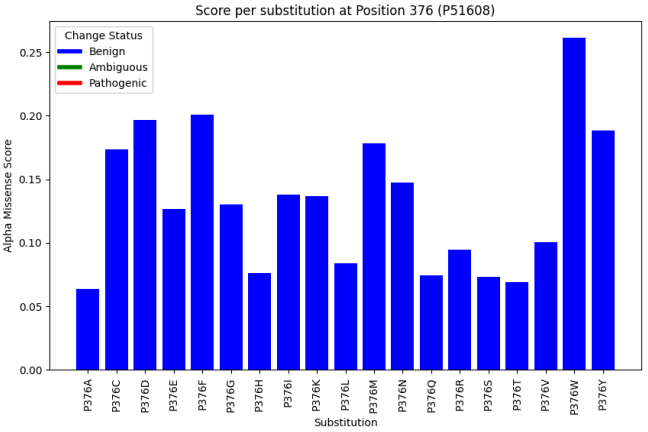
g. 

i. Structure 5, with and without mutation
ii. RMSD - 16.992

AlphaMissense is another software we used. AlphaMissense is based off of AlphaFold but is fine-tuned on labels distinguishing variants found in human and closely-related primate populations. The software doesn't predict the change in protein structure upon mutation but instead produces between zero and one approximately rating the likelihood of a variant being pathogenic.

AlphaMissense Figures:





We looked at these alpha missense predictions to get a sense of where the important parts of the protein are, and where our mutation is rated relative to those parts. While some predicted structures in AlphaFold were noted to possibly be altering the global structure of the gene. However, AlphaMissense shows us that the mutation may not be in a very important part of the protein.

In addition to using AlphaFold to assess the impact of the mutation, we performed a sequence alignment of this protein across many species to test the hypothesis that this mutation is selected against. If the mutation is absent in our alignment, it's likely that, for some evolutionary reason, it is detrimental to cells. In order to get the alignment, we downloaded all sequences of MECP2 from the National Center for Biotechnology Information orthologs. We then ran the sequences through Kalign, a multiple sequence alignment server and received an aligned output.[7]

Kalign Figures:

The mutation we are looking for is a change from P to S, proline to serine, so the fact that S cannot be found in column 153 may be evidence that the mutation is actually detrimental to cells.

Discussion

Based on our research, there appears to be some conflicting analysis on how detrimental the mutation is. The sequence alignment and the AlphaFold structure predictions both support that the mutation is more pathogenic. However, the AlphaMissense figures suggest that mutations at position 153 are completely benign. These ambiguous results highlight the necessity for a deeper dive of research into the MECP2 gene and the significance of its mutations.

References

1. Chen, C. H., Cheng, M. C., Huang, A., Hu, T. M., Ping, L. Y., & Chang, Y. S. (2020). Detection of Rare Methyl-CpG Binding Protein 2 Gene Missense Mutations in Patients With Schizophrenia. Frontiers in genetics, 11, 476. https://doi.org/10.3389/fgene.2020.00476
2. Singh, T., Poterba, T., Curtis, D., Akil, H., Al Eissa, M., Barchas, J. D., Bass, N., Bigdeli, T. B., Breen, G., Bromet, E. J., Buckley, P. F., Bunney, W. E., Bybjerg-Grauholm, J., Byerley, W. F., Chapman, S. B., Chen, W. J., Churchhouse, C., Craddock, N., Cusick, C. M., DeLisi, L., … Daly, M. J. (2022). Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature, 604(7906), 509–516. https://doi.org/10.1038/s41586-022-04556-w
3. 5082 associations were downloaded from the NHGRI-EBI GWAS Catalog (Sollis et al., 2022) on 05/23/2024 under EFO ID MONDO_0005090 (schizophrenia).
4. Stöber, G., Ben-Shachar, D., Cardon, M., Falkai, P., Fonteh, A. N., Gawlik, M., … Riederer, P. (2009). Schizophrenia: From the brain to peripheral markers. A consensus paper of the WFSBP task force on biological markers. The World Journal of Biological Psychiatry, 10(2), 127–155. https://doi.org/10.1080/15622970902898980
5. Jumper, J. et al. “Highly accurate protein structure prediction with AlphaFold.” Nature, 596, pages 583–589 (2021). DOI: 10.1038/s41586-021-03819-2
6. Mirdita, M. et al. “ColabFold: Making protein folding accessible to all.” Nature Methods, 19, pages 679–682 (2022). DOI: 10.1038/s41592-022-01488-1
7. Lassmann, T., Sonnhammer, E.L. Kalign – an accurate and fast multiple sequence alignment algorithm. BMC Bioinformatics 6, 298 (2005). https://doi.org/10.1186/1471-2105-6-298