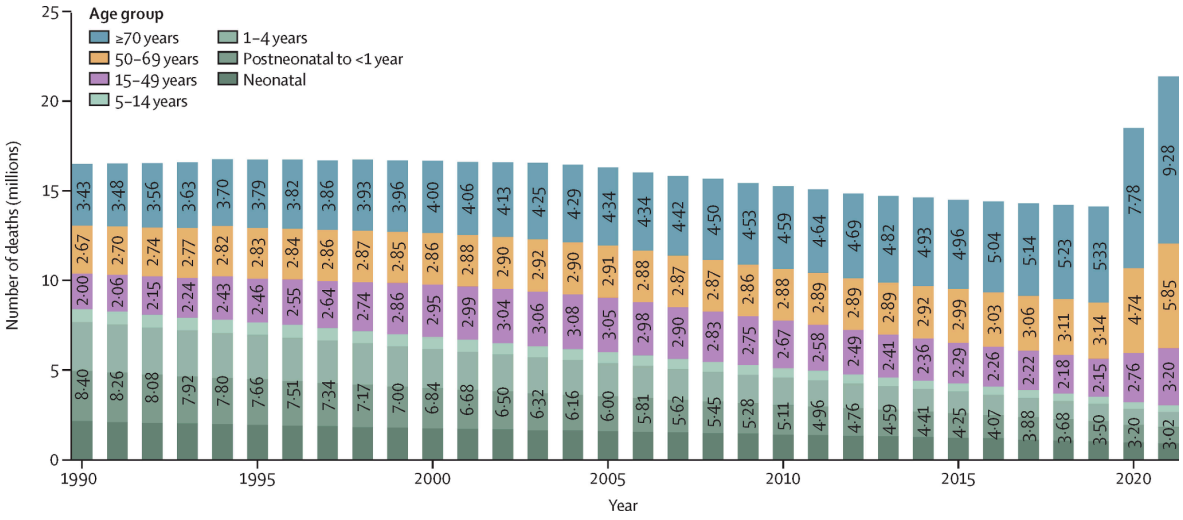
CYSF 2025/2026

ZAFINA ZAMAN

Grade 11

Project Type: Research/ Study

Combating Antifungal Resistance Through Secondary "Rescue" Mutations in *Candida auris*

DATE AND ENTRY	NOTES
2025-08-16 I began looking into the problem of Bacterial Resistance, or antimicrobial resistance (AMR). It's a really relevant problem since if it's not solved, people could start dying from the common cold just like they did in the 1800s and onward. After people came up with vaccines and whatnot, the death rate went down. However, due to AMR, simple viruses could become	--to fight AMR, ppl have created the Pioneering Antimicrobial Subscription To End Upsurging Resistance (PASTEUR) --to create new antibiotics, it would be almost $6 mill for taxpayers!!1 --92 million deaths associated with (AMR) will occur worldwide between 2025 and 2050…2 --from 1990-2021 deaths from AMR decreased by about 50% for kids younger than 5.2 --methicillin-resistant *Staphylococcus aureus* AMR mortality rate increased the most globally from 1990-2021 from 261000 associated deaths and 57200 attributable deaths, to 550000 associated deaths and 130000 attributable deaths.2 --WHO Bacterial Priority Pathogens List 2024's is a guide made by the World Health Organization to find bacteria that are most dangerous. Highlights critical need to address gram-negative pathogens, and drug-resistant tuberculosis (RR-TB) as high priorities.3  Image 1 link, from source 2

lethal once more.
**I will be consistently abbreviating antimicrobial resistance as AMR throughout my research and notes.

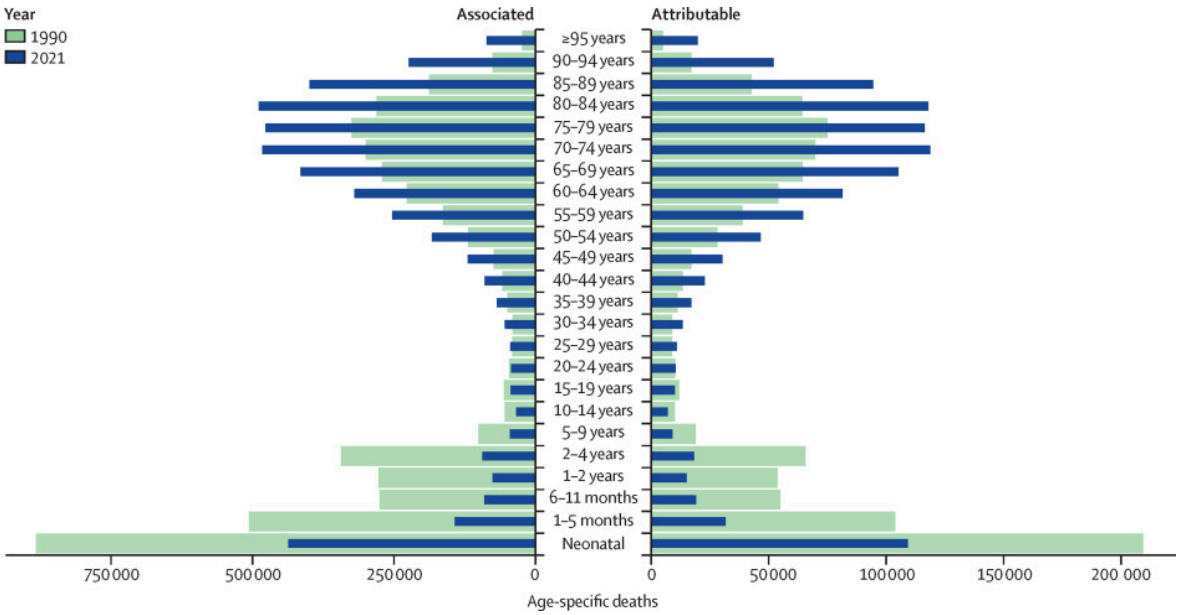


Image 2 link, from source 2

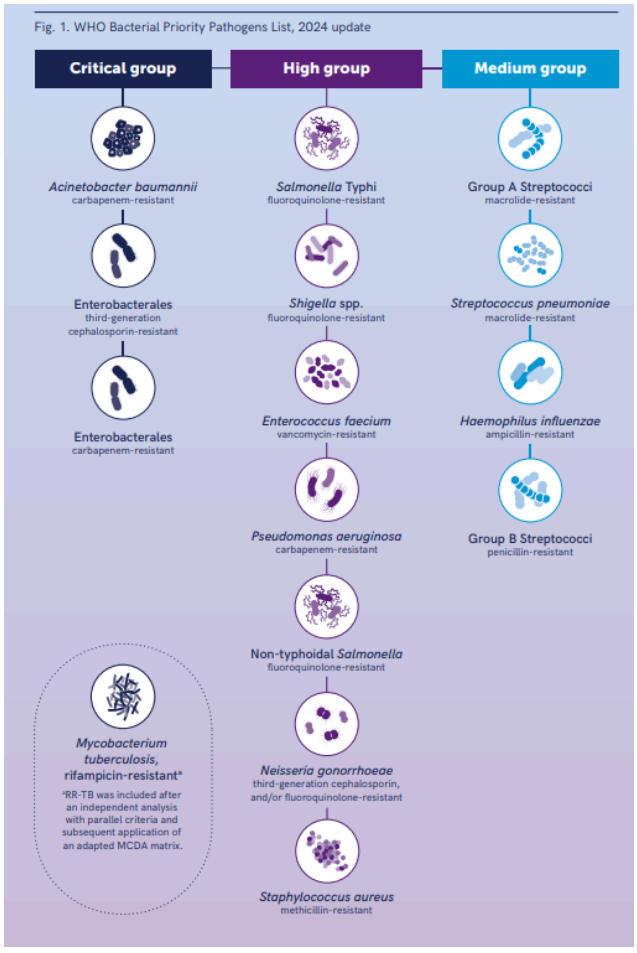


Image 3 link, page 13, source:WHO Bacterial Priority Pathogens List 2024

2025-08-18

Solution	Process
Reduce antibiotic overuse	
Mandate public reporting of antibiotic use and link to pay for performance incentive for lower use and disincentive for higher use	Regulation or legislation
Focus research funds to reduce reliance on antibiotics: • Vaccines to prevent bacteria/fungal infections, and AMR pathogens in particular • Immunotherapies to prevent or treat bacterial/fungal infections • Phages • Microbiome-based strategies to prevent or eliminate AMR colonization • Better infection-prevention implementation • Short-course antibiotic therapy studies • Rapid diagnostics and biomarkers to reduce antibiotic prescriptions • Novel psychological approaches, such as nudges, to reduce antibiotic prescriptions	Regulation
Require special training or certification to prescribe new antibiotics and those addressing unmet need	Regulation or legislation
Ensuring future antibiotic R&D	
Establish a board including clinical experts to better target existing incentives	Legislation
Establish endowed nonprofits to conduct early stage discovery and R&D of AMR solutions	Legislation

Abbreviations: AMR, antimicrobial resistance; R&D, research and development.

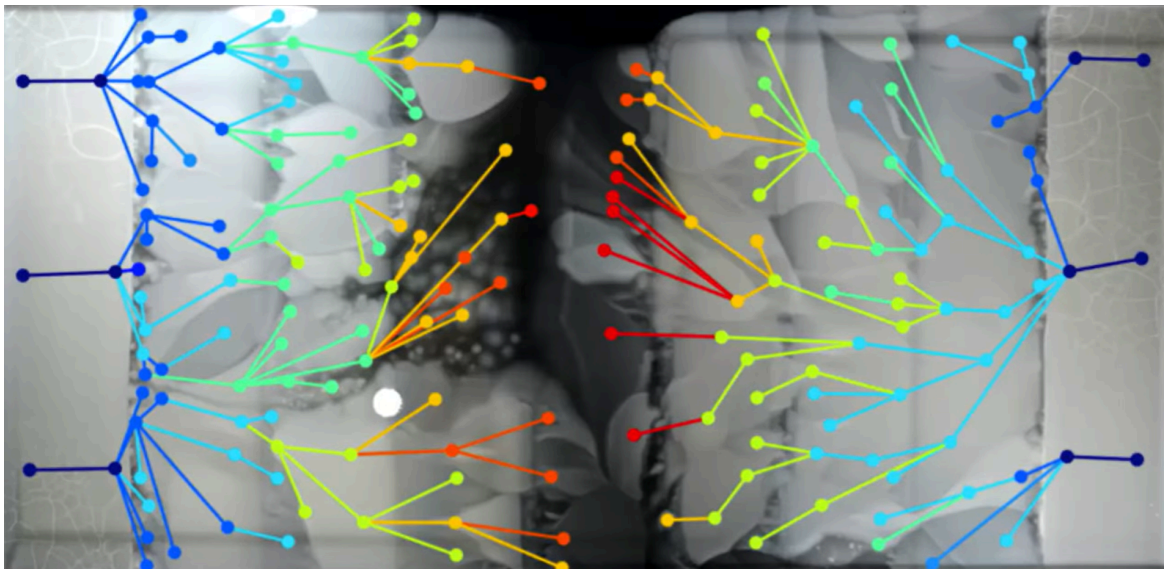
Image 4 link, table 3, source:1

--impact of antibiotic concentration towards AMR→higher concentration leads to stagnation (death), while lower allows for multiplication
--bacterial biofilms→jelly-like colony structures inside where bacteria are suspended. Acts like a protective barrier, so the bacteria only receives a fraction of the antibiotic. *Harder to target*.
-- mutations alter bacteria characteristics like cell walls/ enzyme activity which makes antibiotics *less effective*.
EX: Enzyme deactivation (bacteria develop enzymes that deactivate antibiotics), antibiotic ejection, bacterial wall prevention (mutations make it difficult for antibiotics to enter bacteria).
--mutated bacteria transfer "survival skills" onto new bacteria. Now they're both resistant.

EXAMPLES OF ANTIBIOTIC RESISTANCE
--Carbapenem-resistant enterobacteriaceae (CRE).. some strains are incurable and *resistant to all antibiotics.* Mortality rate is **50%** for patients with bloodstream infections.
--Clostridium difficile (C. difficile)... commonly found in hospitals and group homes; associated symptoms fatal in elderly. *Resistant to many antibiotics.* Makes spores which are hard to kill
--Neisseria Gonorrhoeae…sexually transmitted. **2nd more common infection in North America**. Abt **30**% of infections are now resistant. Serious *reproductive complications.*
Source for above paragraph (Biomed Microbio Slides)

--resistance arises from spontaneous mutations (inevitable), horizontal gene transfer, and antibiotic overuse
--preventative measures include using *narrow-spectrum antibiotics*→an antibiotic that is only able to kill or inhibit limited species of bacteria.4
--COVID pandemic redirected resources and scientific attention away from AMR..5

Evolution of bacteria video
https://en.wikipedia.org/wiki/File:Kishony_lab-The_Evolution_of_Bacteria_on_a_Mega-Plate.webm



2025-08-19

Today I will get into the nitty gritty and actually look at the mechanisms behind how AMR develops/ how it works. That'll allow me to deepen my understanding of this problem.

THE SIX MECHANISMS OF ANTIMICROBIAL RESISTANCE 6

1. Drug inactivation/ modification

- Enzymatic deactivation of penicillin G in some penicillin resistant bacteria thru the production of B-lactamases. Can be done chemically through the addition of *functional groups* (substitute in a molecule causing its reactions) by *transferase* enzymes (catalyse transfer of functional groups). For example, acetylation, phosphorylation, or adenylation are common resistance mechanisms to aminoglycosides. Acetylation is the most widely used mechanism and can affect a number of drug classes.

2. Alteration of Target/ Binding Site

- Alteration of PBP (binding site of penicillins) in MRSA (a penicillin resistant bacteria).

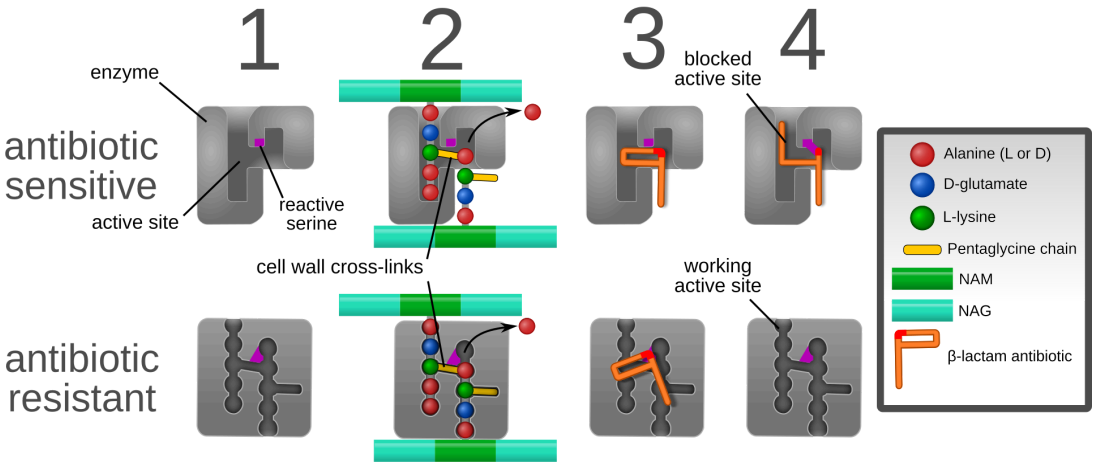


image 5 link, from source 6

3. Ribosomal Protection Proteins

- Protect the bacterial cells from antibiotics that target the cell's ribosomes to inhibit protein synthesis.
- Involves binding of ribosomal protection proteins to the ribosomes of cells. So it changes its shape. Allows ribosomes to continue synthesizing proteins essential to cells but also prevent antibiotics from binding. (more info at source 7)

4. Alteration of Metabolic Pathway

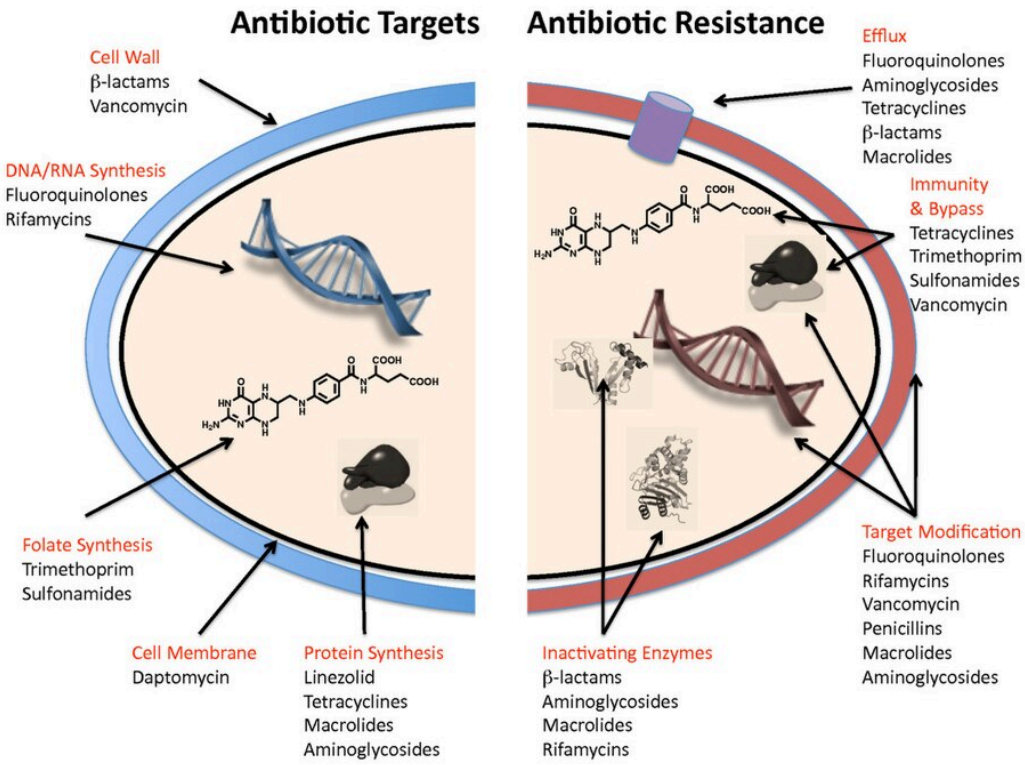
- Some sulfonamide resistant bacteria don't require para-aminobenzoic acid..PABA..(important for the synthesis of folic acid and nucleic acids in bacteria inhibited by sulfonamides. But like mammalian cells they turn to using preformed folic acid.

