Asymmetrical dose-responses shape the evolutionary trade-off between antifungal resistance and nutrient use

Philippe C Després[1,2,3,4]*, Angel F Cisneros[1,2,3,4], Emilie MM Alexander[1,2,3,4], Ria Sonigara[1,2,4,5], Cynthia Gagné-Thivierge[1,2,3,4,5], Alexandre K Dubé[1,2,3,4,5], Christian R Landry[1,2,3,4,5]*

1. Département de Biochimie, de Microbiologie et de Bio-informatique, Faculté des Sciences et de Génie, Université Laval, G1V 0A6, Canada

2. Institut de Biologie Intégrative et des Systèmes, Université Laval, G1V 0A6, Canada

3. PROTEO, Le regroupement québécois de recherche sur la fonction, l'ingénierie et les applications des protéines, Université Laval, G1V 0A6, Canada

4. Centre de Recherche sur les Données Massives, Université Laval, G1V 0A6, Canada

5. Département de Biologie, Faculté des Sciences et de Génie, Université Laval, G1V 0A6, Canada

* To whom correspondence should be addressed: PCD: philippe.despres.3@ulaval.ca, CRL: Tel: 1-418-656-3954, Fax 1-418-656-7176, christian.landry@bio.ulaval.ca

Abstract

Antimicrobial resistance is an emerging threat for public health. The success of resistance mutations depends on the trade-off between the benefits and costs they incur. This trade-off is largely unknown and uncharacterized for antifungals. Here, we systematically catalog the effect of all amino acid substitutions in the yeast cytosine deaminase *FCY1*, the target of the antifungal 5-FC. We identify over 900 missense mutations granting resistance to 5-FC, a large fraction of which appear to act through destabilisation of the protein. The relationship between 5-FC resistance and growth sustained by cytosine deamination is characterized by a sharp trade-off, such that small gains in resistance universally lead to large losses in canonical enzyme function. We show that this steep relationship can be explained by differences in the dose-response function of 5-FC and cytosine. Our results provide a powerful resource and platform for interpreting drug target variants in fungal pathogens as well as unprecedented insights into resistance-function trade-offs.

Introduction

Pathogenic fungi are an emerging threat to human health, agriculture and wildlife[1]. Therapeutic options remain limited and resistance to all antifungals classes has been observed in multiple species[2]. Very little is known about the functional landscape of antifungal targets: the genotype of a strain is rarely enough to infer its phenotype. Recent efforts to catalog mutations associated with resistance have yielded a handful of mutations for a few targets, mostly derived from small scale studies and not always experimentally validated[3]. More systematic efforts are being deployed to catalog resistance mutations in a comprehensive manner[4,5], but information remains sparse for most targets.

The benefit of resistance to a given drug is not the only factor that will affect the long-term outcome of a mutation: the associated cost is also a key parameter[6]. For instance, a mutation in an essential protein that confers resistance but leads to a decrease in function would rapidly be lost upon treatment interruption. For example, resistance to the antifungal nystatin has been shown to have significant fitness costs in stressful environments[7], which means they would actively be selected against when nystatin concentration drops. In *Candida albicans*, aneuploidies are associated with antifungal resistance but are deleterious in many environments[8]. Accordingly, both a better inventory of the mutations granting resistance and the trade-offs associated with them are key to improve our ability to predict or interpret clinically relevant genotypes in complex environments such as in a human patient. Ideally, we would need approaches that allow us to measure this type of trade-off in a systematic and comprehensive manner for resistance mutations.

Most antifungals used in the clinic are drugs that bind their targets and disrupt their activity or impair their structural integrity[9,10]. For instance, azoles bind to the active sites of their targets and inhibit their function. Resistance mutations therefore tend to occur in the substrate binding pocket[11]. An exception to this is 5-fluorocytosine (5-FC, also known as flucytosine), which is administered as a pro-drug that is deaminated into the cytotoxic nucleotide analog 5-fluorouracil (5-FU) inside fungal cells. The accumulation of 5-FU leads to defects in mRNA synthesis and protein translation as well as DNA damage through thymine starvation[12], resulting in dose-dependent fungistatic or fungicidal activity. As one of the oldest antifungals, 5-FC is on the WHO list of essential medicines and is often used to treat cryptococcal meningitis in combination with amphotericin B[13]. Although resistance to 5-FC is often thought to arise more rapidly than for some other antifungals, the rate at which mutants occur *in vitro* has been shown to be lower than for fluconazole in *Candida glabrata*[14].

In yeast, the cytosine deaminase Fcy1p (Fca1p in *Candida albicans*) forms a homodimeric enzyme that converts 5-FC to 5-FU, and loss of function of this protein leads to 5-FC resistance[15]. Fcy1p is widely conserved throughout the fungal tree of life, and has a distinct fold compared to *Escherichia coli* cytosine deaminase[16]. As vertebrates lack a *FCY1* ortholog, 5-FC can be administered as a prodrug that will be selectively converted by fungal cells into 5-FU. The lack of cytosine deaminase activity in humans has made the enzyme a promising candidate for pro-drug gene therapy, where it is targeted to cancer cells to induce tumor specific toxicity[17,18]. This has led to extensive engineering of Fcy1p to optimize its properties as a pro-drug gene therapeutic agent[19–21]. Because of this wealth of knowledge,

Fcy1p is a powerful model to study the evolution of resistance and the potentially associated trade-offs.