5. Reduced Drug Accumulation

- Decrease drug permeability or increasing pumping out. Pumps out antibiotics before they do anything…

6. Ribosome Splitting and Recycling

- Liberation of the ribosome from the drug allows further translation and more resistance to the drug



h

image 6 link, from source **6**

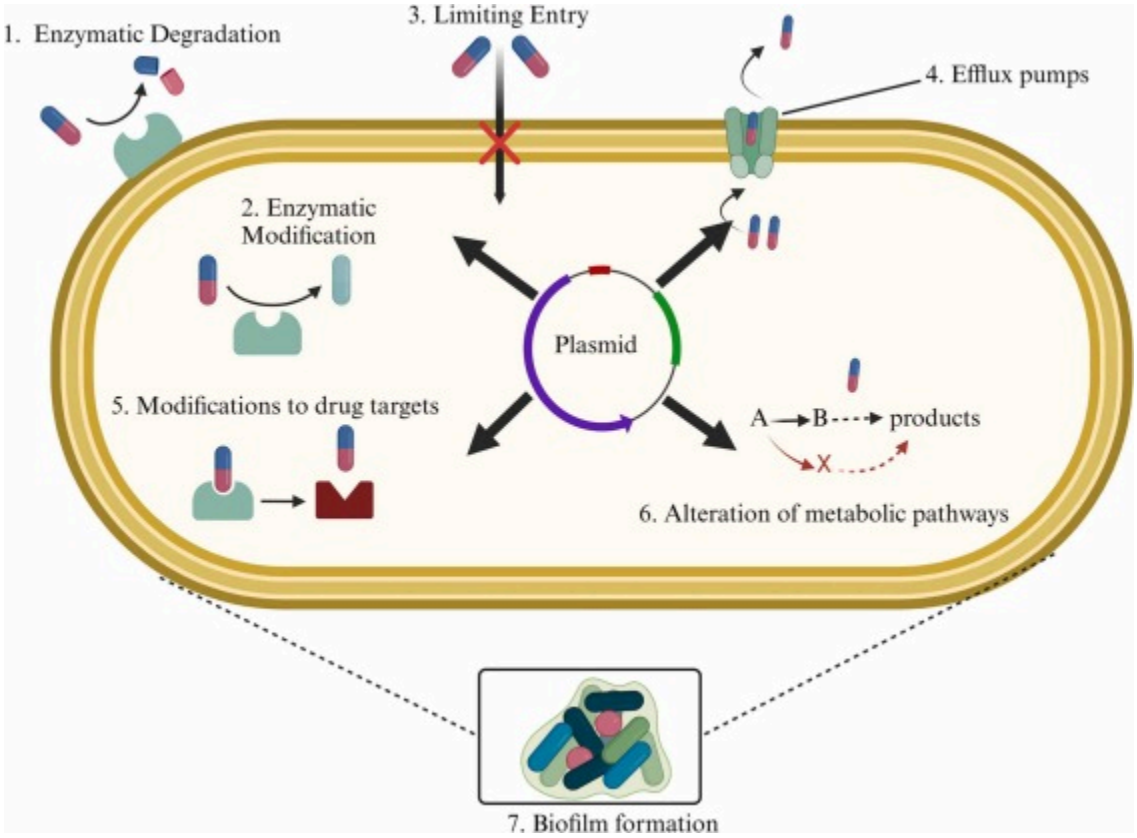


image 7 link, from source 7, figure 1

STRATEGIES TO COMBAT AMR FROM SOURCE 8

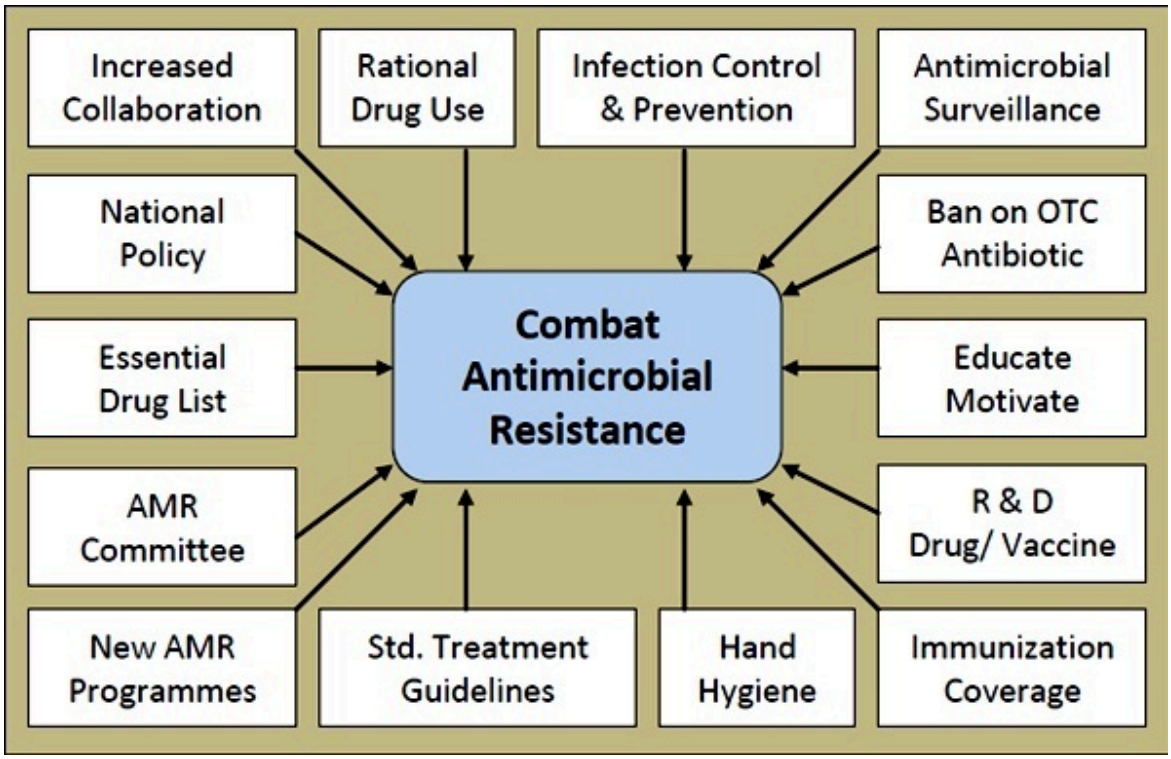


image 8 link

CURRENT DEADLIEST ANTIBIOTIC-RESISTANT BACTERIA 9

- Escherichia coli (E. coli).
- Staphylococcus aureus (S. aureus).
- Klebsiella pneumoniae (K. pneumoniae).
- Streptococcus pneumonia (S. pneumoniae).
- Acinetobacter baumannii (A. baumannii).
- Pseudomonas aeruginosa (P. aeruginosa)

CURRENT DEADLIEST SUPERBUGS 9

- C. diff (*Clostridioides difficile*).
- Drug-resistant gonorrhea.
- Methicillin-resistant *Staphylococcus aureus* (MRSA).
- Multidrug-resistant *Mycobacterium tuberculosis* (MDR-TB).
- Vancomycin-resistant *Enterococci.*

2025-08-20

Today I'll look deeper into the top deadliest AMR bacteria, E. coli. Hopefully that'll allow me to really understand the two and the

E. COLI-- AKA ESCHERICHIA COLI 10

- Gram negative bacteria commonly found in the lower intestine of warm-blooded organisms
- Part of the normal microbiota of the gut, where they are about 0.1% along with other *facultative anaerobes* (organisms making ATP by aerobic respiration).
    - These bacteria are mostly harmless. Can be beneficial. Some strains of E. coli benefit by making vitamin K2 or by preventing colonization of the intestine from bad pathogens.
- Some stereotypes like EPEC and ETEC are pathogenic and cause serious food poisoning in hosts.

role antibiotics play (or doesn't in this case…)

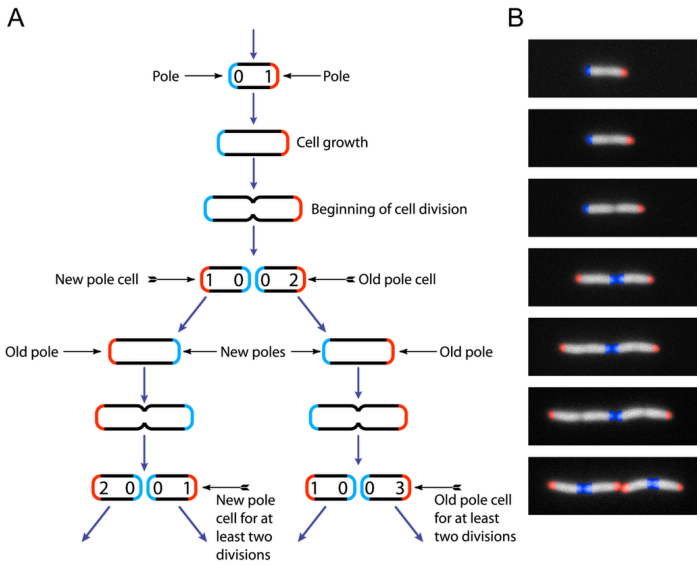


image 9 link

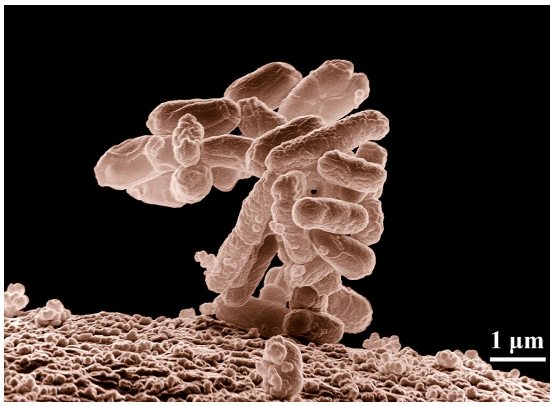


image 10 link

- Transfer DNA through horizontal gene transfer.
- E. coli is already implemented in a multitude of national AMR surveillance programs.

After this finding i decided to move to a different bacteria----->

I found an actual meeting record from the WHO on an AMR Surveillance meeting. It occurred on March 21-23, 2023 in Buenos

AMR SURVEILLANCE WHO MEETING 11

- A particular concern was voiced in the meeting about a pathogen called *Candida auris*, which suddenly emerged and has already been reported worldwide. It's highly transmissible and highly resistant to antifungals (page 14)

**GLASS→Global Antimicrobial Resistance Surveillance Programme

- Moving forward they decided to:
  - Strengthen fungal lab capacity for countries
  - Work on developing this part of the new WHO AMR labb strategy
  - Develop a Quality Control (QC) program for labs

Aires, Argentina.This is actually so cool! https://cdn.who.int/media/docs/default-source/antimicrobial-resistance/amr-spc-sel-glass/4th-meeting-report-(upload).pdf?sfvrsn=fc00167c_3&download=true	• I noticed that they didn't mention any clear steps on making new antibiotics for it.. Maybe due to costs for production? • I noticed that solving the entire issue of AMR is very difficult because it's a very large problem. Each pathogen has a different antibiotic and responds to it differently. They all have different mechanisms too. So I think I need to really hone into one thing.

2025-08-21
So *Candida auris* is a very new and scary thing. I was just searching it up, and there are news articles about it spreading and how deadly it is from just **23 HOURS AGO…** so this is very scary indeed!
Today I'll just be learning about this fungus, and maybe see where I can tie it to AMR.
	CANDIDA AURIS 12
	• Commonly spreads in hospitals, and long-term care facilities
	• Can cause:
	  ○ Ear infections
	  ○ Wound infections
	  ○ UTIs
	  ○ Infections in your blood that can spread to other places in your body
	• C. auris doesn't live naturally on your body-- you get it from contaminated surfaces/ others who have it
	• It's a rare disease, but it is emerging… there has been an increase since 2019
	• Symptoms:
	  ○ Fever
	  ○ Chills
	  ○ Lethargy (extreme tiredness)
	  ○ Low BP
	  ○ High heart rate
	  ○ Hypothermia
	  ○ Pain, pressure, or feeling fullness in your ear..?
	• Complications once it spreads to your blood:
	  ○ Sepsis
	  ○ Shock
	  ○ Organ failure
	• Strains of C. auris are becoming resistant…making it harder to treat
	ARTICLES
	
	CANDIDA AURIS: KILLER FUNGUS WARNING AS "GLOBAL THREAT" ALREADY DETECTED IN UK 13---(this article was written on August 11th, 2025… which was 10 days ago…)
	• Has been detected in more than 40 countries, with b/w 30%-60% being **FATAL**

image 11 link

- Fungus death rate→1 in 3
- WHO listed it as among the 19 lethal fungi and said its a “threat to global public health”
- Resistant to disinfectant and meds..
- Spreads to
    - Blood
    - Brain
    - Spinal chords
    - Bones
- More and more resistant due to overuse of antifungal agents

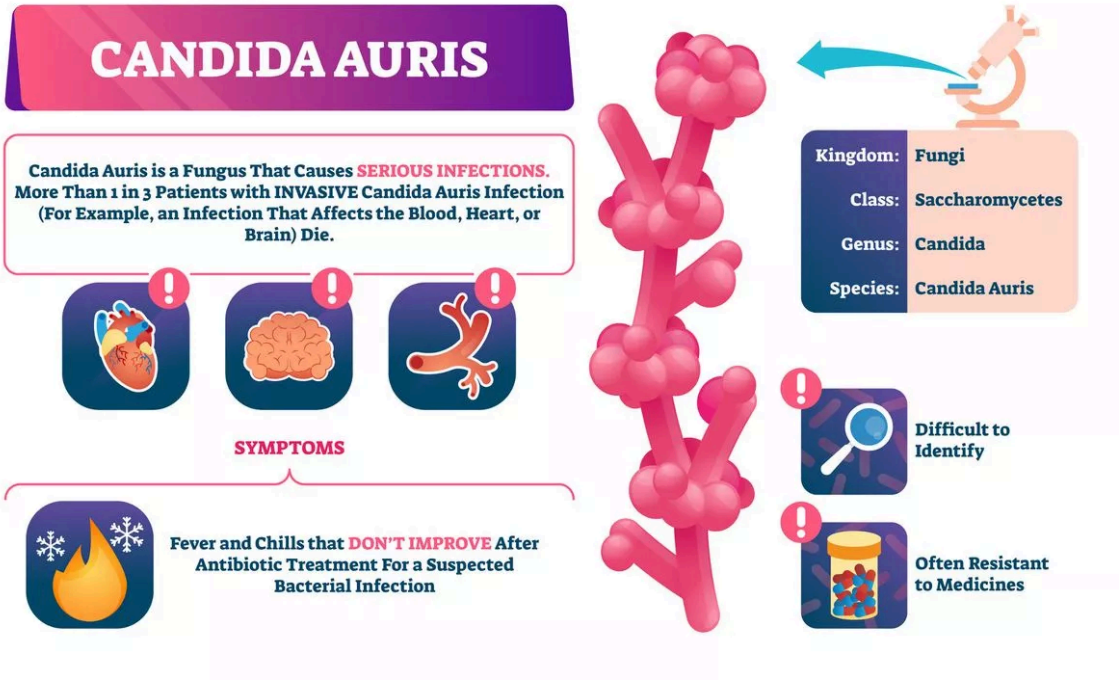


image 12 link

DOCTORS ISSUE FRESH WARNING OVER KILLER FUNGAL INFECTION POSING A “SERIOUS THREAT TO HUMANITY” 14

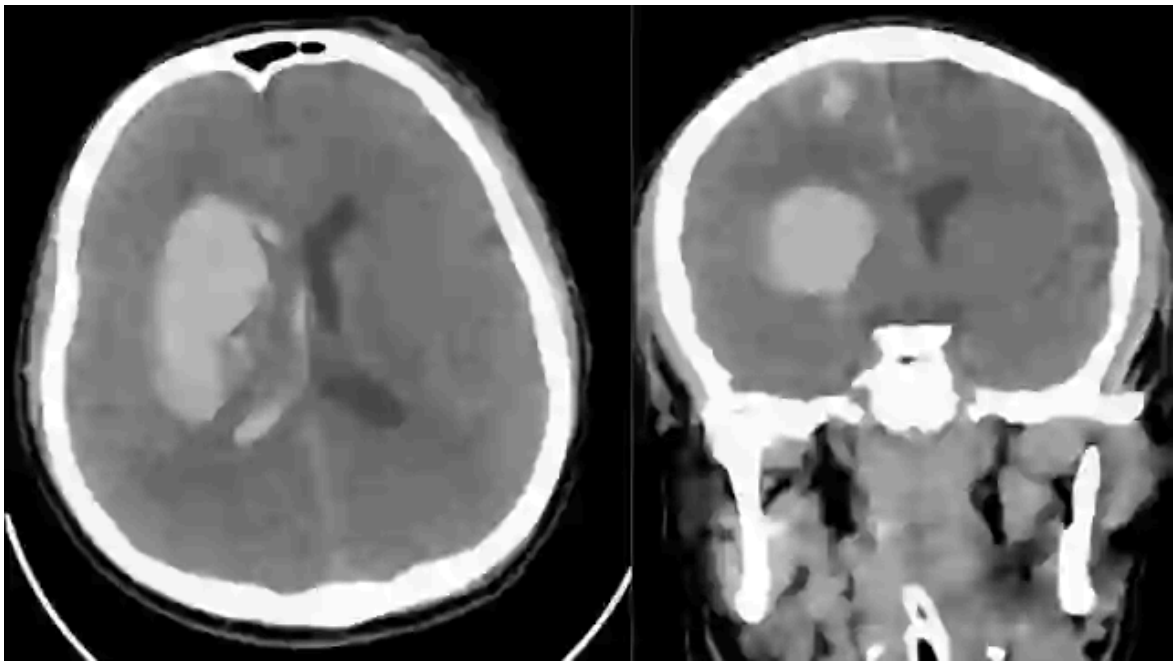


image 13 link

- First found in 2009
- Survives on skin and surfaces for long time

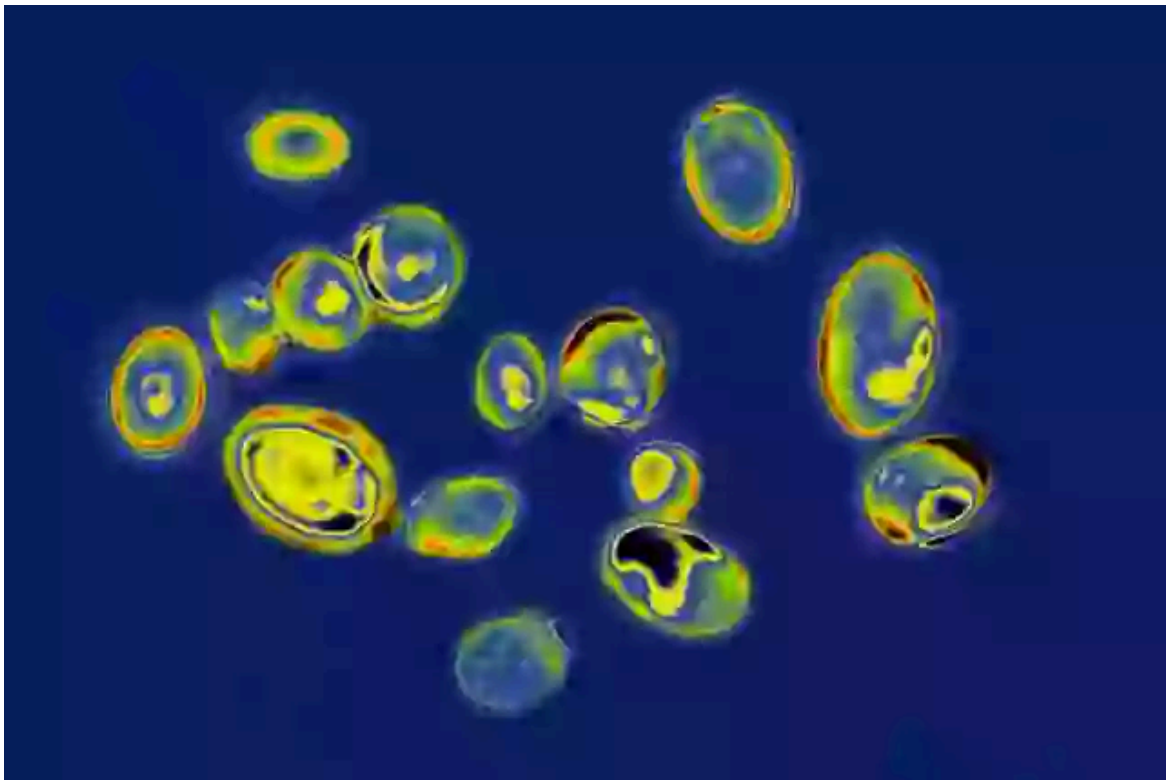


image 14 link

- Ppl with weakened immune systems at more risk
- Can develop antifungal resistance quickly, and some strains have developed resistance to ALL strings

CANDIDA AURIS CASES NEARLY TRIPLE AS DEADLY FUNGUS SPREADS TO NEW STATES

15--(this news article was written 23 hours ago…)

- Cases have tripled in the last 3 ½ months
- End of april-- 1052 cases, in aug 9--2809 cases across 21 states (this is US only data)
- CDC said its "**an urgent antimicrobial resistance threat"** (AMR MENTIONED)

CURRENT DRUGS USED TO TREAT-- FROM THE CDC 15

Echinocandin Drug	Adult dosing	Pediatric dosing ≥ 2 months of age
Anidulafungin	loading dose 200 mg IV, then 100 mg IV daily	not approved for use in children
Caspofungin	loading dose 70 mg IV, then 50 mg IV daily	loading dose 70mg/m^2/day IV, then 50mg/m^2/day IV (based on body surface area)
Micafungin	100 mg IV daily	2mg/kg/day IV with option to increase to 4mg/kg/day IV in children at least 40 kg

2025-08-24

- Echinocandins are the initial recommended treatment for adults and children over 2 months.
- Amphotericin B deoxycholate at 1 mg/kg daily is the initial recommended treatment for infants under 2 months.
- Reports of echinocandin-resistant and pan-resistant cases are increasing.

image 15 link

I used Chat GPT to find some free online research for gene modelling and found these. This is the link to the conversation I have-- it also helped me find possible research questions

https://chatgpt.com/share/68c22953-8678-800d-840d-18e0e30c20b3

Useful free tools & data sources (you can use them from home)

- **NCBI Genome / SRA** — raw reads and assemblies for *C. auris* (download sequences for comparative work). NCBI +1
- **NCBI PubMed / PMC** — review articles and primary studies about resistance mechanisms. Nature +1
- **Candida Genome Database (CGD)** — gene pages and annotations for *Candida auris*. Great for gene IDs and literature links. PubMed +1
- **ColabFold / AlphaFold Colab** — free notebooks to predict protein structures without heavy local compute. Perfect for student projects. Google Colab +1
- **Galaxy (usegalaxy.org)** — web platform for beginner-friendly sequence QC, mapping, variant calling and more (no installation). Studies have used Galaxy to analyze *C. auris* SRA reads. PMC
- **MAFFT / Clustal / MEGA / IQ-TREE** — alignment and phylogeny tools (MEGA is GUI-friendly; IQ-TREE is good for more advanced phylogenies).
- **Figshare / public genome collections** — some groups publish global *C. auris* read sets you can reuse. Figshare

4. **Protein structure modeling** — pick a *C. auris* protein implicated in resistance (e.g., Erg11) and predict its 3D structure with AlphaFold/ColabFold; then compare wild-type vs published resistance mutation(s) to visualize possible structural impacts. (No lab work, purely in silico.) Google Colab +1

It gave a lot of info on project ideas and resources but I decided not to dig any deeper since I really didn't want it to be giving me answers/ actual steps.

I liked this topic idea. I'd like to look into the molecular mechanisms in C. Auris's resistance and its proteins related. From there I can edit the genes and perhaps mutate the virus?

Protein modeling & mutagenesis tools

- **SwissSidechain Mutate Tool** (SwissSidechain/SwissSidechain online services) — lets you introduce point mutations into proteins and visualize effects.
- **SwissSidechain DeepView (SwissSidechain Viewer)** — free software for protein visualization + mutagenesis.
- **UCSF ChimeraX** — powerful free visualization software; you can load a protein (e.g., Erg11), introduce mutations, and inspect binding sites.
- **SwissSidechain PredictProtein / HOPE** — online tools that predict structural and functional effects of point mutations.

Structural prediction / modification

- **ColabFold (free Google Colab notebooks)** — based on AlphaFold, predicts 3D structures from sequence. You can edit the sequence (add your mutation) and rerun to see structural changes.
- **SwissSidechain SwissSidechain (SwissSidechain)** — allows simple point-mutation modeling on existing structures.

Drug binding & docking

- **SwissDock** — free online docking platform where you can test how antifungal drugs (if you download the structures from PubChem) bind to wild-type vs mutated versions of the protein.
- **PyRx (free software)** — virtual screening/docking tool with a GUI, good for beginners.

Databases for sequences & structures

- **NCBI GenBank** — for DNA/protein sequences of *C. auris* resistance genes (ERG11, FKS1, etc.).
- **Protein Data Bank (PDB)** — for crystal structures or predicted models of fungal proteins (some already available).
- **Candida Genome Database (CGD)** — annotated gene info, resistance-related mutations.

2021-09-11

Start to dig deep into the virus and understand its mechanisms.

Candida Auris: Epidemiology, biology, antifungal resistance, and virulence (link)

- Belongs to the CTG clade-- a group of fungi, primarily yeasts from the Candida genus, that share a distinctive genetic trait: they translate the DNA codon CUG into the amino acid serine instead of the standard leucine. supports theories of genetic code evolution and expands the protein diversity and adaptability of these fungi.
- Related to the often multi-drug resistance species *Candida haemulonii*
- It was first isolated from the ear discharge of a female patient in Japan, 2009
- 40% mortality rate
- C. Auris as been isolated in over 40 countries across 6 continents
- Outbreaks in hospitals across the globe
- Nosocomial fungal infection
- *C. auris* grows well at 42 degrees celsius, so you can use it to differentiate it from other *Candida* species.
- There are amphotericin B- and fluconazole-resistant isolates
- *C. auris* can form biofilms, undergo filamentation, and phenotypically change between specific cell types→ associated w/ virulence, antifungal tolerance, and survival natural *and* host niches
- Thermotolerance and osmotolerance→ can survive on biotic and abiotic surfaces for long periods of time (can tolerate salt concentration >10% NaCl, wt/vol)
- Morphological plasticity. Rapidly adapts to environmental changes. Can switch b/w cell types in response to environmental cues
- Resistant to fluconazole

- Ergosterol is the major sterol component of fungal membranes and is the target of fluconazole and amphotericin B
- Fluconazole inhibits cellular ergosterol biosynthesis
- ERG11 encodes lanosterol demethylase

    - Many antifungals (fluconazole) directly target the ERG11 protein
    - By binding to the enzyme's active site, azoles block the removal of the methyl group, **halting ergosterol production.** This causes toxic sterol buildup and weakens fungal cell membranes
    - Resistance in *Candida* often comes from **mutations in ERG11** → these mutations change the enzyme's structure so azoles don't bind as well.
    - Three mutations (Y132F, K143R, and F126L or VF125AL) have been found in ERG11 in fluconazole resistant *candida*

2025-09-12

Dig deeper into the resistance to fluconazole

The molecular and genetic basis of antifungal resistance in the emerging fungal pathogen Candida auris (link)

- Flucanazole works by inhibiting sterol demethylase→ ERG11 is a mechanism for fluconazole resistance since it actively encodes for sterol demethylase.
    - ERG11 encodes the amino acid substitutions Y132F, K143R, and F126L or VF125AL

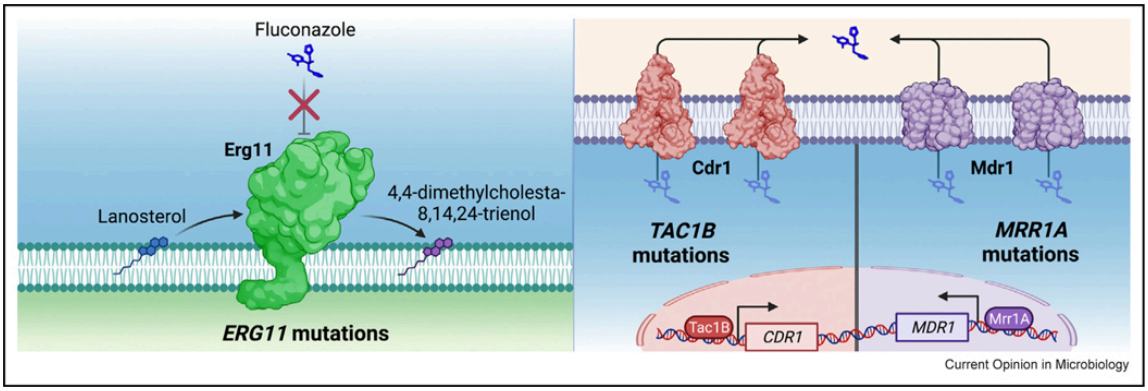


image link

→mutations in ERG11, TAC1B, and MRR1A contribute to clinical fluconazole resistance

- *C. auris* fluconazole resistance is caused by both ERG11 mutations and ERG11 overexpression (from extra gene copies)

Characterized mechanisms of triazole resistance and their associated impact on fluconazole MIC.

Triazole-resistance mechanism	Amino acid substitutions	Impact on fluconazole MIC if known
ERG11 mutations	–	–
Predominant	**VF125AL, Y132F, K143R**	8–16-fold increase
Less-commonly reported	L43H, Q357K, **F444L**, G459S, I466M, Y501H	Fourfold increase
TAC1B mutations	–	–
Predominant	**A640V**, A657V, F862_N866del	16-fold increase
Less-commonly reported	A15T, S192N, S195C, F214S, K247E, R495G, A583S, P595L/H, **S611P**, A651T, M653V	Fourfold increase
MRR1A mutations	–	–
Predominant	**N647T**	Fourfold increase

Mutations in bold have been characterized and impact on fluconazole MIC is shown in the adjacent column.

table link

Fold increase = how many times higher the MIC is compared to the baseline (wild type).
A higher fold increase means the fungus can tolerate much more drugs before being inhibited.

2025-09-15

Another method for *C. auris's* drug resistance--Biofilm formation

Drug Resistance in Candida Biofilms (link)

Biofilm resistance is stronger than planktonic resistance

- Planktonic (free-floating) cells use mechanisms like efflux pumps, ERG11 mutations, ERG3 loss-of-function, and FKS1 mutations.

- But when the same strains grow as biofilms, those mechanisms explain only part of the resistance. Biofilms add extra, unique resistance strategies.

Biofilm-specific resistance mechanisms:

- **Extracellular matrix (ECM):**
  - Makes up most of the protective barrier.
  - Composed of proteins (55%), carbohydrates (25%), lipids (15%), and extracellular DNA (5%)
  - Acts as both a physical barrier and a drug-sequestration system.
  - Glucans and mannans form complexes that trap antifungal drugs, preventing them from reaching targets.
    Fks1-driven glucan synthesis and glucanases (Bgl2, Xog1) are crucial for ECM assembly and drug sequestration.
  - This ECM protection is conserved across Candida species, including *C. auris*.
- Efflux pump upregulation (without drug exposure):
  - Biofilm cells increase expression of efflux pumps, pushing drugs out of the cells before they can act.
- Persister cells:
  A small subpopulation of metabolically dormant cells within biofilms that can survive high antifungal doses and repopulate after treatment.
- Biofilms have high cell density & quorum sensing→ enhance community-level tolerance.
- Stress response upregulation→ makes cells harder to kill.

Extracellular vesicles (EVs):

- Candida species release EVs during biofilm growth.
- These vesicles transport glucan and mannan to reinforce the ECM and help with drug resistance.
- Mutants that can't properly use EVs are more susceptible to antifungals in *C. auris* and other Candida species.

Immune evasion through ECM composition:

- Matrix mannans shield biofilms from neutrophil killing.
- Without mannans, biofilms become more vulnerable and trigger more neutrophil extracellular trap (NET) release.

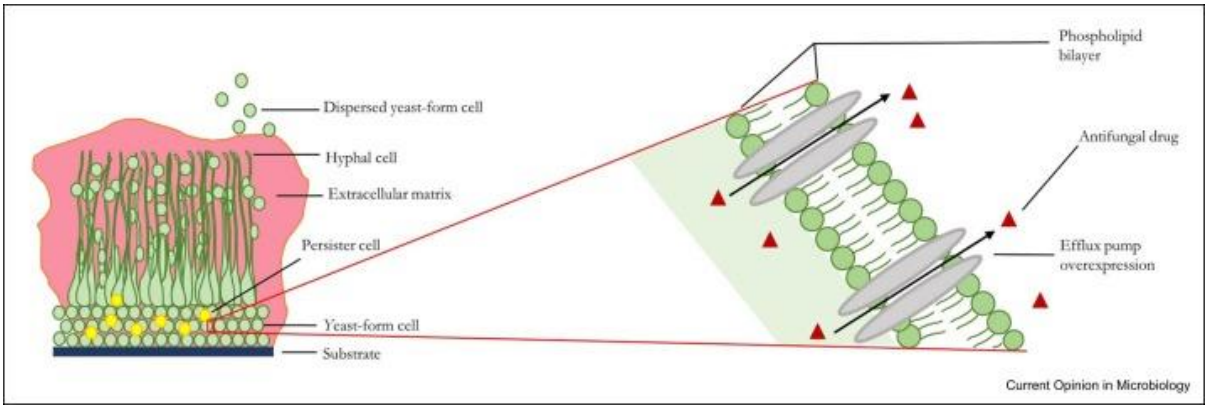


image link

LONG STORY SHORT→*Candida auris* biofilms resist antifungals and immune attack mainly by building a drug-trapping extracellular matrix (rich in glucans and mannans), upregulating efflux pumps, and maintaining persister cells. Extracellular vesicles help distribute matrix components, and the specific composition of the matrix

	also helps the fungus evade host immune defenses.
2025-09-15 **TESTABLE QUESTION DEVELOPMENT**	- I've done some background research into *Candida* and the ERG11 mutations still seem the most interesting to me. I got all those resources from ChatGPT (previously posted) that can help with modelling and drug docking. I was wanting to see if there was a way to reduce a fungus's resistance, because then antifungals would actually work and the mortality rate of *Candida* wouldn't be a whooping 40%. However, I don't actually know where I'd even start with that. Researching into ERG11 has made me wonder more about the actual drug docking process. Maybe there's a way to reverse it to how it was before a fungus's mutation? - Can we identify (in silico) secondary mutations that restore predicted fluconazole binding to a resistant ERG11 variant? - I like the above question, because it feels achievable with all my resources online. I did more research and these types of "reverse mutations" can be referred to as rescue mutations→secondary genetic changes that can suppress or reverse the harmful effects of a primary mutation, leading to a milder disease phenotype or restored function. (link) A Short Blurb On Rescue Mutations: ● Some mutations in ERG11 (like Y132F) make *Candida auris* resistant to fluconazole because they weaken drug binding in the active site. ● But those resistance mutations often come with a cost→ the enzyme might work less efficiently, or the pocket is slightly "broken" ● In evolution, sometimes a second mutation happens elsewhere in the protein that compensates — it "rescues" the protein's function, stability, or even drug binding. ● Scientists call these suppressor mutations. They don't necessarily put the protein fully back to wild-type, but they can partially restore lost interactions (or make the enzyme fold better, so the drug can still get in). ● **Wild-type ERG11 = fluconazole binds normally → sensitive.** ● **Mutant ERG11 (Y132F) = fluconazole binds poorly → resistant.** ● **Mutant+rescue (Y132F + X??) = maybe fluconazole binding is improved compared to plain Y132F → resistance is reduced.** - If I start with a resistant version of ERG11 (like Y132F), could adding a second, different mutation somewhere else in the protein partially reverse the resistance effect and restore drug binding affinity? - I like this version better, but I feel like it needs to be refined so it actually sounds more *ScIeNcY*. I also realize that I want to include the keywords: rescue mutation, fluconazole (duh), and the fact that I'm using computational docking and not creating ANOTHER global pandemic…and that I need to actually calculate the stability of the proteins so everything doesn't fall apart lol - Can a secondary "rescue" mutation in ERG11 restore predicted fluconazole binding affinity to a clinically resistant ERG11 variant (Y132F), as measured by computational docking and stability metrics? - I really really like this one!

2025-09-16

For the proposal it asks that I use AI to identify experimental errors, improvements, or obstacles

Link to chat: https://chatgpt.com/share/68ca231f-49b0-800d-b0c3-5f9f7485001c

⚠️ Potential Obstacles & Solutions

- **Obstacle: Lack of structural data for ERG11 with mutations.**
  → **Solution:** Use **in silico mutagenesis + MD minimization** to generate reliable structural models.
- **Obstacle: Interpreting docking scores meaningfully.**
  → **Solution:** Compare relative changes between WT, Y132F, and rescue mutants, not absolute docking energies.
- **Obstacle: Computational resource limitations.**
  → **Solution:**
  - Use free MD/docking resources (Google Colab with MD packages like GROMACS/AmberTools).
  - Focus on a small set of rescue mutations predicted from literature (not random scanning).
- **Obstacle: False positives in "rescued" binding.**
  → **Solution:** If possible, cross-validate docking with a second software (e.g., SwissDock, GOLD, or Schrödinger Glide) to confirm trends.

🔬 Crystal structures

- A **crystal structure** is the 3D atomic arrangement of a protein (or other molecule) determined experimentally, usually by **X-ray crystallography** (sometimes cryo-EM or NMR).
- Proteins are crystallized, X-rays are fired at them, and the diffraction pattern is used to reconstruct electron density → from which atom positions are mapped.
- These structures are deposited in the **Protein Data Bank (PDB)** and are considered the "gold standard" for computational modeling.
- Example: If someone has solved the structure of *Candida albicans* ERG11 bound to fluconazole, you could download it and directly use it for docking.

🧩 Homology models

- Many proteins (like *Candida auris* ERG11) don't have solved crystal structures.
- In that case, scientists build a **homology model**: a predicted structure based on a similar protein whose structure *is* known.
- Idea: Proteins with similar amino acid sequences usually fold into similar shapes.
- Programs (like SWISS-MODEL, Phyre2, AlphaFold) use known structures as templates to model your protein.
- Homology models are **less accurate** than crystal structures, especially if your protein has big insertions or low sequence identity with the template.

📉 Ramachandran plot

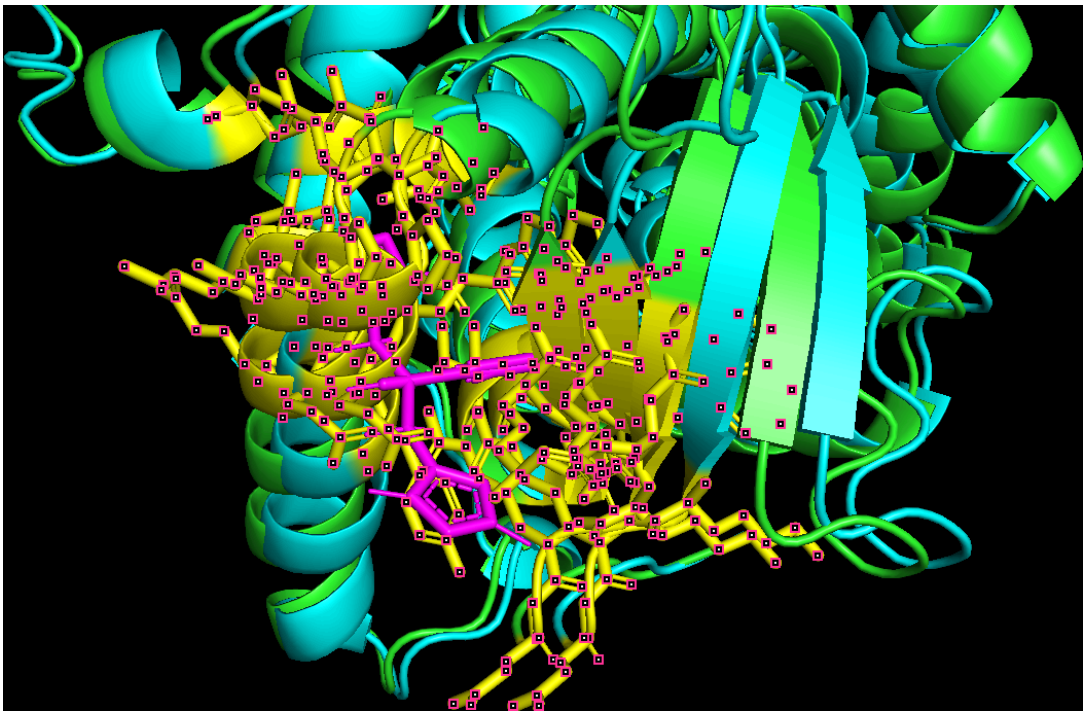
- A **Ramachandran plot** is a quality check for protein structures.
- It maps the **backbone dihedral angles** (ϕ and ψ) of each amino acid in the protein.
- Since proteins fold into specific shapes, only certain angle combinations are physically possible (because of steric clashes).
- The plot shows allowed vs. disallowed regions:
  - **Favored regions**: where most amino acids should fall (like α-helices and β-sheets).
  - **Outliers**: residues in forbidden regions → may mean errors in the model or unrealistic geometry.
- When you build or mutate a protein structure, you check the Ramachandran plot to see if the geometry makes sense.