FCY1 deletion was shown to confer full 5-FC resistance to *S. cerevisiae*[15], with little effect on fitness in multiple lab conditions[22,23]. Gene loss of function, for instance through premature stop codons, therefore appears to be an easily accessible mechanism of resistance. However, since Fcy1p is a dispensable enzyme in standard growth media, it is unclear if 5-FC resistance through loss of function has any trade-offs since enzyme function is generally decoupled from growth. In contrast, other antifungal targets, such as *ERG11* or *FKS1*, are considered essential for growth and thus for virulence in *C. albicans*[24]. Interestingly, *FCY1* can be conditionally essential in many laboratory strains where the *de novo* uracil synthesis pathway has been disrupted. In this scenario, cells must rely on the pyrimidine salvage pathway, where Fcy1p deaminates cytosine into uracil, complementing the auxotrophy. For example, in the lab strain BY4743, *FCY1* buffers the loss of function of *URA3* in minimal media supplemented with cytosine[25]. By changing environmental conditions, the interplay between resistance and Fcy1p activity can thus be used as a model for an essential gene imposing strict constraints on cell metabolism.

An important knowledge gap regarding the emergence of resistance is whether complete loss of function of Fcy1p is necessary or if partial loss of activity is sufficient, which would dramatically increase the mutational target size. Because the natural substrate, cytosine, and the toxic analog 5-FC have very similar structures, it is likely that a loss in the ability to process 5-FC will also lead to a loss in the ability to metabolize cytosine. If the trade-off is weak, there may be mutations that lead to resistance while maintaining the ability of the cell to utilize cytosine. These mutations could in principle lead to resistance without seriously affecting fitness in the absence of the drug. The relationship could be linear, in which case a 50% increase in resistance would lead to a 50% decrease in growth rate in cytosine media. At the other end of the spectrum, there may be a strong trade-off, which means that all mutations leading to 5-FC resistance impede cytosine use. This would then greatly favor the emergence of loss of function alleles at the expense of Fcy1p activity. For some types of coding sequence mutations, such as frameshifts or premature stop codons, loss of function can be easy to predict. However, significant challenges and uncertainty remain when trying to predict the effect of missense mutations distal from the active site. Possible consequences could include allosteric effects on enzymatic function, or simply destabilising protein structure or assembly into complexes. These perturbations would lower the effective protein concentration, potentially leading to loss of function phenotypes. Sensitivity to destabilisation would then be another key parameter in determining the mutational target size for resistance mutations.

Recent advances in genome editing and DNA sequencing have greatly increased the throughput at which protein variants can be assayed, allowing to draw a high-resolution map of fitness effects and potential trade-offs. Here, we use Deep Mutational Scanning (DMS) to explore the functional landscape of Fcy1p at its endogenous locus in the context of 5-FC resistance and cytosine deamination. We systematically assess protein variants to determine which mutations lead to drug resistance and explore which protein properties are linked to it. We examine how resistance mutations are distributed on the enzyme and we characterize the trade-off between resistance and cytosine use.

Results

Deep mutational scanning of Fcy1p

To generate the mutant *FCY1* library, we used the megaprimer method starting from the *Saccharomyces cerevisiae* amino acid sequence to create all possible single codon changes[26] (see Methods). This approach generates the full set of premature stop codons, silent mutations and single position missense mutants (often called protein variants). We integrated the mutants at the native *FCY1* locus by using a “bait and switch” genome editing strategy[27,28]. Since our approach requires direct sequencing of the coding sequence and because *FCY1* is too large to sequence with a single read pair using high-throughput sequencing, we split our mutants into three pools each covering partly overlapping sections of the coding sequence (Figure 1A). We used two different media compositions: media with 5-FC to select for resistance, and media without uracil but enriched with cytosine (cytosine media) to select for cytosine deamination into uracil. In cytosine media, the reaction is essential for growth. Using this approach, growth rate can be used as a proxy for Fcy1p flux for both enzymatic reactions in pooled competition assays (Figure 1B). Comparing the results would allow us to measure the strength of the trade-off between the two phenotypes across the mutational landscape for all single-step mutations (Figure 1C). To determine the screen conditions, we identified which 5-FC or cytosine concentrations reduce growth rate to ~50% of the maximum (also known as IC_{50}, Figure S1). This corresponds to roughly 12 µM and 84 µM for 5-FC and cytosine, respectively.

During preliminary experiments, we observed growth patterns on solid media suggesting that *fcy1Δ* could grow on cytosine media if wild-type strains were also present on the plates (Figure S2). We hypothesized this could be due to cross-feeding between wild-type and mutant cells, which could potentially impact selection strength during the pooled competition assay. Previous work has indeed shown that the yeast plasma membrane can leak uracil in the surrounding media[29]. The main contributor to uracil uptake in yeast is the transporter *FUR4*[30], the deletion of which is lethal when combined with a non functional *URA3* allele in standard media[31]. As uracil auxotrophy is essential to one of our selection conditions, we could not disrupt *FUR4* to prevent cross-feeding. Instead, we quantified the impact of potential uracil leakage on selection intensity in liquid cultures. Using a cytometry based competition assay, we measured the effect of different cytosine concentrations and population compositions on the selection coefficient of a *FCY1* deletion. We found that high cytosine concentrations and a higher initial proportion of wild-type cells tended to lower selection strength against *fcy1Δ* (Figure S3). Reassuringly, the selection coefficient remained above 0.3 at the cytosine IC_{50} (84 µM), even in populations with as much as 90% wild-type cells. This is well above the threshold required to detect fitness effects in a high-throughput sequencing assay[32].

We extracted genomic DNA and amplified the appropriate *FCY1* region for each pool and condition, and measured changes in allele frequencies between the start and the end of the competition. Synonymous codons were used as internal replicates for individual amino acids: $\log_2$ fold-change for each mutant was calculated as the median of the codon level fold-changes. Spearman’s correlation coefficient for replicate pairs ranged from 0.814 to 0.945 (Figure S4). To adjust for the differences between the different mutant pools (each

covering a third of the coding sequence), we harmonized the log_2 fold-change to pool 2 using the variants in the pool overlaps. The log_2 fold-changes of mutants in these two regions (codons 49 to 67 and 93 to 110 respectively) were well correlated between pools (Figure S5 and S6A-F). We then used the distribution of log_2 fold changes from the synonymous and nonsense mutants to put both conditions on the same scale to obtain the DMS score (Figure S6G and H). This allowed us to obtain a full-length functional landscape of Fcy1p that covered 99.96% (2892/2893) of possible missense variants in both selection conditions (Figure 2A and B).

The distributions of DMS scores are bimodal, with the two modes corresponding to WT sequences and premature stop codons (Figure S7A-C). Stop codon mutants can have different effects based on their position in the coding sequence, with generally weaker effects near the end of the coding sequence[33]. We observed that only the last amino acid of the C-terminus was partially dispensable. Surprisingly, stop codons that occured before the methionine in position 6 did not grant 5-FC resistance, suggesting that Fcy1p translation from the alternative start site remains efficient enough to confer sensitivity. This provides further evidence for the current canonical translation start site, which was challenged by comparative genomics work showing that the alternative start codon was conserved in close species such as *S. bayanus* and *S. mikatae*[34]. As such, we excluded stop codon mutants from positions 2-5 and 158 from downstream analyses.

Across the mutant library, 31% (922/2982) of missense mutants had 5-FC DMS scores that fell in the range observed for nonsense mutants (5-FC DMS score ≥ 0.931, 1st percentile of nonsense scores). Of the resistant mutants, 276 could be recapitulated by a single nucleotide change in the *S. cerevisiae FCY1* coding sequence, representing a large set of easily accessible mutations leading to resistance (Figure S8). The mutational target size of resistance is therefore much larger than stop codon mutations. For mutants in pool 2, which cover most of the substrate binding and catalytic residues of Fcy1p, we also performed the experiments in an additional condition to investigate whether some mutations could improve cytosine specificity. To test this, we used cytosine media with 24 µM 5-FC and 113 µM cytosine, in which wild-type cells grew at a similar rate as in the two other conditions. Mutant scores were highly correlated between the cytosine and 5-FC+cytosine condition (Figure S7D), with the only difference being better signal in the latter, probably due to more stringent selection. Based on this comparison, we did not identify any mutants in our library for which we could infer increased specificity for cytosine over 5-FC compared to the wild-type. Having catalogued which amino acid changes conferred resistance or maintained Fcy1p activity, we next sought to better understand the mechanisms behind mutational effects.

5-FC Resistance mutations are distributed across the protein

Mapping the DMS scores on the Fcy1p structure (Figure 2C and 2D), we found that the active site and the residues in close contact with the substrate were extremely sensitive to mutations. Our results corroborate what has been observed in small-scale studies on Fcy1p biochemistry. For example, Glu 64 is known to be essential for catalytic activity, and we did not observe any variant at this position that maintained 5-FC sensitivity or growth in cytosine[35]. Focusing on functional sites known to be involved in the catalytic cycle, we found that the core components of Fcy1p were strongly intolerant to mutations with respect to cytosine use (Figure 3A). Surprisingly, most resistance mutations do not reside in the active

site of the enzyme, suggesting that the enzyme is very sensitive to amino acid changes, either by allosterically affecting the enzymatic activity or perturbing the stability of the complex.

To better understand the constraints linked with maintaining Fcy1p function, we assembled a set of 215 orthologs (Figure S9) and hypothesized that conserved sites would be more sensitive to mutations. On average, amino acids from *S. cerevisiae* Fcy1p were present in 60% of orthologous sequences at the homologous positions, but the distribution of occupancy was bimodal (Figure S10A). As much as 18.5% of residues were observed in more than 95% of sequences, indicating strong structural and functional constraints maintaining residues for millions of years of evolution. Sites not part of the active site or involved in the dimer interface accounted for 41% of highly conserved amino acids, highlighting the importance of distal sites for Fcy1p function.

We directly investigated the effect of missense mutants in our library for which an equivalent substitution could be found in orthologs. Variants observed in orthologous sequences had much lower 5-FC scores and higher cytosine scores compared to those not observed (Figure 3B). Again the distributions are bimodal, which means for instance that some amino acids are compatible with the WT phenotype in *S. cerevisiae* but are not seen in orthologs. Since the number of orthologs is limited and does not cover the entire sequence space compatible with this enzyme function, this is not unexpected. There are few cases where amino acids present in orthologs lead to resistance to 5-FC when introduced in *S. cerevisiae* but those mainly concern amino acids present in the most distant species, which means that they could have been compensated for by other amino acid substitutions in the sequence[36]. For example, mutating the interface residue Glu 69 to an Arg leads to 5-FC resistance even if Arg 69 is relatively frequent in the alignment (12%, 25/215 orthologs). However, all the orthologs with Arg 69 are from outside the fungi kingdom (22 bacteria, 2 archaea and 1 plantae). We therefore cannot exclude functional divergence as an explanation for the conservation of Arg 69. Overall, only 55 amino acid changes observed in orthologs conferred resistance, with most of them being unique to a specific sequence (median occupancy=0.47%).