2025-09-22 **STEP ONE!** https://chatgpt.com/share/68d4a526-1c68-800d-bf37-6e47f42803e9	NCBI Protein Database • I will use the NCBI Protein Database to actually download the protein sequences for the wild-type (resistant to fluconazole) ERG11, and the Y132 variant (resistant one) • I was totally unsure how to use this software so I used ChatGPT to give an actual outline **Candida auris ERG11 Wild-type PROTEIN:** >UNE56009.1 Erg11 [Candidoza auris] MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPALGKGVIYDCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFKMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSL MTNSTYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYNEVLSVLAEKGGSLKD LAYDDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRAN EFDPHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRF KKDGSLPPVDYQSMVTLPMEPAEIEWEKRETCVY • ISSUE→ Hit a major roadblock→ the NCBI genbank doesn't have the Y132F variant…. So I have to figure sth out.. • A FASTA file is legit just a text file with fasta at the end. So I'm gonna try and download another file as a text file with fasta…? **Candida auris ERG11 Y132F Variant:** >UNE56009.1 Erg11 [Candidoza auris] MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPALGKGVIFDCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFKMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSL MTNSTYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYNEVLSVLAEKGGSLKD LAYDDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRAN EFDPHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRF KKDGSLPPVDYQSMVTLPMEPAEIEWEKRETCVY • GUYS IT WORKED • Ok so now I need to use Collab Fold to start docking stuff https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=kOblAo-xetgx
2025-09-29 Link to tutorial video: https://www.yo	• The CollabFold software was incredibly confusing so I found a youtube video to help me • In the video, they used UniProt for their sequences→ This is another open-access database with protein sequences and another information • I noticed that for CollabFold, i don't even need the downloaded FASTA files, only the letter sequence with

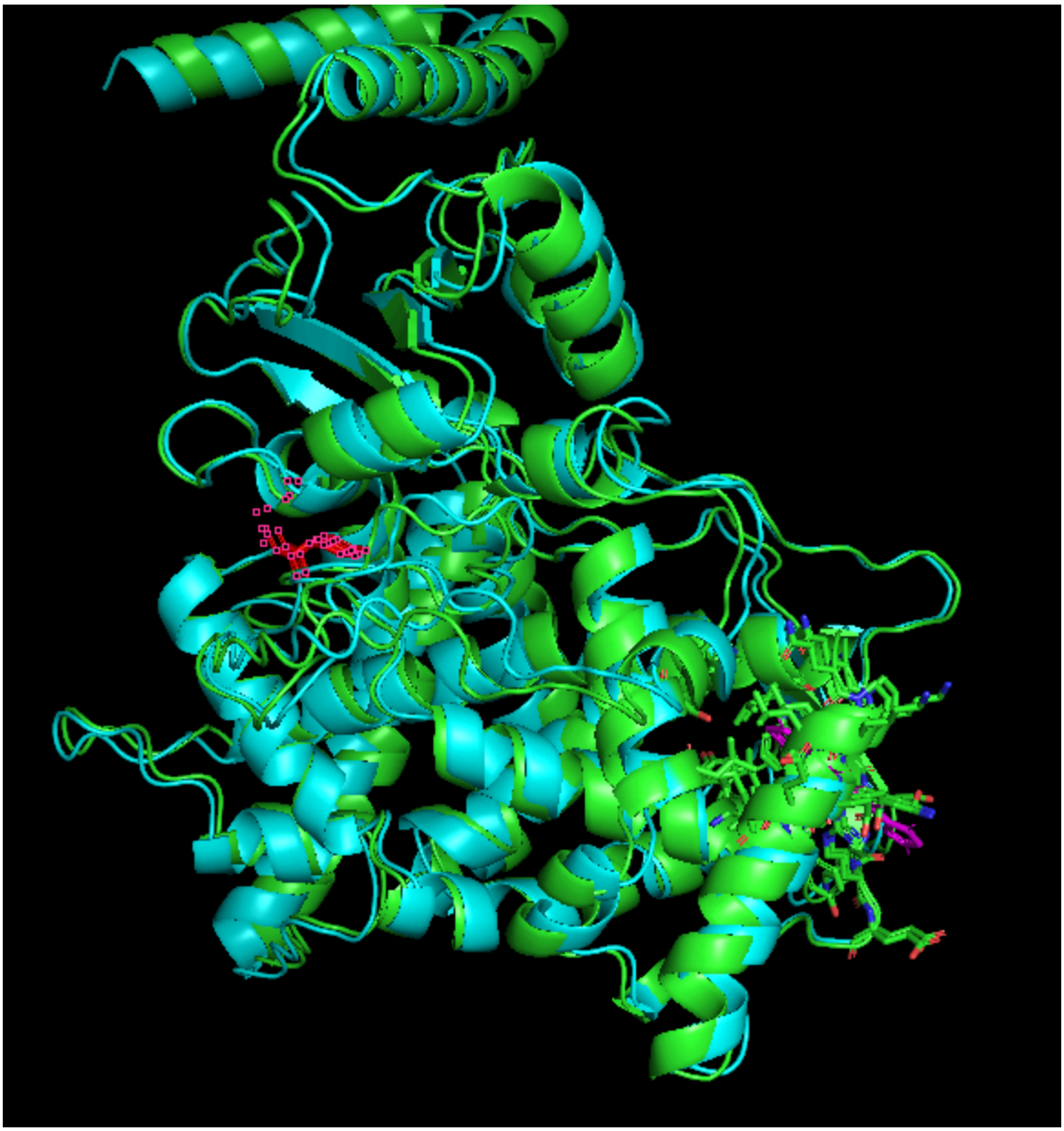
utube.com/watch?v=eLy7PdzRgLs	no spaces - So for this first run, I am just testing the known Y132F variant, just to see how it works and stuff. - I need to use PyMOL to make a 3D model of it and actually see what's happening
2025-09-30 Procedure Development (checkpoint #2) and safety plan https://chatgpt.com/share/68db3b16-b090-800d-ab8f-d139ffda438e	**Procedure:** 1) Use UniProt softwares to obtain both the Wild-Type and variant DNA 2) Use CollabFold AlphaFold 2 to predict the 3D structures of both 3) Use the PyMOL software to visualize the 3D models and see the active sites 4) Repeat steps 1-3 with a new, secondary mutation 5) Observe the effects on the fungus and record data 6) Use a Ramachadran plot to measure the stability **[possible further steps]** 7) Research the usage of mycoviruses (viruses that infect fungi) to try and *theoretically* find a way to administer this secondary mutation. **mycoviruses don't naturally spread *well* between fungal cells, so I'd need a way to make them transmissible OR 8) Antifungal development with the help of Radiance Compounding Pharmacy **Safety Plan:** **Scope** - This project is entirely in silico. No live fungi, viruses, or pathogenic materials will be handled. - All data is sourced from public sequence databases (NCBI, UniProt, PDB). **Software Safety** - Use trusted, official sources (NCBI, UniProt, ColabFold, PyMOL). - Keep software updated to avoid security risks. - Store data on secure, password-protected devices. **Data Ethics** - Only use publicly available sequences and structures. - Cite all databases and tools properly. - Do not attempt to engineer or distribute modified genetic material. **Biosafety** - No wet-lab experiments with fungi, viruses, or DNA will be performed. - Any discussion of mycoviruses is limited to literature review and theory only. - Clearly state in your report that experimental validation would require a BSL-2/3 facility, trained professionals, and institutional approval.
2025-10-23 Today I aim to figure out the PyMOL software. What i have: Variant structure, WT structure, fluconazole downloaded Tutorial videos i followed:	Right now I have the candida auris wild type and y132f variant files downloaded from collabfold onto my computer. I need to find out how to use pymol to actually visualize the docking sites and take data so that when I make changes I can compare. What I want to look for: 1. The binding site for fluconazole→ since this'll change between the WT, variant, and secondary mutation 2. Stability 3. Structural differences

https://www.youtube.com/watch?v=IvLI9ECRMPI

https://www.youtube.com/watch?v=mBlMI82JRfI&list=PLUMhYZpMLta1_Z7to3by2ATHP-cI4ma5X



GREEN→ WT
BLUE→ Y132F Variant
Pink thing→ fluconazole
YELLOW→ amino acid residues within 5 Å of the fluconazole molecule (the binding pocket). part of the protein that interacts/ could interact) with fluconazole. This is the functional site where inhibition or mutation affects matter most.



Red thing on the left is the actual mutation relative to the fluconazole (right purple-pink thing)

WHAT AM I LOOKING FOR

Primary (most important)

1. Docking score (predicted binding affinity) — numeric (kcal·mol^{-1}).
   → Shows whether predicted binding gets stronger (more negative) with rescue.

2. ΔΔG of folding (stability change) — numeric (kcal·mol^{-1}).
   → Ensures rescue isn't strongly destabilizing the enzyme.

Secondary Stuff

3. Number of hydrogen bonds between ligand and protein in best pose (integer).
→ More H-bonds often = stronger binding.

My next step:

- Measure the distances between fluconazole and the *same binding site residue (132)* in both models — the WT (ERG11_WT) and the Y132F_Variant — so I can compare how the mutation affects binding proximity
  - Shorter distances (2–3 Å) between the drug and active-site atoms generally mean stronger interactions and higher binding affinity.
  - Longer distances (4–6 Å or more) suggest weaker binding — often what happens in resistant variants like Y132F.

2025-10-24

Goal and purpose:
using distance measurements to figure out how close certain atoms are between:

- The side chain of residue 132, and
- The atoms of fluconazole in the binding site.

WHY

- Short distances (abt 2–4 Å) suggest strong hydrogen bonds or hydrophobic contacts.
- Longer distances (abt 6 Å) suggest weak or no interaction.

So, if in the (Y132) structure, the *OH group* of tyrosine is close (abt 3.2 Å) to fluconazole→ means a strong interaction.
But if in the (F132) structure, that distance increases (abt 6.5 Å) — the mutation has disrupted that interaction, explaining the drug resistance.
→quantify how the mutation changes fluconazole's position or strength of binding.

Measuring the distance **PURPOSE→** If the distance between these atoms gets larger in the Y132F variant, it suggests that the mutation weakened fluconazole binding (the atoms drift farther apart).
**when testing rescue mutation→ if the distance becomes shorter again, that would suggest partial restoration of binding affinity.

RESULTS MEANING→ The OH of Tyr132 in the wild type is about 38.0–40.8 Å away from fluconazole atoms

ISSUE…..--> Distances of ~38–42 Å are *huge* on the molecular scale. Typical interacting atoms (hydrogen bonds, van der Waals contacts) are in the ~2.5–4.0 Å range. AlphaFold predicted protein structure, not ligand placement… so basically these readings aren't even helpful. In the current PyMOL model, the fluconazole molecule is not positioned in the enzyme's active site near Tyr132. it's sitting far away. **I need to use a different software like SwissDock or AutoDock Vina…**

2025-10-25

PyMOL Data and Current Progress:

STEPS COMPLETED	OBSERVATIONS	CURRENT UNDERSTANDING

• Loaded and visualized ERG11_WT, ERG11_variant (Y132F), and fluconazole in PyMOL.
• Highlighted the Tyr132 residue (WT) and its OH group, the site involved in fluconazole binding.
• Identified all relevant atomic indices for Tyr132 and fluconazole atoms.
• Measured distances between Tyr132's OH group (index 1051) and key atoms on fluconazole:
  ◦ To fluconazole O (index 20): ~39.0 Å
  ◦ To fluconazole N
	• The large distances indicate no close interaction between the protein and ligand in the current model.
	• Confirms that the AlphaFold structures are ligand-free and require molecular docking to simulate true binding positions.
	• Successfully determined reference atom positions and residue sites for future binding or docking analyses.
		• Tyr132 (WT) is located near the fluconazole binding pocket in *ERG11*, but in the un-docked structure, it does not form direct contact.
		• Mutation at Tyr132 (to Phe) likely alters binding affinity by removing the hydroxyl group that can hydrogen bond with fluconazole.
		• I have the atom indices, residue identification, and measurement protocol established for both WT and variant

	(index 1): ~41.7 Å		

Next steps: do the docking on Swiss Dock and collect data after modelling.
Tutorial used: https://www.youtube.com/watch?v=gztIfgMALuY&t=40s
https://www.youtube.com/watch?v=Z4iw-63yW4A

2025-10-26

Understand swissdock!

Notes:

- So swissdock actually requires **NO** coding… meaning that all those crashouts over PyMOL were for naught (jk I got some pretty good data from there)

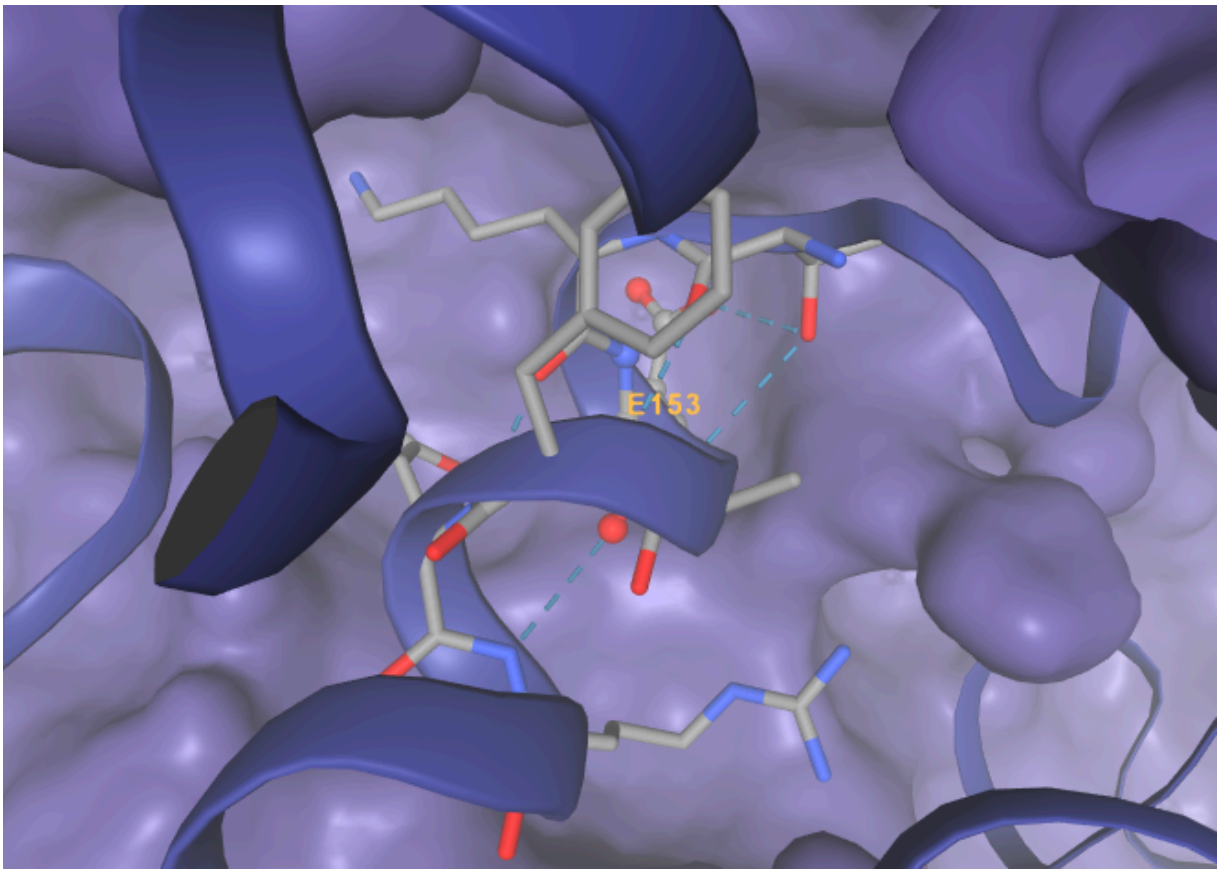
- Must clean up "hetero atoms" → atoms that can hinder the process. Like water molecules, preoccupied ligans, etc. Need the "clean" protein to be docked
- Download structure of protein compound (so my fluconazole) in " canonical SMILES" format→ found on PubChem. This is what swiss dock uses. **I paste the SMILES code in the "Ligand" section in SwissDock**

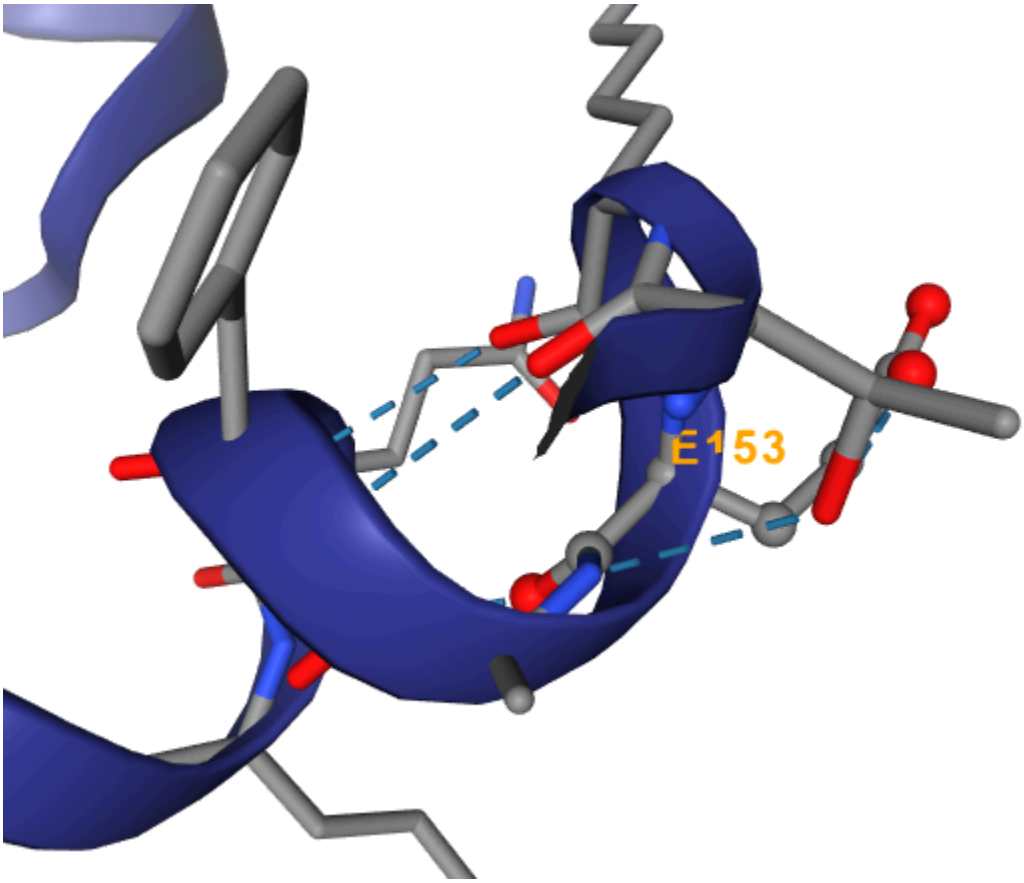
Fluconazole in SMILES format→ OC(Cn1cncn1)(Cn1cncn1)c1ccc(F)cc1F

- Use Protein Database (PDB) to get the PDB id of target (So my candida WT and variant)
  - Issue→ PDB had multiple candida auris WT models…
  - Solution→ I used the downloaded WT file i already had from the PyMOL (looks like all that wasn't for nothing :))
- More negative→ better affinity score since it runs based on energy

RESULTS FOR WT

Cluster number	Cluster member	AC Score	SwissParam Score
0	1	-32.116355	-7.6802
1	1	-31.212586	-7.5077
2	1	-29.228661	-7.3784
3	1	-29.110482	-7.3028
4	1	-29.079224	-7.4344
5	1	-28.467575	-7.4495
6	1	-28.368782	-7.3037
7	1	-28.088412	-7.7408
8	1	-27.887821	-7.4528
9	1	-27.875340	-7.9076





- Cluster→ A grouping of similar docked poses (binding conformations)
- Cluster member→ number of poses in that cluster (1 = only one pose)
- AC Score→ "FullFitness" or approximate binding free energy (kcal/mol). More negative = stronger predicted binding. **BINDING AFFINITY SCORE**
- SwissParam Score→ Internal parameterization score, usually used to evaluate molecular energy; smaller

(more negative) is more stable.