Protein complex destabilisation is a major mechanism of resistance

Many of the highly conserved residues that lead to resistance when mutated are away from the active site. This is possibly due to the importance of structural integrity of the complex or instances of allosteric regulation of enzyme function. Using an *in silico* approach[37], we predicted the effect on Fcy1p structure and dimer interface stability of all missense mutants (see methods). The predicted effects of amino acid mutants were strongly correlated with their DMS scores, being positively correlated with 5-FC resistance ($\rho=0.514$, $p=2.14\times10^{-199}$, $n=2964$) and negatively correlated with cytosine DMS scores ($\rho=-0.405$, $p=2.7\times10^{-117}$, $n=2964$). Protein destabilisation is therefore a key determinant of 5-FC resistance.

The dimer interface also appears to play a role in maintaining Fcy1p function. While protein engineering efforts have investigated mutations predicted to stabilise the dimer[20], whether dimerization is essential for activity remains unknown. The alpha-helix at the core of the

interface is highly sensitive to mutations, suggesting that dimerization is important (Figure 2C and D). We found that mutants of residues in contact with the other subunit tended to have lower cytosine DMS scores compared to those near (within 6 Å of a residue from the other chain) or distant from the interface (Mood's median test, $p= 6.36 \times 10^{-6}$ at interface vs near interface, $p=8.84 \times 10^{-15}$ near interface vs distant). Interestingly, we did not observe the same relationship for the 5-FC DMS scores, where only the difference in median score between mutants located near the interface and those distant was significant (Mood's median test, p= 0.775 at interface vs near interface, p=0.024 near interface vs distant). Leveraging our set of orthologs, we calculated the evolutionary rate for each position in the alignment[38] (Figure S10B), which revealed that both the core dimer interface and the interface rim evolved at a slower rate compared to other regions of the protein (Figure 3D). While homodimer formation appears to be essential to Fcy1p function, perturbing the interaction appears insufficient to grant resistance.

The asymmetric effects of mutations affecting the Fcy1p dimer interface in the DMS data prompted us to validate the observations made with the *in silico* predictions. We used an assay that reports on the amount of protein complex formed in the cell. Dihydrofolate reductase protein complementation assay (DHFR-PCA) is an *in vivo* assay that allows quantitative measurement of a protein-protein interaction (PPI) by linking it to the reconstitution of a methotrexate-resistant DHFR[39]. Growth quantitatively varies as a function of both the abundance of the proteins interacting and PPI strength[39,40]. This provides information on the abundance of the Fcy1p dimer in conditions in which FCY1 is dispensable so as to not confound the PPI signal with the fitness of the mutants. Starting from a set of individually reconstructed mutants (see next section), we examined 54 Fcy1p variant to measure PPI abundance. We measured the interaction between a wild-type copy of Fcy1p and each mutant, which reflects the stability and ability of the WT-mutant complex to form.

The *in silico* predicted effects on structure and interface stability were highly correlated between the mutant-mutant and WT-mutant dimer models (Figure S11A and B). Replicate control and methotrexate growth rates of the DHFR-PCA assays were well correlated between variant replicates (Figure S12A and B) and did not reveal any bias due to lower overall growth of some mutants (Figure S12C). As expected if destabilisation were causing resistance, variants that clustered with silent mutations in the DMS assays generally had detectable PPI, while those that clustered with stop codon mutants generally lost the interaction (Figure 3E, p=0.006, Mood's median test). Importantly, PPI intensity was well correlated with FoldX predicted effects on protein structure stability ($\rho = -0.475$, p=0.0003, Figure S12D), as well as with variant growth in cytosine ($\rho=0.375$, p=0.009, Figure S12E) and negatively correlated with 5-FC resistance ($\rho =-0.477$, p=0.0005, Figure S12F). These results directly link altered protein complex abundance with mutant phenotypes, and validate the use of *in silico* tools to predict the effect of mutations on Fcy1p stability. On the other hand, there was no correlation between PPI intensity and *in silico* changes in PPI stability as predicted by FoldX ($\rho=0.156$ and p=0.259 for 54 mutants), limiting the relevance of such metric for Fcy1p. Based on these observations, structure destabilisation leading to lower protein complex abundance appears to be a frequent mechanism driving Fcy1p loss of function.

To have a wider picture of variant and residue characteristics that might influence mutational effects beyond destabilisation, we generated the all by all correlation matrix of mutant

residue properties and DMS scores (Figure 3E). We examined both biochemical properties such as predicted effect on stability or residue solvent accessibility and evolutionary information from our set of orthologs. Because of the high similarity between the 5-FC and cytosine deamination reactions, we would expect most relationships to be mirrored between the two conditions. Indeed, we found that for most properties, the correlation between the property and the measured DMS effect in the two conditions was roughly symmetric (opposite direction and approximately same ρ). Correlations were in general stronger for 5-FC compared to cytosine DMS scores, potentially due to noise introduced by uracil leaking that led to weaker selection. Exceptions to these rules were the properties linked with the dimer interface, such as its presence or proximity as well as the predicted dimer interaction ΔΔG. As observed earlier, these were negatively correlated with cytosine DMS scores but were not correlated with 5-FC DMS scores. FoldX predicted protein destabilisation was a strong predictor of DMS scores, indicating that protein complex destabilisation is a key mechanism explaining Fcy1p loss of function, which extends the mutational target size well beyond the active site and changes creating stop codons.