My data meaning:

- The best docking pose has an AC Score of -32.116 kcal/mol (Cluster 0)
- This is the strongest predicted binding among your poses.
- The range (-32 to -27 kcal/mol) shows that fluconazole binds fairly tightly in all poses→ which is very very predictable, since this is the WT. I can already predict that this will be significantly different for the Variant

Now I will do the same for the Y132F Variant, and then compare results!

RESULTS FOR VARIANT

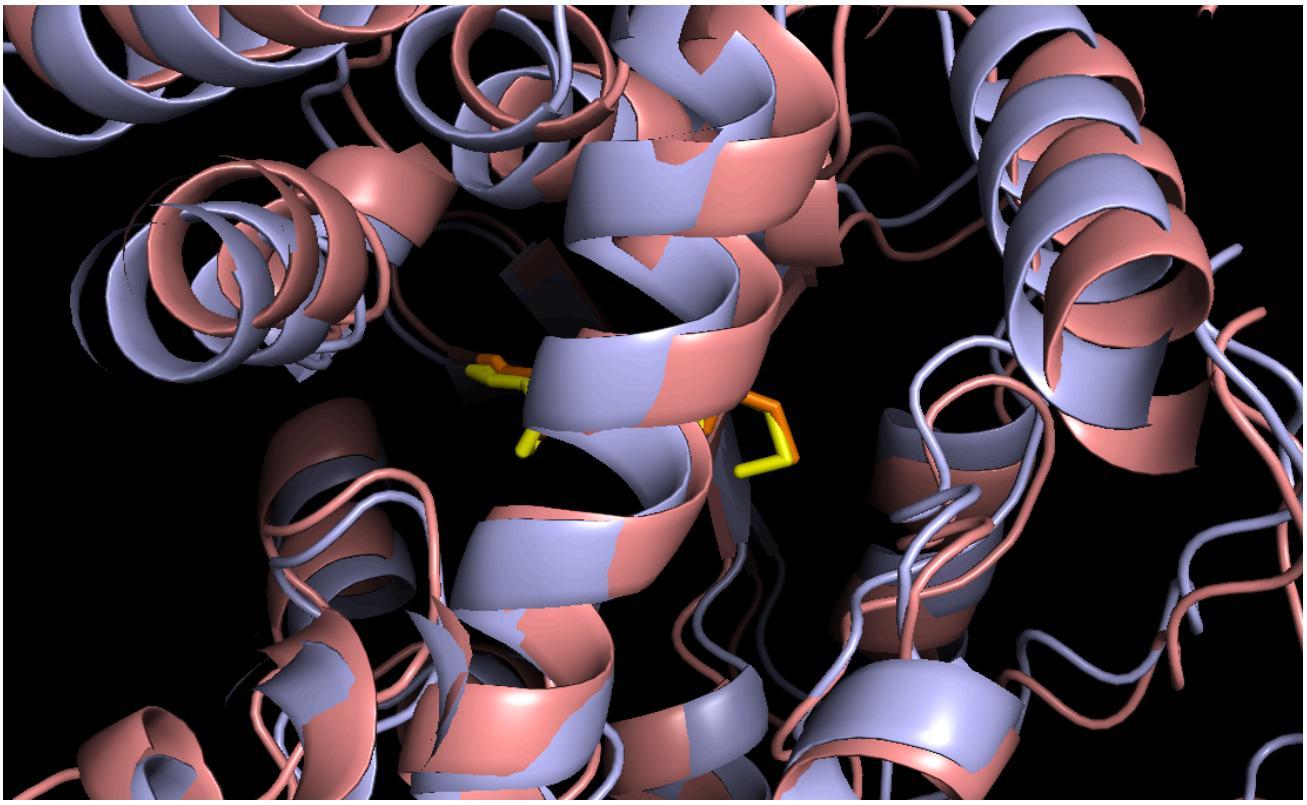
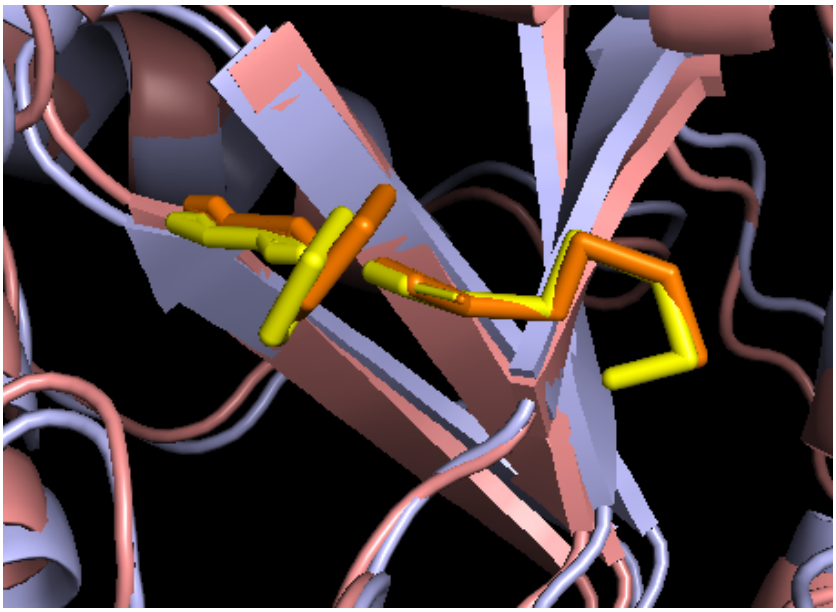
Cluster number	Cluster member	AC Score	SwissParam Score
0	1	-32.909148	-7.7914
1	1	-32.352840	-7.6757
2	1	-31.717886	-7.3859
3	1	-31.214959	-7.8617
4	1	-31.208539	-7.3902
5	1	-30.623716	-7.5994
6	1	-29.656282	-7.3286
7	1	-29.364178	-7.6848
8	1	-29.099384	-7.4578
9	1	-29.027533	-7.5809

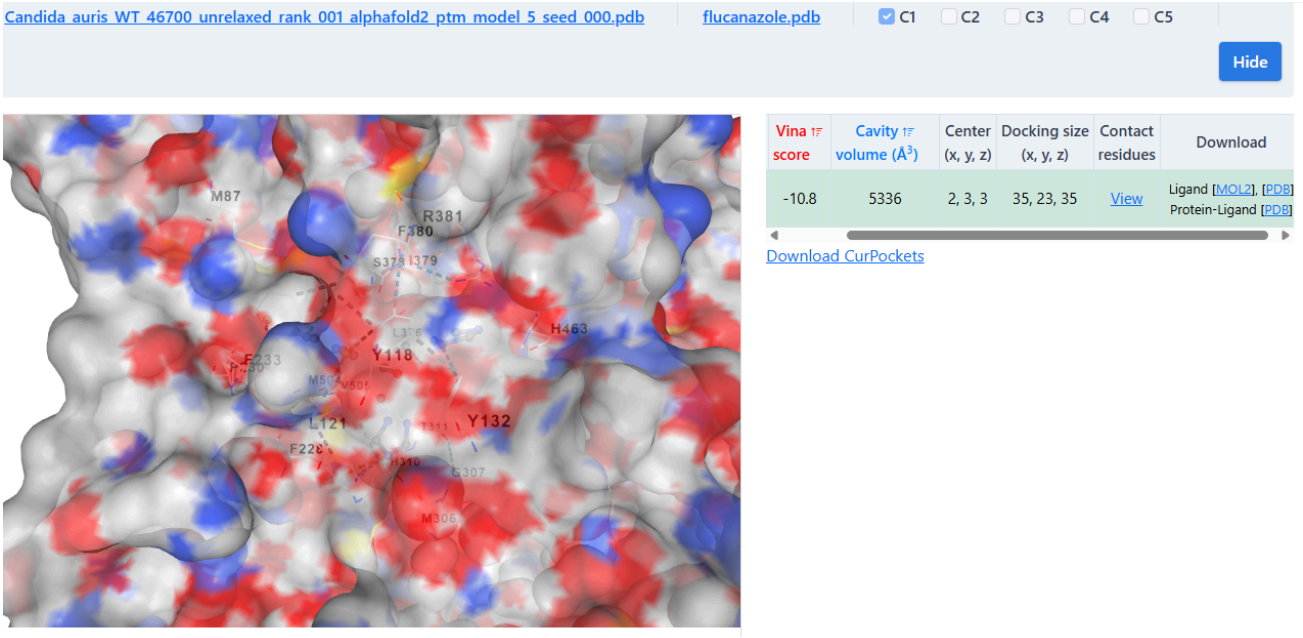
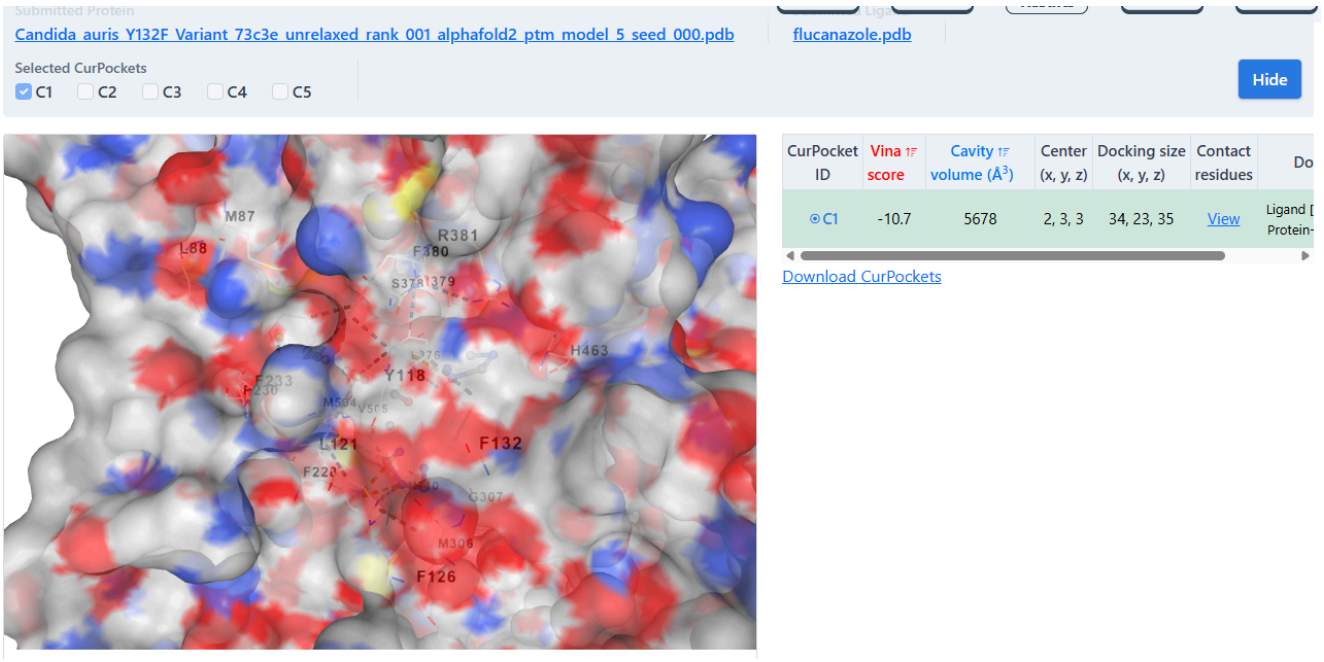
ISSUE→ I couldn't find the fluconazole docking at all on the model
SOLUTION→ will use PyMOL again (yay) to get a correct "box size" to narrow into the search so that SwissDock can be more precise. I'll use my code to find an accurate xyz parameter "box center".

Cluster number	Cluster member	AC Score	SwissParam Score
0	1	-31.714820	-7.3382
1	1	-31.363759	-7.2270
2	1	-30.536442	-7.4106
3	1	-30.032848	-7.3457
4	1	-29.848995	-7.3399
5	1	-29.766900	-7.5848
6	1	-29.356989	-7.4506
7	1	-29.084845	-7.2941
8	1	-29.024551	-7.4116
9	1	-28.684106	-7.2834

- I faced the same issue… the docked fluconazole is nowhere to be seen! I can't just use this data since it isn't accurate. The AC score (binding affinity) should be decreasing significantly due to the mutation.
- I will do it one last time with a different window, and if this doesn't work I will switch to CB-Dock2

Tutorial used:
https://www.youtube.com/watch?v=i3cdgCpCmvU

2025-11-08	• Retrying SwissDock with a larger box size (30-30-30) • Parameters are the same as the second run: (x = -11.711, y = 10.125, z = -4.594) ○ It didn't work again • Onto CB-Dock2 now.. I will basically redo what I did in swissdock but here
2025-11-10	 WT→ blueish purple Variant→ salmon WT docking→ yellow Variant docking→ orange  The fluconazole were rotated→ same binding pocket, but the variant is shifted and slightly rotated This can: • Weaken the binding affinity (less optimal orientation for hydrogen bonds or van der Waals forces) • Alter how the drug fits into the pocket (possibly causing resistance)

	• Change the accessibility of nearby residues (some contacts may be lost or new ones formed)
2025-11-12	**WT Results from CB-Dock 2**  • The VINA SCORE (in red) is about the binding affinity/ strength. For the WT, its -10.8 **Variant Results from CB-Dock 2**  • For Variant, Vina score is -10.7 - The Vina docking score represents the predicted binding affinity between the ligand (fluconazole) and the protein — a lower (more negative) score indicates stronger binding. - WT has a slightly better (more negative) docking score than variant→ This suggests that the mutation slightly reduces fluconazole's binding affinity, meaning the drug binds a bit less tightly to the variant protein. This was very predictable, and pretty much expected Next steps→ use PLIP to visualize and list the hydrogen bonds/ contacts and compare
https://chatgpt.com/share/69175fbc-1170-800d-bfa7-bad64	

9d92bd5	
2025-11-13	**PLIP Results!** **WT** Hydrophobic Interactions ···· 	Index	Residue	AA	Distance	Ligand Atom	Protein Atom
---	---	---	---	---	---
1	118A	TYR	3.76	26	965
2	233A	PHE	3.60	27	1888
3	380A	PHE	3.94	27	3055	 Hydrogen Bonds — 	Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
---	---	---	---	---	---	---	---	---	---
1	118A	TYR	3.34	4.02	125.52	×	✓	14 [N3]	968 [O3]
2	132A	TYR	2.62	3.03	106.97	✓	✓	1075 [O3]	14 [N3]
3	310A	HIS	2.54	3.14	119.00	✓	✓	2496 [Npl]	12 [N3]	 - 3 hydrophobic interactions (Tyr118, Phe233, Phe380) - 3 hydrogen bonds, including one directly involving Y132→, helps stabilize fluconazole in the binding pocket. - These multiple H-bonds and close hydrophobic contacts indicate a strong, well-anchored binding pose. - Overall, the WT shows a solid interaction network that supports tight fluconazole binding. **VARIANT** Hydrophobic Interactions ···· 	Index	Residue	AA	Distance	Ligand Atom	Protein Atom
---	---	---	---	---	---
1	118A	TYR	3.80	24	965
2	121A	LEU	3.90	26	992
3	233A	PHE	3.56	27	1887
4	380A	PHE	3.82	27	3054	 Hydrogen Bonds — 	Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
---	---	---	---	---	---	---	---	---	---
1	310A	HIS	2.21	3.02	138.73	✓	✓	2495 [Nar]	12 [N3]
2	381A	ARG	2.86	3.79	157.44	✓	✓	3065 [Ng+]	15 [N3] | π-Stacking ···· ···· | Index	Residue	AA	Distance	Angle	Offset	Stacking Type	Ligand Atoms
---	---	---	---	---	---	---	---
1	228A	PHE	4.27	24.12	0.82	P	10, 11, 12, 17, 18, 19	 π-Cation Interactions ···· 	Index	Residue	AA	Distance	Offset	Protein charged?	Ligand Group	Ligand Atoms
---	---	---	---	---	---	---	---
1	228A	PHE	4.73	1.37	×	tertamine	12	 - Still hydrophobic contacts, but the pattern changes: now have 4 hydrophobic interactions and only 2 hydrogen bonds, and Y132-related hydrogen bonding is lost because phenylalanine cannot donate H-bonds. - The variant additionally forms π-stacking and π-cation interactions, but these are weaker or less directional than the original Y132 hydrogen bond.

	○ Pi stacking: produce relatively strong and orientation sensitive bonding sometimes called “pancake bonding”. These interactions are particularly well suited for transferring electrons from molecule to molecule ○ Pi cation: result from the attraction of the charge borne by a positive center towards the quadrupole created by the π-electron cloud of an aromatic side chain ● Overall, the mutation reduces the specific stabilizing interactions (especially H-bonding), which explains the weaker binding seen in the docking score.
2025-11-14 Understand the differences between the tables	**CONCRETE DIFFERENCES** 1. Hydrogen bonds (directional, strong interactions) ● WT: 3 H-bonds detected. *One is Y132 → ligand* with H–A = 2.62 Å and D–A = 3.03 Å (PLIP labels: residue 132A TYR). ● Variant: 2 H-bonds detected. *The Y132 H-bond is absent* (Y → F removes the OH donor). The variant’s H-bonds are to HIS310 (H–A = 2.21 Å) and ARG381 (H–A = 2.86 Å, D–A = 3.79 Å) ● Why it matters: losing the Y132 H-bond (2.62 Å) removes a short, directional contact that anchored the ligand in WT. The variant’s remaining bonds are fewer and one is geometrically longer (ARG381 D–A = 3.79 Å), which is weaker for hydrogen bonding. 2. Hydrophobic contacts (shape/packing) ● WT: Hydrophobic contacts with TYR118, PHE233, PHE380 (distances ~3.6–3.9 Å). ● Variant: Hydrophobic contacts with TYR118, LEU121, PHE233, PHE380 (distances ~ 3.56 -- 3.90 Å) — one extra hydrophobic contact (LEU121) appears in the variant. ● Why it matters: the variant’s ligand re-packs against hydrophobic side chains (gaining LEU121), indicating the ligand **shifted** to make more van der Waals contacts→ these interactions help stabilize binding but are less specific than H-bonds. 3. Aromatic / π interactions (weaker, less directional) ● WT: No π-stacking/π-cation reported in the table you provided. ● Variant: π-stacking (PHE228, 4.27 Å) and a π-cation interaction (PHE228, 4.73 Å) are present. ● Why it matters: the variant’s ligand rotates to exploit aromatic surfaces (PHE228), forming π interactions that partially compensate energy-wise but are less directional and less able to substitute for the lost Y132 H-bond when it comes to specific inhibitory action. **MEANING** ● In the wild type, fluconazole is held by three stabilizing hydrogen bonds including a short, directional Y132→ligand contact (H–A = 2.62 Å). In the Y132F variant that H-bond is lost; the ligand instead makes fewer hydrogen bonds (and one is geometrically weaker) while shifting to increase hydrophobic and π interactions (an extra hydrophobic contact with LEU121 and π-stacking/π-cation at PHE228). Together these changes explain why the variant shows weaker, less specific binding: the ligand is reoriented to rely on non-directional contacts that cannot fully replace the lost Y132 hydrogen bond. ● PLIP analysis shows WT ERG11 forms a key Y132 hydrogen bond to fluconazole (H–A = 2.62 Å), which is absent in the Y132F variant; the variant instead gains extra hydrophobic and π contacts (e.g., LEU121, PHE228), indicating the ligand is reoriented and stabilized by weaker, non-directional interactions — consistent with reduced effective binding despite similar raw docking scores

2025-11-17

Start developing secondary mutations

Brainstorming

- The WT has a short, directional hydrogen bond that anchors the ligand. A successful rescue mechanism would re-introduce a polar group (an H-bond donor or acceptor) into the binding pocket at a way that can approach the same ligand atom at abt 2.6–3.2 Å
- PLIP analysis shows the Y132F variant loses a directional hydrogen bond (WT: Y132→ligand, H–A = 2.62 Å), and the ligand re-orients to increase hydrophobic and π contacts (e.g., LEU121, PHE228). A successful rescue would restore a short (≈2.6–3.2 Å), directional polar contact to the same ligand heteroatom, thereby returning the ligand toward the WT geometry
    - Since it needs to be within 2.6–3.2 Å, I'll change the 5 letters on either side beside the Y132F variant

What Each Letter Means in the Protein Sequence (link)

Letter	Amino Acid Name	Properties
A	Alanine	Small, non polar
R	Arginine	Positively charged (basic)
N	Asparagine	Polar, uncharged
D	Aspartic acid	Negatively charged (acidic)
C	Cysteine	Polar. Can form disulfide bonds
E	Glutamic acid	Negatively charged (acidic)
Q	Glutamine	Polar, uncharged
G	Glycine	Very small, flexible
H	Histidine	Positive/neutral (can switch), often catalytic, aromatic
I	Isoleucine	Hydrophobic
L	Leucine	Hydrophobic
K	Lysine	Positively charged (basic)
M	Methionine	Hydrophobic, sulfur-containing
F	Phenylalanine	Aromatic, hydrophobic
P	Proline	Rigid, causes kinks in protein chain
S	Serine	Polar, can form H-bonds
T	Threonine	Polar, can form H-bonds
W	Tryptophan	Large aromatic, hydrophobic
Y	Tyrosine	Aromatic, polar OH group
V	Valine	Hydrophobic

	Aromatic amino acids **Y132F Variant Sequence Again For Reference:** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● The highlighted area (the F) is the variant. It used to be a Y, but mutated to an F ○ Y→ Tyrosine→ aromatic, polar OH group ○ F→ Phenylalanine→ aromatic, hydrophobic ● So the mutation removed that OH group that was Y before, which removes that contact **CANDIDATE ONE** ● I will change the I (Isoleucine) to Y (Tyrosine) to re-add that OH group and have that contact for fluconazole ● Mutation→ I131Y (the first character in mutation is capital i, not 1) **Candidate 1→ I131Y** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVYFDCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY
2025-11-27	● PLAN: Use collabfold to dock it and download pdb file, upload to CB Dock and dock fluconazole, and then put into PyMOL and combine, and then go to PLIP ○ Alphacollab predicts the 3D protein structure from the amino acids The Vina Score was -10.5, which is less than the WT (-10.8), so that didn't work… **CANDIDATE TWO** ● Change the N (Asparagine) to C (Cysteine) since both are polar, and C can form disulfide bonds. Maybe this can replace the lost OH bond. **Candidate 2→ N132C** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPCSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● Binding score was -10.2…

	• It's just getting worse with my changes lmao
2025-11-28	**CANDIDATE THREE** • I noticed that the Tyrosine change had that OH group, which is the main reason that binding is weaker in the variant. So if i input Serine maybe it'll help since serine can form H bonds • Q (Glutamine) is polar, just like S (Serine). The only difference is that while Q is uncharged, Serine can form H bonds and is *also* uncharged. Really crossing my fingers for this one… **Candidate 3→Q142S** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM ESKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY • Binding score -10.3… so just slightly better than the last but still bad. • Gonna try changing it to T since it's also polar and can form h bonds **CANDIDATE FOUR** **Candidate 4→Q142T** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM ETKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY • -10.4…
2025-11-29	**CANDIDATE FIVE** • I noticed that both Y and F ((the two that were switched in the initial variant) were both aromatic. So is W (tryptophan) **Candidate 5→V130W** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGWIFDCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY • -10.6! We're getting there :D **CANDIDATE SIX** • I (Isoleucine) is right beside the mutated F, and is hydrophobic just like W. both also large and nonpolar. **Candidate 6→ I131W** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVWFDCPNSRL MEQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDM RARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMT NSTYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLA

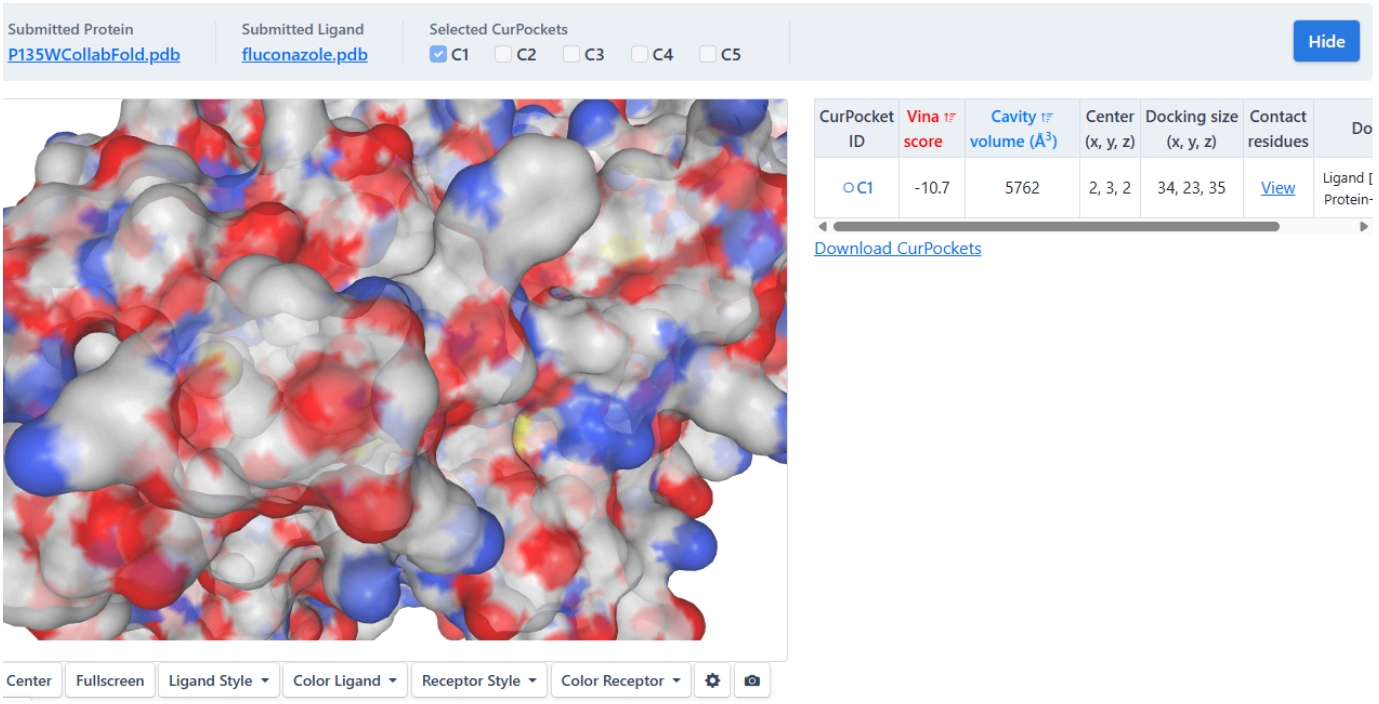
YDDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEF
DPHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKD
GSLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.5… it went down again

Candidate 7→ P135W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCWNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.7! Technically back to the beginning lmao…



- I'm going to download the complex and compare in PLIP! See what i did
  - Remember to first combine in pyMOL as a complex

Hydrogen Bonds

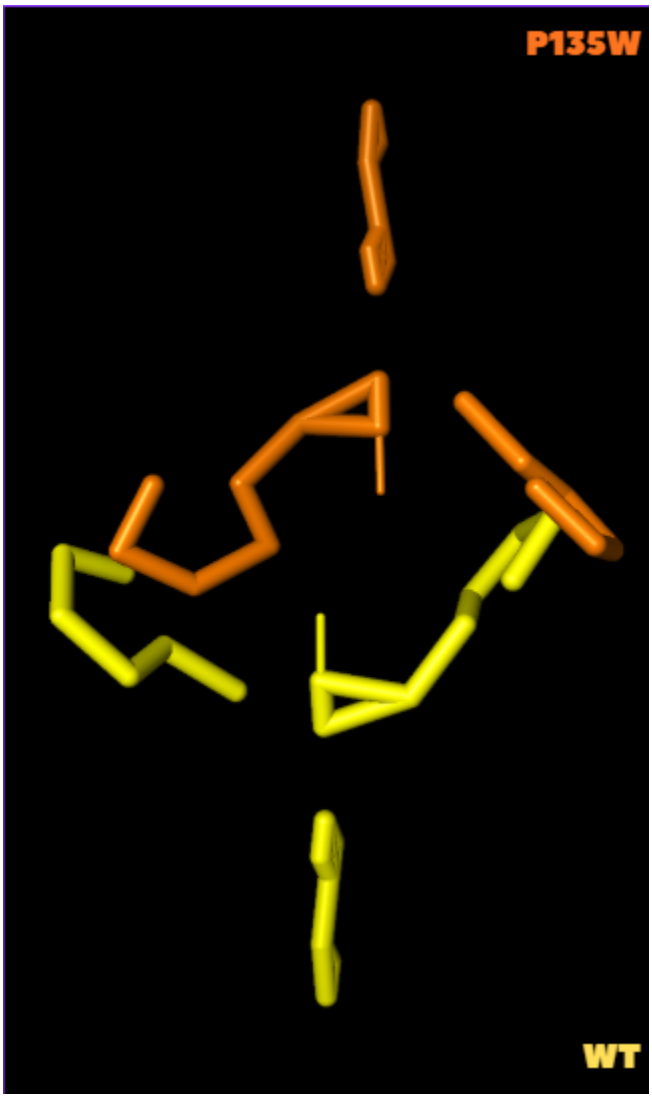
Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	310A	HIS	2.73	3.59	146.56	✓	✓	2502 [Nar]	15 [N3]
2	381A	ARG	2.57	3.33	134.10	✓	✓	3071 [Ng+]	12 [N3]
3	381A	ARG	2.23	3.07	143.38	✓	✓	3072 [Ng+]	12 [N3]
4	463A	HIS	2.95	3.62	123.94	×	×	11 [N3]	3718 [O2]

Halogen Bonds

Index	Residue	AA	Distance	Donor Angle	Acceptor Angle	Donor Atom	Acceptor Atom
1	142A	GLN	3.02	148.83	94.64	31 [F]	1156 [O2]

Halogen bonds are even stronger!!

- This is actually a big breakthrough and a huge step---> halogen bonds are stronger than h bonds. So this one

	actually made that connection whereas the others didn't.  I uploaded the WT and P135W into pymol to visualize the difference… and for some reason it flipped 180 degrees..?? ● Fluconazole is binding in a non-native pose ● The docking algorithm accepts it ● But the ligand is under strain ● The protein pocket is deformed Docking penalizes that strain → score stalls at −10.7.
2025-12-27→ 2026-01-01 (happy new years lol)	**Candidate 8→ N136C** ● Cysteine can form disulfide bonds MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPCSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● -10.2… **Candidate 9→ Y118W**

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAWSHLTTPVFGKGVIFDCPNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.7 again lol

Candidate 10→ C134W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDWPNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.7 again omg

Candidate 11→ R381W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFWKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -9.8..this is the worst one so far!

Candidate 12→ R381S

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFSKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.3

Candidate 13→ R381T

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ
FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLM
EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR
ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN
STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY
DDLQKMPLINQTIKETLRLHMPLHSIFTKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD
PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG
SLPPVDYQSMVTLPMEPAEIEWEKRETCVY

- -10.0

Candidate 14→ D133W

	MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFWCPNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● -10.0 again Collecting data: ● I found that using aromatic amino acids were successful for the most part (F,Y,W,H) ○ P135W made halogen bonds ■ Proline (p) is a non-essential amino acid ■ Nonpolar and rigid. ■ No backbone h-bond donor ■ Induces Kinks & Turns: Its rigidity introduces bends and turns in protein secondary structures, crucial for loops and turns ○ P is also aromatic, i wanna see what happens if i do P135H, instead of W **Candidate 15→ P135H** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCHNSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● -10.7 again.
2026-01-22	**What all the -10.7s have in common** ● I'm noticing i'm always getting sooo close to -10.8 but never truly there ● Im realizing that those mutations restore local contacts, but they do NOT restore the original pocket geometry ○ Docking scores come from Shape complementarity, Pocket pre-organization, Ligand strain energy… not just raw interaction count→ i've been compensating energetically, but not geometrically ■ **More bonds ≠ better binding if the ligand has to distort to make them** ● I'm realizing that secondary mutations rarely restore *full* WT affinity… they stabilize or partially compensate. ● Secondary mutations that increase hydrophobic or aromatic contacts can partially compensate for resistance, but cannot restore wild-type binding unless they also restore native ligand orientation. ● The WT had perfectly tuned interactions and geometry, which I realize will be close to impossible to recreate with a secondary mutation… it was a mix of perfect pocket volume, Aromatic spacing, Helix geometry, and Ligand orientation. I need to not strain the actual backbone of the structure, and see if i can bring back the original ligand orientation

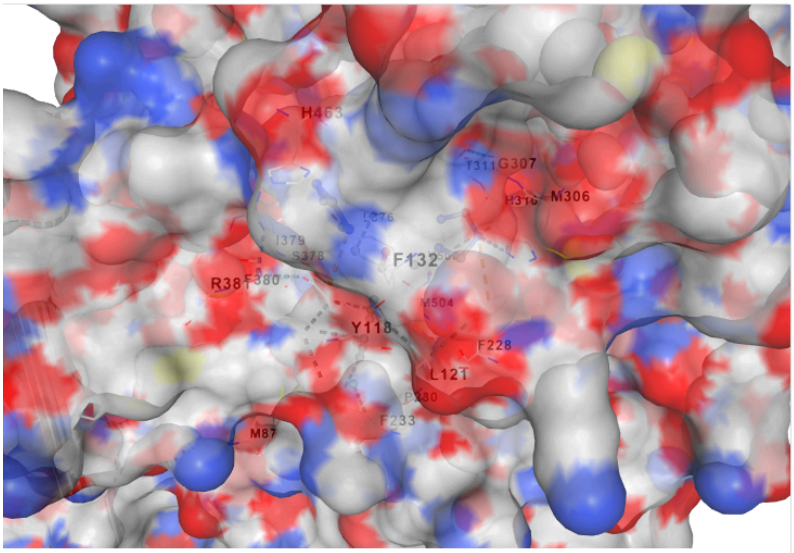
	SOLID FINDINGS: - Resistance alters binding geometry - Secondary mutations can increase contacts - But cannot fully restore native affinity - And may even force non-physiological binding poses So if I want to continue with this almost gamble, I'm realizing that the amino acids I'm using are just large, and require the structure to compensate for its size. On the contrary, Histidine (H) is a relatively small aromatic amino acid…although I literally did try that for my last one. ● Histidine is rigid and side chain driven, which might be why this didn't work ○ **helps binding contacts but does not restore backbone geometry or pocket pre-organization** So then I need to find a way to get a secondary mutation that fixes this by reducing steric constraint, allowing backbone relaxation, and removing side-chain dominance → basically a mutation that *removes structure* rather than adds it. OR A mutation that does not contact the ligand at all, but stabilizes the helix or loop that positions the pocket.
2026-01-24 Trying to find secondary mutations once again	Small Amino Acids that can help relax the backbone and fix geometry ● A→ Alanine→ small an non-polar ● G→Glycine→ small and flexible. Smallest amino acid. ● S→Serine ● T→ Threonine **Candidate 16→ P135A** MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQ FFESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCANSRLM EQKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMR ARFDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTN STYKDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAY DDLQKMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFD PHRWDEETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDG SLPPVDYQSMVTLPMEPAEIEWEKRETCVY ● **AHHHH I DID IT IT WORKED**

Submitted Protein
P135A_d78fd_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000.pdb

Submitted Ligand
fluconazole.pdb

Selected CurPockets
C1 C2 C3 C4 C5

Hide



Center | Fullscreen | Ligand Style | Color Ligand | Receptor Style | Color Receptor

CurPocket ID	Vina score	Cavity volume (Å³)	Center (x, y, z)	Docking size (x, y, z)	Contact residues	Do
C1	-10.8	5022	3, 3, 3	35, 23, 35	View	Ligand Protein

Download CurPockets

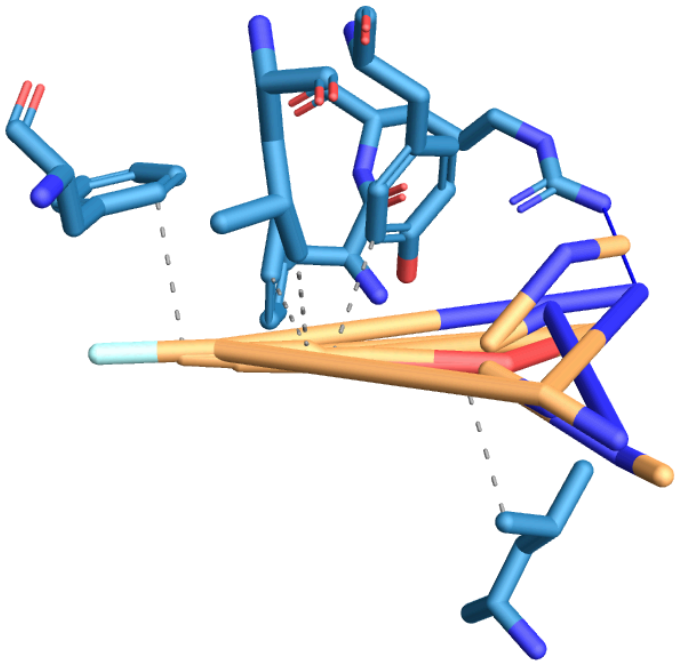
- I GOT IT TO -10.8
- Removing proline rigidity restored pocket pre-organization
  - Ligand no longer needs to flip or strain
  - Docking crossed the energetic barrier without adding interactions

Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	118A	TYR	3.96	24	965
2	121A	LEU	3.72	26	992
3	233A	PHE	3.83	27	1885
4	376A	LEU	3.91	16	3017
5	380A	PHE	3.70	26	3052

Hydrogen Bonds

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	381A	ARG	Distance between hydrogen and acceptor atoms	✓	✓	3063 [Ng+]	15 [N3]		



- PLIP analysis of the P135A rescue mutant shows a distributed network of hydrophobic interactions and a single directional hydrogen bond that stabilize fluconazole in a relaxed orientation, indicating restoration of binding-pocket geometry rather than introduction of new ligand contacts
    - Orange→ fluconazole
    - Blue→ protein side chain residues
    - Dashed lines→ protein interactions (mostly those hydrophobic ones)

I want to understand why this actually worked..