A strong trade-off exists between 5-FC resistance and cytosine use

At first glance, 5-FC and cytosine DMS scores appear negatively correlated with each other. This is visible in Figure 2: sites that lead to resistance to 5-FC when mutated generally have reduced fitness in cytosine media. However, as shown in Figure 3F, the scores are only weakly correlated (ρ=-0.22, $p=7.5x10^{-35}$, n=2892 missense mutants). A potential explanation for this would be a very sharp trade-off resulting in a front minimum that leads to near vertical and horizontal curves outside of the inflexion point (i.e. the black line in Figure 1C). To verify this hypothesis and validate the DMS scores, we selected 72 missense mutants across the two-dimensional phenotypic landscape inferred from DMS (Figure S13A) and generated the corresponding strain using genome editing by inserting the individual mutations at the *FCY1* locus. We included most mutants that showed large variation from the main clusters or that appeared to have high scores in both 5-FC and cytosine, as such outliers could have important implications for the emergence of antifungal resistance. Based on the similarity of their DMS scores with nonsense or synonymous mutations, we assigned 43 mutants to clusters with another 29 that were considered outliers.

For each mutant, we measured growth rates in 5-FC and cytosine screen conditions in duplicate. Growth rate replicates were highly correlated (Figure S13B and S13C). The 5-FC DMS scores were well correlated with validation growth rates (Figure S13D, ρ=0.72, $8.77x10^{-13}$), but the correlation was much weaker for cytosine scores (Figure S13E, rho=0.29, 0.015), potentially due to noise introduced by uracil leaking in the large-scale experiment. This is supported by a stronger correlation between growth in cytosine and 5-FC+cytosine DMS scores for the mutants that were part of pool 2 (Figure S13F). We observed a striking pattern in the relationship between a mutant's growth rate in 5-FC and cytosine: all variants, including the outliers from the DMS screen, fell along the same front (Figure 4A). The relationship was well described by a power law equation with a negative k, characteristic of a strong trade-off, for which we determined parameters using a linear regression on the log_2 transformed growth rates of individual mutants. In effect, this strong trade-off means that a Fcy1p mutation that increases resistance to 5-FC by ~25% will incur a

cost equivalent to ~67% of the cytosine media growth rate of the wild-type. The steepness of the relationship generates a front minimum where a large fraction of mutants are unfit in both 5-FC and cytosine media, which we could also observe in the DMS data (Figure S14A). We found that 42/43 of the mutants to which we had assigned a cluster recapitulated their cluster from the DMS experiments in the validation. Finally, we note that the trade-off escape that was observed for some outliers of the DMS experiments could not be validated with our growth curve data.

Looking back at the DMS data, we drew a similar front in the DMS two dimensional landscape. Mutant clusters along the trade-off, which covered around 70% of our dataset, showed interesting variation in the distribution of structural and genetic properties (Figure S14B). Mutations clustering with nonsense mutants tend to have large predicted effects on protein stability and occur at highly conserved positions, with the exact opposite occurring for mutations clustering with synonymous mutations. Destabilisation values of mutants at the front minimum often fell in the ambiguous range of 1-2 Kcal/mol, which corresponds to thresholds frequently used to define significant effects on protein folding[41]. This group of mutations appears to not destabilize the protein enough to provide resistance, but affect metabolic flux enough to prevent the use of cytosine for cell growth.

This difference in the consequences of mutations on phenotype might be the result of large differences in the catalytic efficiency of Fcy1p when deaminating cytosine instead of 5-FC. However, based on reported values in BRENDA[42], this is not the case: the distributions of catalytic efficiencies reported for the two substrates overlap, and the K_{cat}/K_M of the median values shows only a small increase in efficiency for 5-FC deamination (111.9 $mM^{-1}s^{-1}$ vs 104.4 $mM^{-1}s^{-1}$, Figure S15). Furthermore, our dataset also contains some mutants whose presence in the front minima would be difficult to attribute to change in the catalytic efficiency of the enzyme. The stop codons mutants from positions 2-5, which did not confer 5-FC resistance, cluster with the front minimum mutants in the DMS data. These mutants are expected to affect phenotype through lower protein levels due to the alternative translation start and not through changes in the *FCY1* coding sequence. As such, these observations pointed toward a difference in the interplay between protein complex abundance and the response to 5-FC and cytosine levels that could explain the shape of the trade-off.

Because 5-FC is a pro-drug, the total inhibitory effect on growth rate will depend both on concentration and the metabolic flux of its deamination by Fcy1p. In turn, metabolic flux will depend on the effective enzyme concentration and its catalytic efficiency. Logically, mutations that confer full resistance must therefore reduce flux below a certain threshold through changes in protein stability reducing the effective protein concentration or lower catalytic efficiency. As cytosine deamination by Fcy1p is the limiting factor determining growth rate in cytosine media, reductions in flux would instead result in slower growth. Because we observed that many resistance mutations were destabilizing, we hypothesized the shape of the trade-off could come from the fact that growth was more sensitive to slight changes in cytosine flux than in 5-FC flux near the IC_{50}. Mutations reducing protein abundance would result in higher loss in growth rate in cytosine than the magnitude of the gain in 5-FC. In effect, for mutants exposed to the same effective levels of 5-FC and cytosine, changes in growth rate should reflect the change in flux caused by the amino acid change. For example, a mutation that reduces Fcy1p activity within cells by 50% (through any mechanism) should result in a growth rate equivalent to a 50% reduction in compound

concentration. Furthermore, at IC_{50}, the substrate concentration in the media is many orders of magnitude under the known K_M of wild-type Fcy1p for both 5-FC and cytosine. Based on Michaelis-Menten kinetics, this means flux will vary as a function of substrate concentration. As our mutants are expressed at the endogenous level, this should remain true for them as well. This implies the phenotypes of mutants with reduced flux will change based on 5-FC and cytosine concentrations in a way that reflects their level of loss of function and dose-response kinetics. Using this framework, mutations can be conceptualized as having effects on the effective drug or metabolite concentration.

While the dose-response curves of Figure S1 were sufficient to determine IC_{50} in preparation of our DMS experiments, they did not cover the full range of response to 5-FC and had only minimal coverage of the range of cytosine concentrations where most changes occurred. We therefore performed a second round of dose-response curves with a greater emphasis on precisely determining curve parameters (Figure 4B and 4C). The IC_{50} and Hill coefficient we had first determined were close to the values we found in our second round of experiments (5-FC: 12 µM vs 9.14 µM, cytosine: 84 µM vs 71 µM). The most striking difference between 5-FC and cytosine response lies in the steepness of changes around the IC_{50}. Growth inhibition by 5-FC varies along a gradient that spans a 2048-fold difference in drug concentration: for cytosine, variation occurs within a range spanning only a 4-fold difference. The dose-response curves explain the shape of the trade-off between 5-FC resistance and cytosine use we observe. Because of the important difference in steepness around the IC_{50}, the same level of loss of function will lead to different outcomes (Figure 5A). For example, based on our growth-response data, a 2-fold decrease in effective 5-FC concentration starting from the IC_{50} would result in 36% (instead of 50%, Figure 5B) growth inhibition. The same change in effective cytosine concentration instead leads to loss of the ability to grow (from 50% to 1.4% relative growth, Figure 5C), explaining the front minimum we observed in the DMS experiment. This difference in dose-response should result in strong unidirectional effects, so that mutants having a full resistance phenotype would never recover the ability to use cytosine at higher concentrations.

To test our hypothesis, we performed growth curves and measured the area under the curve (AUC) at 24h for all 72 validation mutants in five additional 5-FC and cytosine concentrations higher than screen and validation conditions (Figure 5D, full growth curves shown in Figures S16 and S17). Large increases in cytosine concentrations (450 µM and 901 µM) could restore growth comparable to the wild-type for multiple mutants that showed no growth at concentrations closer to the IC_{50}. Conversely, a few mutants showed some level of sensitivity to higher levels of 5-FC (i.e L37F, T60F, T60R), supporting our model based on Fcy1p flux. Many front minima mutants (i.e. G76H, K80L, D92C) had sharp on/off dynamics in cytosine, consistent with the shifted dose-response curve our model predicts. We observed differences in growth between 450 µM and 901 µM cytosine for some mutants, indicating that the activity of the cytosine permease *FCY2* was not limiting in our assay. Importantly, the relationship between responses to 5-FC and cytosine were correctly predicted. No fully resistant mutant recovered the ability to use cytosine, and resistance gains were always linked to greater losses in cytosine use. The phenotypes of our Fcy1p mutants therefore support our trade-off model based on metabolic flux and shaped by different response patterns to 5-FC and cytosine.

Finally, we also found that the predicted effect of a mutation on structure stability was a strong predictor of whether it behaved as a complete loss of function (no response to either 5-FC or cytosine), as would be expected in a situation where protein complex abundance is critical. Mutations with a predicted ΔΔG from 1 to 5 Kcal/mol were significantly more likely to retain sensitivity or growth at higher compound concentrations compared to those with ΔΔG > 5 Kcal/mol (19/27 vs 6/18, odds ratio = 7.125, p= 0.002, two-sided Fisher's exact test). Thus, we find that changes in protein stability, which translates into lower complex abundance, are an important factor in determining the level of loss of function of Fcy1p and predicting resistance.

Discussion

Systematically characterizing the impact of mutations on drug target resistance and function is a key aspect of improving our ability to design new molecules, manage existing ones and potentially exploit trade-offs associated with resistance mutations. These goals include interpreting mutations that have never been reported before and understanding the mutational target sizes of different drug targets. Using genome editing and DMS, we systematically characterized the impact on 5-FC resistance and canonical enzyme function of all single amino acid mutations in *S. cerevisiae* Fcy1p and this at its endogenous locus. We catalogued 922 missense and 151 nonsense mutations that conferred 5-FC resistance, nearly 33% of all amino acid mutations. Resistance mutations were found all across the enzyme and not only near the active site. Our computational work and experimental assays suggest that this is largely due to effects on protein stability. The fact that destabilizing mutations in Fcy1p can be sufficient to lead to resistance points to the importance of maintaining the endogenous transcription levels for these assays as overexpression could mask some of the effects.

Our results provide definitive evidence for loss of function as the main 5-FC resistance mechanism at the amino acid level for Fcy1p. Adaptation through loss of function is not specific to our study system. Loss of function is known to be an important driver of adaptation to new environments[43,44], and can play a role in increasing the virulence of fungal pathogens[45]. For example, recent work in *C. glabrata* has identified many predicted loss of function mutations associated with resistance to fluconazole or anidulafungin[4]. Despite how widespread *FCY1* orthologs are, not much is known about their contribution to cell survival in natural or host environments, except that it can provide increased metabolic efficiency in some conditions[46]. The conservation of the enzyme across fungi suggests it plays a key role, which contrasts with the fact that it is dispensable in laboratory conditions.