The previous -10.7 ones…

- Added energy
- Reduced entropy
- Forced pose
- Created non-physiological binding

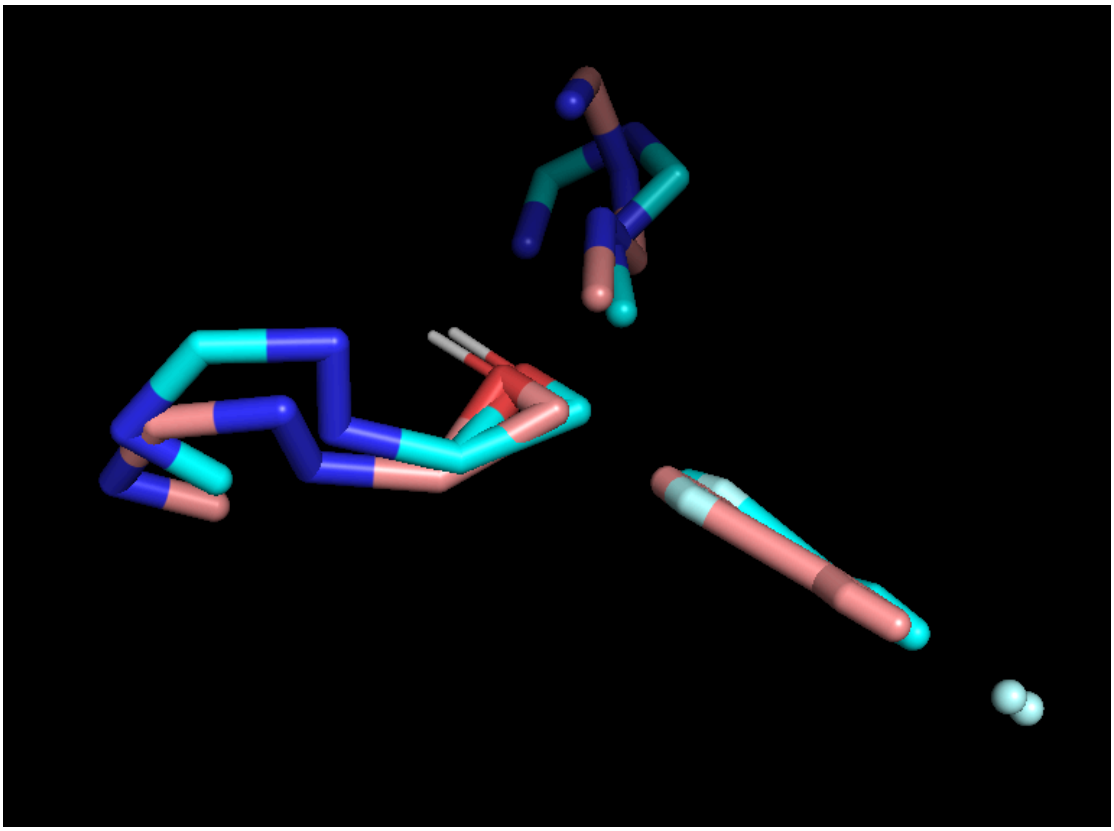
This one…

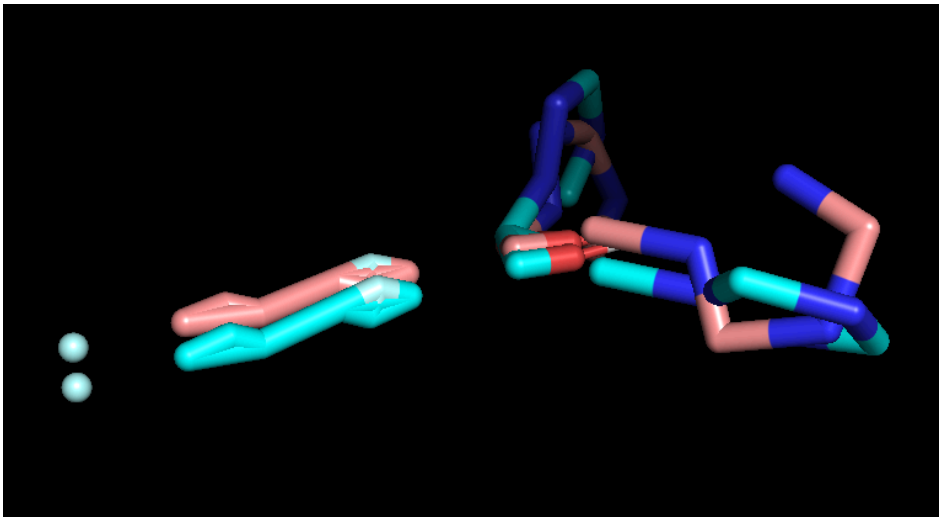
- Removed constraint
- Restored backbone freedom
- Allowed native-like ligand orientation
- Reduced ligand strain
- Preserved pocket volume

Docking scores are global free-energy estimates.

I've officially discovered that Resistance is a geometric problem, not a contact problem

And that Restoring WT affinity does not require restoring WT interactions





WT is pink, P135A is blue

CONCLUSIONS: The resistance mutation altered pocket geometry rather than eliminating interactions.
A secondary mutation (P135A) restored binding affinity by relieving backbone constraint and re-symmetrizing the binding cavity, without introducing new ligand contacts

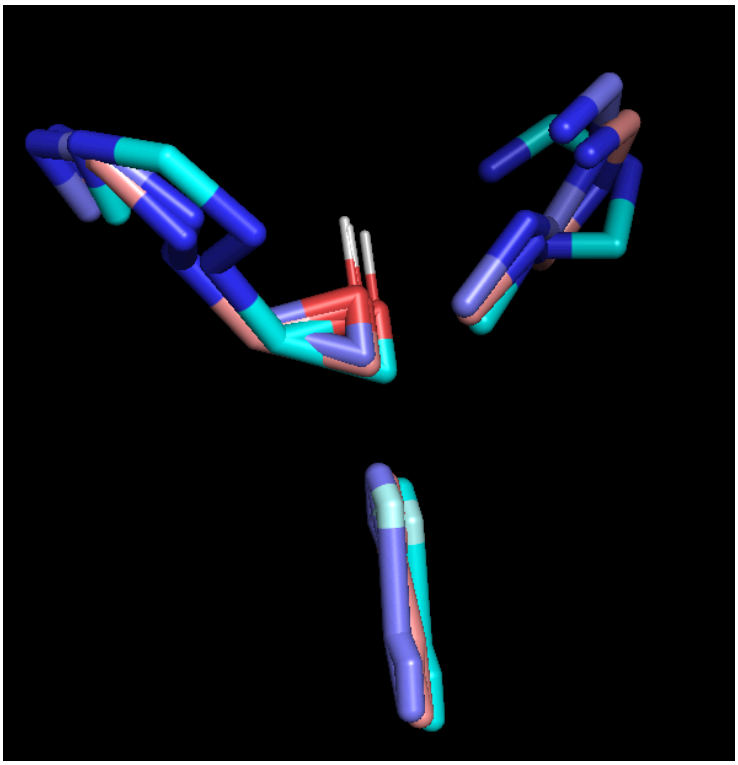
Previous hypothesis: A secondary rescue mutation will partially restore predicted fluconazole binding to the Y132F variant by stabilizing the protein structure and binding pocket

- And my findings actually support this!

Results:

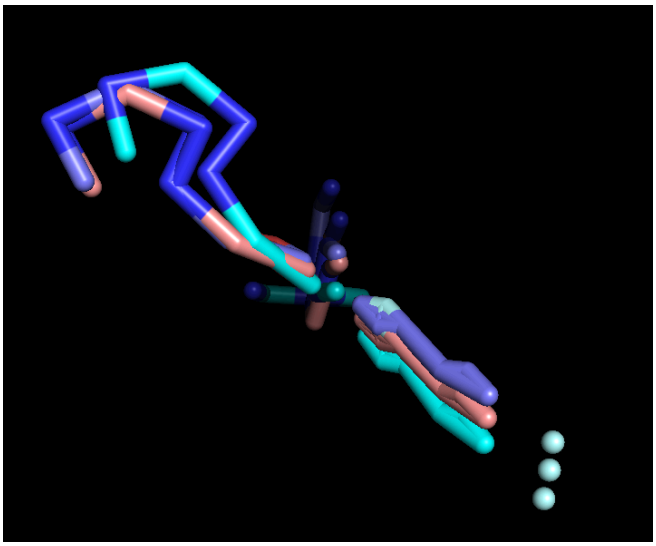
- Partial restoration confirmed
  - Y132F alone → resistant geometry, reduced docking
  - Y132F + P135A → docking score improves past the −10.7 ceiling
  - Binding is not identical to the WT but is clearly "rescued"
- Structural stabilization strongly confirmed
  - In PLIP:
    - Fewer interactions
    - Proper distances
    - No extraneous and over the top H-bond interactions
  - In PyMOL:
    - No ligand 180 flipping or constrain
  - Residue behavior:
    - P135A doesn't directly contact fluconazole, but provides rescue by helping the protein's geometry
      - **by <u>removing</u> destabilizing constrain**
- P135A compensates for Y132F's distorted geometry→ not only did P135A stabilize the pocket, it actually restored the correct energetic landscape for binding

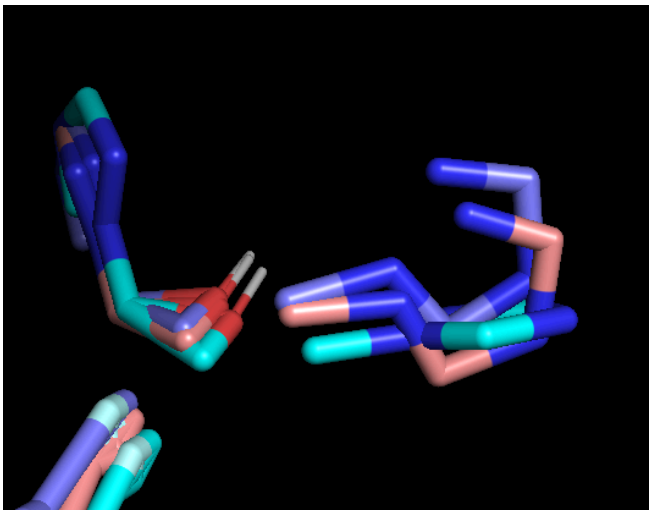
MY findings support the hypothesis. The P135A secondary mutation partially rescues fluconazole binding in the Y132F background by restoring binding-pocket geometry and pre-organization, rather than by introducing new ligand interactions.

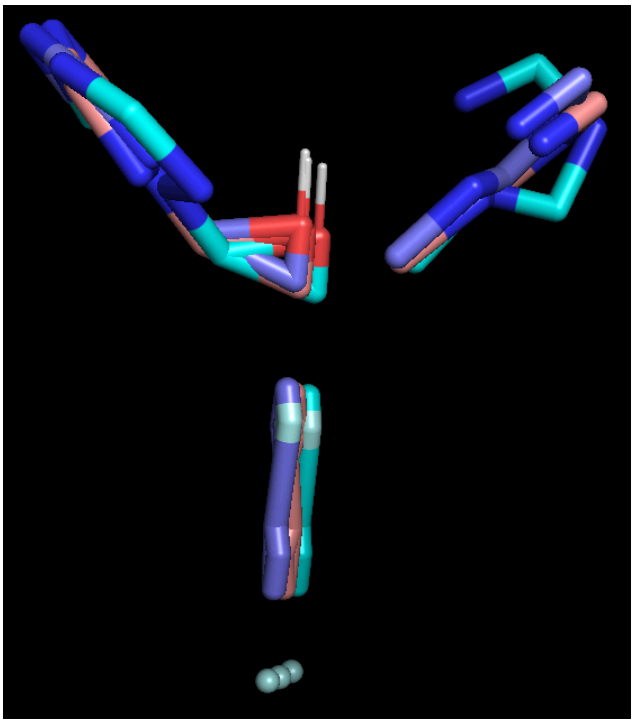


Pink→ WT
Purple→ Y132F
Blue→ P135A

- Im noticing that the WT and Y132F kind of mirrored each other, but a little rotated
- The P135A did the same for the most part, but on that right side top part, it's totally different



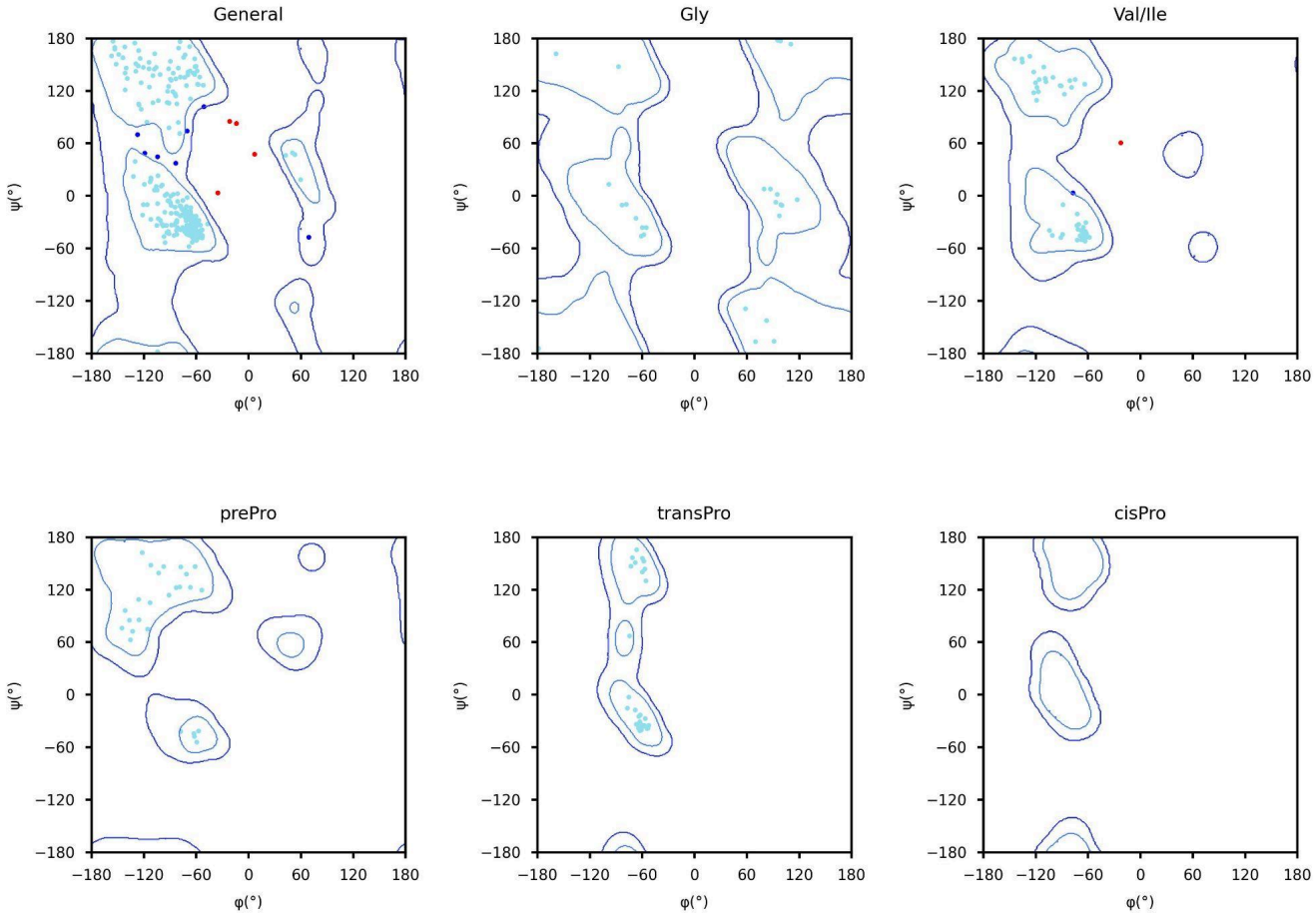




2026-01-25

Testing for stability using Ramachadran plot (www.ramplot.in)

- Basically what a Ramachadran plot is trying to say is “Does this protein fold in a realistic, stable way?”



Color key

- **Light blue dots** → residues in allowed/favored regions
- **Blue contour lines** → statistically preferred backbone angles
- **Red dots (if present)** → outliers / strained conformations

1) General Graph:

- This means most residues

Dense clusters in:

- α-helix region ($\varphi \approx -60$, $\psi \approx -45$)
- β-sheet region ($\varphi \approx -120$, $\psi \approx 120$)
- Very few outliers
    - The protein backbone is well-behaved
    - No global unfolding
    - No major modeling artifacts

      This supports structural stability

2) Gly

Glycine is super flexible, so this plot looks more spread out.

- Points are still inside allowed contours
- No wild, isolated outliers

Meaning:

- Local flexibility exists (normal)
- No unrealistic torsion angles

3) Val/Ile

These are bulky residues and *very sensitive* to steric clashes.

- Tight clustering
- Almost everything inside favored zones

suggests:

The protein core remains well-packed despite mutation nearby.

4) prePro / transPro / cisPro

These are special cases that often cause problems.

Key observation:

- Very few points
- All in expected regions

Meaning:

- No weird peptide bond distortions
- No forced backbone flips

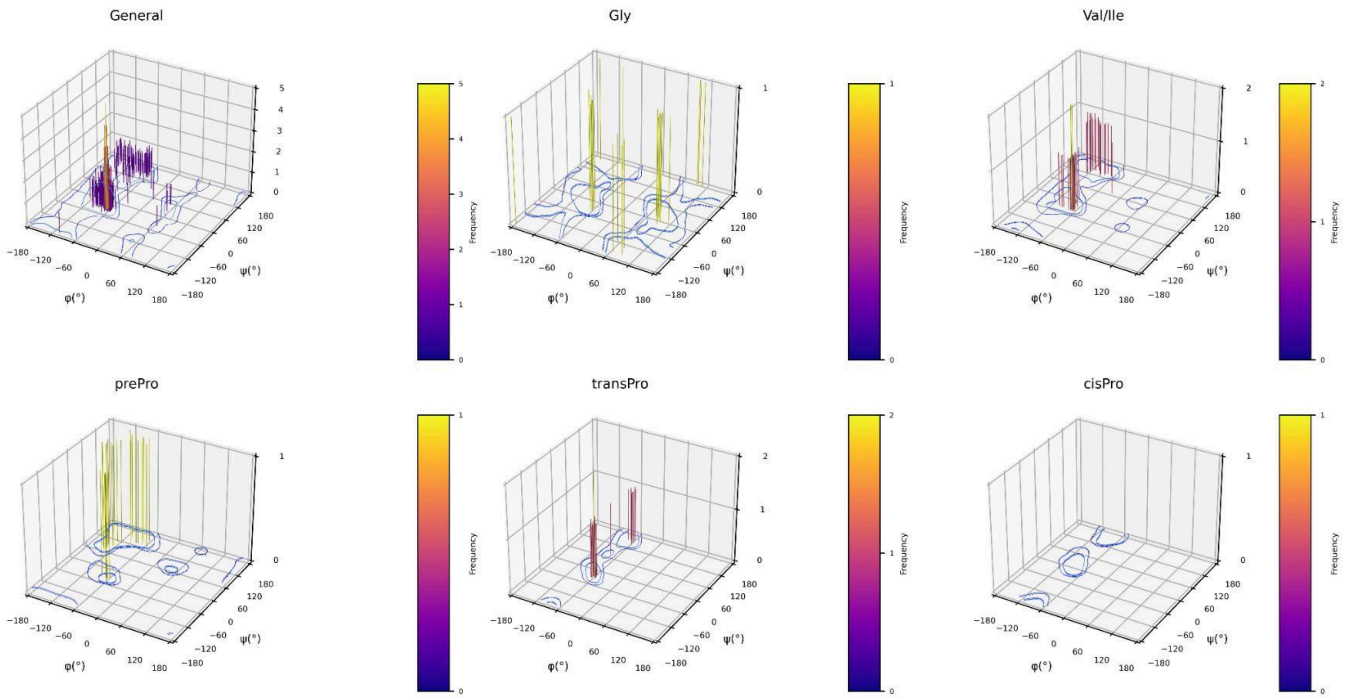
THE BIG PICTURE

- No new strain introduced
- No compensatory distortions elsewhere
- Clean backbone geometry

P135A is a “quiet” mutation — it adjusts local packing without breaking the fold.