An important factor influencing the fate of resistance mutations is the cost associated with them. Understanding the mechanisms that determine the relationship between resistance and growth in the absence of drugs is important because it could eventually open the door to using conditions that maximize the cost of mutations. Here we show an example of an extremely steep resistance-function relationship and provide a mechanistic explanation. Our results demonstrate that the shape of a trade-off is inherently linked to how steep the function linking growth to flux through the pathway is compared to the function linking flux and inhibition by the toxic analog. During treatment, the target serum concentration of 5-FC

is between 25 and 100 µg/ml (194 to 775 µM)[47]. In our dose-response curve, this corresponds to the upper plateau. Because of its low steepness, a mutation would have to reduce Fcy1p flux by over 400-fold to make the effective 5-FC concentration inhibit growth by less than 10%. At this level of loss of function, the remaining cytosine deaminase activity would be irrelevant to cell metabolism. The yeast membrane's propensity for uracil leakage[29] would also be expected to reduce the impact of Fcy1p loss of function, as a mutant would probably be surrounded by wild-type cells that could provide it with enough uracil to compensate for partial loss of the salvage pathway, alleviating the strength of the trade-off. The cost of resistance mutations could therefore be frequency dependent. How the shape of the trade-off we uncover affects evolution *in vivo* therefore needs to be investigated.

The characterization of trade-offs may help design better combination therapy in the future. For instance, combining 5-FC with another compound that targets the *de novo* pyrimidine synthesis pathway could provide a way to reduce or abolish its ability to buffer loss of functions in the salvage pathways. For example, olorofim is a newly developed antifungal currently undergoing phase IIb clinical trials that inhibits fungal dihydroorotate dehydrogenase, which is part of the *de novo* pathway[48,49]. While it has no activity against *Candida* and *Cryptococcus* species, olorofim has strong activity against molds such as *Aspergillus* species, many of which are also sensitive to 5-FC *in vivo*[50,51]. New compounds such as these could provide opportunities for combination regimens that take advantage of the potency of 5-FC while slowing the emergence of resistance during treatment. One key result of our screen is that we did not observe any mutants that escaped the strong resistance-function trade-off. This suggests that gaining resistance inevitably leads to loss in the ability to use cytosine. Arguably, one limitation is that we investigated single-step mutations and multiple-step mutations may be able to break this trade-off. In a situation where mutation rate is relatively weak and population sizes are limited, single-step mutations may be more representative of what could happen during treatment.

Antifungal resistance is a growing threat to both human health and food supply, but there is still limited data on the molecular mechanisms underlying it. By systematically exploring the effects of all possible single amino acid mutants, we hope to facilitate the interpretation of new genotypes in the clinical or industrial setting. Numerous efforts have been made in this sense in the field of antibiotic resistance in bacteria, where antibiotic resistance genes such as the β-lactamase TEM-1 have been thoroughly dissected and evolutionary pathways to resistance have been characterized[52]. In human genetics, DMS studies have helped shed light on the phenotypic effects of thousands of variants of unknown significance in disease or cancer genes[53]. In effect, the 1,073 mutants we identified will more than quadruple the number of amino acid substitutions recorded in the MARDy database, which catalogs all known antifungal resistance mutations[3].

Although we performed our screen in a model species, we observed a strong negative correlation between residue evolutionary rate and resistance to 5-FC: mutating highly conserved residues generally led to Fcy1p loss of function. Already, we can use our results to help interpret population genetics of clinical isolates or mutants isolated in drug resistance studies. For example, genetic variation within pathogenic lineages sometimes leads to 5-FC resistance even in the absence of treatment, as observed for *Candida dubliniensis* isolates from the C3 clade, which bear a S29L substitution in Fca1p. In our DMS screen, the equivalent mutation in yeast Fcy1p (G35L) confers resistance to 5-FC[54]. In *Cryptococcus*

neoformans, a recent study investigating the emergence of resistance to 5-FC identified a L40F mutation in Fcy1p in one isolate[55]. Again, our DMS data finds that the equivalent mutation, L37F, confers resistance. In both cases, these changes are expected to significantly destabilize Fcy1p structure ($\Delta\Delta G > 10$ kcal/mol) and so probably grant resistance through reduced complex abundance. Based on a review of known resistance associated mutation in fungal pathogens, our results correctly predict the resistance phenotype of 8/9 variants (Supplementary table 1). This strongly supports that the data we obtained for the *S. cerevisiae* protein can help interpret genotypes from across the fungal tree of life. Using this approach on other antifungal targets could help proactively identify mutations of interest before they are observed in the wild.

Methods

Strains, plasmids and culture media

Strains constructed as part of this study are listed in supplementary Table 2, plasmids used or constructed in supplementary Table 3, and oligonucleotides in supplementary Table 4. Media recipes can be found in supplementary Table 5. All yeast growth experiments were performed at 30°C unless stated otherwise and all bacteria were grown at 37°C. Liquid cultures were incubated with 250 RPM agitation. All Sanger sequencing was performed at the CHUL sequencing platform (Université Laval, Québec).

All strains were constructed in the BY4741 and BY4742 genetic backgrounds[56] unless stated otherwise. We used a standard lithium acetate transformation protocol[57] to construct gene deletions strains. The *FCY1* deletion strains BY4741 *fcy1Δ::NatNT2* (strain s_001) and BY4742 *fcy1Δ::HphMX4* (strain s_002) were constructed by amplifying the *NatNT2* or *HphMX4* cassettes from pFA-NatNT2[58] or pAG32[59] respectively using oligonucleotides del_for and del_rev and transforming the product in competent cells of the parental strain. Gene deletions were verified by PCR using oligonucleotides FCY1_A, Nat_B or Hph_B.