Ramachandran analysis supports that P135A is structurally tolerated and does not destabilize the protein, consistent with a compensatory rescue mutation

Ramachandran analysis showed that the P135A mutant maintains favorable backbone conformations, indicating that the rescue mutation improves binding without compromising overall protein stability

	Resistance is geometric, not energetic — and removing structure can rescue binding better than adding interactions 
2026-01-25	**Understanding** ● Previously, I was trying to fix the docking score by forcing more interactions, but I realized that more interactions doesn't necessarily mean stronger docking. Especially after i saw P135W and saw how yes it had more contacts, but it forced fluconazole to flip 180 degrees in order to fit ○ That made me realize that maybe interactions were not the issue, but the geometry→ more strain on how it formed made it weak in the end ● Why the -10.7 ones were close but not good enough was because those amino acids were large and bulky, and overall just rigid. They did provide interactions, but forced the ligands to bond badly. Energetically, they worked. However, geometrically they all ultimately failed. This was when i realized that more contacts isn't the solution ● I decided to hone into the P135 chain. I noticed that P is known for rigidity and causes kinks in the chain. ○ Since it's rigid, I thought that maybe once Y132F happened, the structure wasn't able to fix its geometry since it didn't have flexibility for it. It was almost incompatible with the mutation. That's most likely why my other mutations failed→ i was replacing a dominant rigid amino acid with another dominant residue ● So i realized i needed a change that pretty much removes structure, instead of adding to it ○ Aromatics add volume, interactions, and pretty much all failed ○ Glycine is very flexible, but is also the smallest amino acid. This one felt very dramatic, but I kept it in my back pocket. ○ S/T→ I had already tried a few times, and they didn't really work ○ Alanine→ minimal side chain, no strong chemistry, also very small. Is often compared to glycine, since it's just a bit bigger. Its compact, non-reactive nature makes it ideal for fitting into tight spaces in protein folding, such as within alpha-helices and beta-sheets→ already sounded good. When it comes to Mutagenesis, it is used frequently in alanine scanning because its small size causes minimal disruption to the protein structure while removing the side chain functionality of the original residue. Its small, non-bulky, and chemically inert, allowing for the assessment of the side

	chain's importance without disrupting the backbone conformation Initial secondary mutations focused on increasing local interactions near the Y132F resistance site, which improved docking scores but consistently plateaued due to ligand strain and pocket deformation. This suggested that resistance arose primarily from altered binding geometry rather than loss of contacts. Proline 135 was identified as a rigid backbone constraint adjacent to the resistance site. Substitution with alanine removed this constraint without introducing new ligand interactions, allowing the binding pocket to relax and re-symmetrize. As a result, fluconazole adopted a lower-strain binding orientation, leading to partial restoration of predicted binding affinity **Solidifying Real-World Connection** ● What if we didn't need a brand-new antifungal—what if we could *fix resistance* by understanding how a single amino acid reshapes drug binding? ● Some deadly fungal infections are becoming resistant to our most common antifungal drug, fluconazole. Instead of giving up on the drug and waiting decades for new ones, my research shows that small structural changes in the fungal protein can partially restore how the drug binds. This suggests that resistance isn't always permanent—and that understanding protein structure could help extend the life of existing medicines ● **Why This Matters: Fighting Resistance Without New Drugs"** ● **"Extending the Life of Antifungal Drugs"** ● **"Can Resistance Be Reversed?** **Why This Matters** ● Could help guide **drug redesign** rather than replacement ● May extend the usefulness of existing antifungals ● Especially important in low-resource settings where alternatives don't exist Instead of asking how to kill resistant fungi, my research asks whether we can make existing drugs work again by understanding protein structure **Practical Applications/ Future Research** ● If a small alanine can restore binding, a drug can be redesigned to do the same thing. mutation becomes a structural clue for drug modification→ help redesign drugs to make them efficient, and without putting out more risk for AMR ○ Resistance can be structurally modulated, not just bypassed ● My research provides a structural explanation for how antifungal resistance may be partially reversed by secondary mutations. This information can guide experimental testing, drug redesign, and resistance prediction, helping extend the effectiveness of existing antifungal treatments. This work helps understand *how* resistance changes drug binding, which is essential for designing better treatments Future work could experimentally validate the predicted rescue mutation by introducing P135A into fluconazole-resistant strains and measuring changes in drug susceptibility. If confirmed, this approach could inform structure-guided strategies to design antifungals that exploit compensatory weaknesses in resistant enzymes
2026-02-13	Filling out the CYSF website and working on write-ups

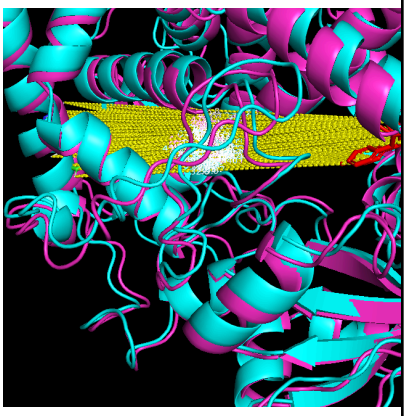
Definitions:

- Clade→ a group of organisms that includes a single common ancestor and all of its descendants

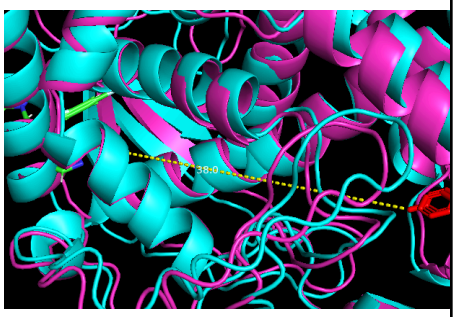
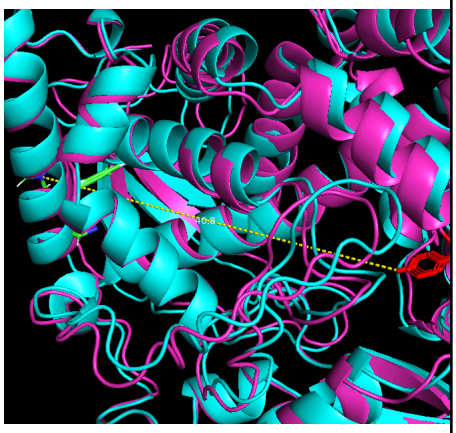
- Nosocomial→ Originating or taking place in a hospital, acquired in a hospital, especially in reference to an infection
- Filamentation→ a bacterial or yeast cell elongating without dividing, often as a stress response to environmental factors or antibiotics
- Thermotolerance→ an organism's ability to withstand high temperatures
- Osmotolerance→ resistance to sugars and salt, or only salt
- Sterol→ a type of steroid and a lipid with a multi ringed structure that functions as a key component of cell membranes in all eukaryotic organisms, regulating fluidity and permeability
- ERG11→ name of a gene found in fungi, including Candida. Encode enzymes needed to build ergosterol.
- Lanosterol Demethylase→ an enzyme which removes a methyl group (-CH3) from the molecule lanosterol during the biosynthesis of ergosterol.
- Ergosterol→ essential for fungal cell membranes. Without it, the membrane is unstable and the fungus cant survive.
- Affinity→how strongly two molecules stick to each other→how tightly fluconazole (the antifungal drug) binds to ERG11 (the fungal enzyme)
- Protein Sequence→ the specific, unique linear order of amino acids linked together by peptide bonds to form a protein molecule, also known as its primary structure
- Å→1 Å = 1 angstrom = 0.1 nanometers = 1×10^{-10} meters→ unit of length used in molecular biology and chemistry to measure distances between atoms.
- OH→ oxygen-hydrogen bond. The OH group is polar, meaning it can form hydrogen bonds with other molecules like fluconazole. These hydrogen bonds are *crucial* for fluconazole to stay bound inside the enzyme's active site
- Tyr132→ Tyr is short for Tyrosine, an amino acid (one of the building blocks of proteins). 132 is its position in the ERG11 enzyme's amino acid chain→ So "Tyr132" means: *the tyrosine residue at position 132 of the ERG11 protein.*
- Canonical SMILES format→ a unique, standardized string representation of a molecule that describes its connectivity and chirality, ensuring that any given molecule always has the same string representation regardless of how it is drawn
- Van der Waales forces→ weak intermolecular forces of attraction between atoms or molecules that are close to one another
- Aromatic→ exceptionally stable organic compounds that have a cyclic structure with delocalized pi electrons
- Quadrupole→ a configuration of four poles (like positive and negative charges or magnets) arranged to create a specific electric or magnetic field

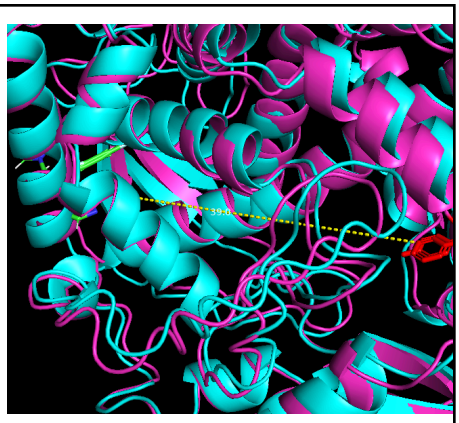
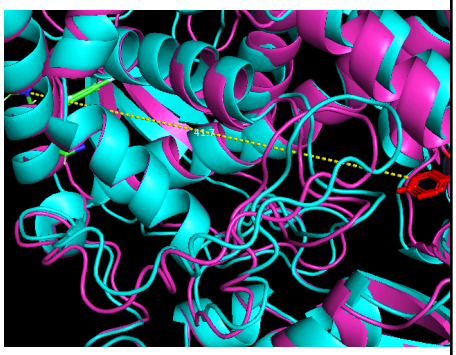
Codes:

CODE	MEANING/ PURPOSE	RESULTS/ NOTES/ SCREENSHOTS

select mut, resi 132 show sticks, mut color red, mut	• Creates a selection called mut for residue 132 (the mutation site Y132F) • Shows it as sticks • Colors it red so I can see it in both WT and variant	
distance wt_contact, Candida_auris_WT and resi 132, fluconazole	• Measures the shortest distance between any atom in residue 132 of the wild type and any atom in fluconazole • Creates a visual dashed line and a label showing the distance (in Å) - Dotted yellow lines connect red residue to fluconazole - White numbers in Å showing measured distance	
iterate (Candida_auris_WT and resi 132), print("WT_atom:", name, "resi:", resi, "chain:", chain, "index:", index, "coords:", x, y, z)	• looped through every atom in residue 132 (the mutation at position 132) of *Candida auris* protein structure and printed its details. Meanings: - CE1 → that atom's name (part of the tyrosine side chain) -resi: 132 → it belongs to residue number 132 -chain: A → in chain A of your protein -index: 1048 → PyMOL's internal atom ID (used in selection)	**Output**: WT_atom: N resi: 132 chain: A index: 1040 WT_atom: CA resi: 132 chain: A index: 1041 WT_atom: C resi: 132 chain: A index: 1042 WT_atom: O resi: 132 chain: A index: 1043 WT_atom: CB resi: 132 chain: A index: 1044 WT_atom: CG resi: 132 chain: A index: 1045 WT_atom: CD1 resi: 132 chain: A index: 1046 WT_atom: CD2 resi: 132 chain: A index: 1047 WT_atom: CE1 resi: 132 chain: A index: 1048

	Tells *which atom coordinates* can reference later (e.g., for distance calculations)	WT_atom: CE2 resi: 132 chain: A index: 1049 WT_atom: CZ resi: 132 chain: A index: 1050 WT_atom: OH resi: 132 chain: A index: 1051 IterateState: iterated over 12 atom coordinate states.
iterate_state 1, fluconazole, print("LIG_atom:", name, "resn:", resn, "resi:", resi, "index:", index)	• did the same as above, but for every atom in your fluconazole molecule (the ligand). • It printed all the atoms in fluconazole (carbon, nitrogen, oxygen, fluorine, hydrogen) and their corresponding indices. **<u>PURPOSE OF THE LAST 2 CODES</u>** Identify the specific atom indices in both the protein residue and fluconazole Then use those to measure the exact distance between them — which tells how strong/close their interaction is (useful for comparing WT vs. mutant)	**Output:** LIG_atom: N resn: UNL resi: 1 index: 1 LIG_atom: N resn: UNL resi: 1 index: 2 LIG_atom: N resn: UNL resi: 1 index: 3 LIG_atom: N resn: UNL resi: 1 index: 4 LIG_atom: N resn: UNL resi: 1 index: 5 LIG_atom: N resn: UNL resi: 1 index: 6 LIG_atom: C resn: UNL resi: 1 index: 7 LIG_atom: C resn: UNL resi: 1 index: 8 LIG_atom: C resn: UNL resi: 1 index: 9 LIG_atom: C resn: UNL resi: 1 index: 10 LIG_atom: C resn: UNL resi: 1 index: 11 LIG_atom: C resn: UNL resi: 1 index: 12 LIG_atom: C resn: UNL resi: 1 index: 13 LIG_atom: C resn: UNL resi: 1 index: 14 LIG_atom: C resn: UNL resi: 1 index: 15 LIG_atom: C resn: UNL resi: 1 index: 16 LIG_atom: C resn: UNL resi: 1 index: 17 LIG_atom: C resn: UNL resi: 1 index: 18 LIG_atom: C resn: UNL resi: 1 index: 19 LIG_atom: O resn: UNL resi: 1 index: 20 LIG_atom: F resn: UNL resi: 1 index: 21 LIG_atom: F resn: UNL resi: 1 index: 22 LIG_atom: H resn: UNL resi: 1 index: 23 LIG_atom: H resn: UNL resi: 1 index: 24 LIG_atom: H resn: UNL resi: 1 index: 25 LIG_atom: H resn: UNL resi: 1 index: 26 LIG_atom: H resn: UNL resi: 1 index: 27 LIG_atom: H resn: UNL resi: 1 index:

		28 LIG_atom: H resn: UNL resi: 1 index: 29 LIG_atom: H resn: UNL resi: 1 index: 30 LIG_atom: H resn: UNL resi: 1 index: 31 LIG_atom: H resn: UNL resi: 1 index: 32 LIG_atom: H resn: UNL resi: 1 index: 33 LIG_atom: H resn: UNL resi: 1 index: 34 IterateState: iterated over 34 atom coordinate states.
# OH of Tyr132 (WT) to the ligand oxygen (likely the drug's polar O) distance WT_OH_to_Lig_O, (ERG11_WT and index 1051), (fluconazole and index 20)	how far apart the hydroxyl oxygen on the Y132 residue of the WT is from a specific oxygen atom on fluconazole. • gives a quantitative measure of how close the drug sits to that critical binding residue→ indicator of binding strength or potential hydrogen bonding.	 38.0
distance WT_OH_to_Lig_N1, (ERG11_WT and index 1051), (fluconazole and index 1)	measures the distance between the same OH atom of Tyr132 and the first nitrogen atom of fluconazole. • see if it's near another potential interaction site — fluconazole has both N and O atoms that can form hydrogen bonds with residues.	 40.8

distance VAR_F_to_Lig_O, (Y132F_Variant and index 1049), (fluconazole and index 20)	Measured the distance (in Å) between the hydroxyl oxygen (OH) of Tyr132 in the WT model and a particular oxygen atom on fluconazole.	 39.0
distance VAR_F_to_Lig_N1, (Y132F_Variant and index 1049), (fluconazole and index 1)	Measured the distance between the same Tyr132 OH and a specific nitrogen atom on fluconazole.	 41.7
iterate_state 1, (Y132F_Variant and chain A and resi 132 and name CA), print("VAR CA coords:", x, y, z) # or by index (if index for variant CA was 1041 or 1049 etc) iterate_state 1, index 1041, print("VAR CA coords:", x, y, z)	• These are the 3D coordinates (x, y, z) of the alpha carbon (CA) in Tyrosine 132 of WT • These coordinates mark the center of the binding pocket region where fluconazole binds and where the Y132F mutation happens in the variant.--> can make swissdock more precise	WT CA coords: -11.711000442504883 10.125 -4.593999862670898 IterateState: iterated over 1 atom coordinate states. MEANING: x = -11.711 y = 10.125 z = -4.594

Sequence “Spell Book”:

Wild-Type Candida Auris ERG11 Sequence

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIYDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Wild-Type Candida Auris Y132F Variant Sequence

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Fluconazole in SMILES Format

OC(Cn1cncn1)(Cn1cncn1)c1ccc(F)cc1F

Candidate 1→ I131F

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVYFDCPNSRLME
QKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRAR
FDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTY
KDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQ
KMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWD
EETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDY
QSMVTLPMEPAEIEWEKRETCVY

Candidate 2→ N132C

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPCSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 3→Q142S

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMES
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 4→Q142T

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMET
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ

SMVTLPMEPAEIEWEKRETCVY

Candidate 5→V130W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGWIFDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 6→ I131W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVWFDCPNSRLME
QKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRAR
FDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTY
KDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQ
KMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWD
EETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDY
QSMVTLPMEPAEIEWEKRETCVY

Candidate 7→ P135W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCWNSRLME
QKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRAR
FDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTY
KDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQ
KMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWD
EETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDY
QSMVTLPMEPAEIEWEKRETCVY

Candidate 8→ N136C

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPCSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 9→ Y118W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAWSHLTTPVFGKGVIFDCPNSRLME
QKKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRAR
FDASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTY
KDGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQ
KMPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWD
EETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDY
QSMVTLPMEPAEIEWEKRETCVY

Candidate 10→ C134W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDWPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 11→ R381W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFWKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 12→ R381S

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFSKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 13→ R381T

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFTKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 14→ D133W

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFWCPNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 15→ P135H

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF

FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCHNSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY

Candidate 16→ P135A

MALKDCIVDVVDRFSALPVPVKLAVLILVPIVYNLVWQFVYSLRKDRAPLVFHWVPWVGSAVVYGMQPYQF
FESCREKYGDVFAFVMLGKVMTVYLGPKGHEFVLNAKLADVSAEAAYSHLTTPVFGKGVIFDCANSRLMEQ
KKFAKTALTKEAFQRYVPRIQEEVLDYFKACSQFRMNERNNGVANVMKTQPEMTILTASKSLMGDDMRARF
DASFAKLYSDLDKGFTPINFVFPHLPLPAYWKRDAAQQKISATYMSLINERRKTGDIVPDRDLIDSLMTNSTYK
DGVKMTDQEVANLLIGVLMGGQHTSASTSAWFLLHLAEQPKLQEELYSEVLSVLADKGGSLKDLAYDDLQK
MPLINQTIKETLRLHMPLHSIFRKVMNPLVVPNTKYVVPKGHYVMVSPGYAQTNEKWFPRANEFDPHRWDE
ETSSNIDTDAVDYGFGKVTKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILATYVYNIKWRFKKDGSLPPVDYQ
SMVTLPMEPAEIEWEKRETCVY