Cytometry competition assay

To measure competitive fitness of wild-type vs *fcy1Δ* cells, we constructed a fluorescent *fcy1Δ* reporter strain where the endogenous copy of *PDC1* is tagged with *mCherry* (s_003). We amplified the mCherry tag from a modified version of plasmid pBS35[60] where the selection marker is *NatMX4* using oligonucleotides PDC1_tag_for and PDC1_tag_rev, and used a standard lithium acetate protocol to transform the cassette in BY4742 *fcy1Δ::HphMX4* and plated the cells on YPD+Nat. Positive transformants were validated by PCR (using oligonucleotides PDC1_C and mCherry_B) and by fluorescence cytometry using a Guava® easyCyte 8HT (excitation: 532 nm, detection 609 nm). To perform the competition assay, four independent pre-cultures of BY4742 and BY4742 *fcy1Δ PDC1-mCherry* were grown overnight in SC media. The following morning, the cells were diluted to 0.1 OD in SC media and allowed to reach exponential phase. Cells from each culture were centrifuged and resuspended in sterile water to a concentration of 1 OD. The culture pairs were then mixed in 1:9, 1:1, 9:1 (*FCY1:fcy1Δ PDC1-mCherry)* and 20 µL of the mix was used to inoculate 180 µL 1.11X SC-ura media with either 225, 113, 84, or 56 µM cytosine so that the media reached 1X with the addition of the cells (final concentration: 0.05 OD) in a Greiner flat bottom transparent 96-well plate. The abundance of each genotype was then measured by counting the number fluorescent vs non-fluorescent yeast cells in each well, and the initial cell density was measured by counting the number of cells/µL. After ~16h incubation at 30°C, the final OD for each well was measured, and the cells were diluted to 0.05 OD in sterile water. The abundance of each genotype and the cell density was measured again for all wells. The selection coefficient was calculated using the following approach for each condition and replicate:

First, we measure the number of doublings for wild-type and *fcy1Δ PDC1-mcherry* by comparing the absolute abundance of each genotype at the start and the end of the competition:

$$Doublings_{WT} = log_2(WT_{start} - WT_{end})$$

$$Doublings_{\Delta} = log_2(Fluo_{start} - Fluo_{end})$$

With the endpoint counts adjusted to the final optical density at 16h. We then use the number of doublings of the wild-type and *fcy1Δ PDC1-mcherry* strains to calculate the relative fitness baked on the number of doubling of the wild-type:

$$w_{WT} = \frac{Doublings_{WT}}{Doublings_{WT}} = 1$$

$$w_{\Delta} = \frac{Doublings_{\Delta}}{Doublings_{WT}}$$

From there, since the selection coefficient is defined as

$$s = 1 - w$$

We have

$$s_{\Delta} = 1 - w_{\Delta}$$

Which gives us the selection coefficient for the *fcy1Δ PDC1-mcherry* strain. As there was little to no overnight growth in the 56 µM cytosine samples, we excluded them from the selection coefficient calculation.

Creations and validation of the *FCY1* DMS library

To be able to differentiate it from the wild-type sequence in eventual diploid assays, we decided to start from a codon optimized *FCY1* sequence to build the DMS library. The yeast codon optimized sequence was synthesized and cloned into a high copy vector (Twist Biosciences, plasmid pTWIST-FCY1_opt). We then used the mega-primer method[26] to generate codon by codon mutant libraries by using oligonucleotides with three degenerate nucleotides (NNN). In the first PCR step, the degenerate forward oligonucleotides are paired with a reverse primer to generate a short amplicon where one codon has been mutated:

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Reverse Primer 10 µM	0.45

Codon-specific NNN oligo 10 µM	0.45
Kapa polymerase	0.1
Plasmid DNA 30 ng/µL	0.23
PCR grade water	10.33
Total	15

With the following PCR cycle:

95 °C	5 mins	
98 °C	20 s	35x
60 °C	15 s	
72 °C	30 s	
72 °C	50 s	

For codons 2 to 57, the reverse oligonucleotide was mega_1, mega_2 for codons 58-133 and mega_3 for codons 114-158. The resulting mega-primer product was verified on agarose gel and used as a primer for the directed mutagenesis PCR reaction with the following parameters:

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Mega-primer	0.75
Kapa polymerase	0.1
Plasmid DNA 30 ng/µL	0.23
PCR grade water	10.48
Total	15

With the following PCR cycle:

95 °C	5 mins	

98 °C	20 s	22x
68 °C	15 s	
72 °C	3 mins	
72 °C	4.5 mins	

After amplification, 0.3 µL 20,000 U/ml DpnI (New England Biolabs) was added to the reaction mix which was then incubated at 37°C for 1 hour to digest the template plasmid. We then transformed 5 µl of the digested mutagenesis reaction into E. coli MC1061 (*[araD139]*$_{B/r}$ *Δ(araA-leu)7697 ΔlacX74 galK16 galE15(GalS) λ-e14-mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2*)[61] using a standard chemical transformation method. This was sufficient to obtain hundreds to thousands of clones per codon, more than enough to cover all 64 possibilities.

The mutagenesis was performed in three batches corresponding to the three different reverse primers in the first PCR steps. For each codon, transformants were recovered by soaking the transformation plate with 5 mL 2YT media and scraping with a glass rake to recover cells in a 15 mL falcon tube. The pools were homogenized by vortexing and then archived by mixing 800 µl bacteria pool and 200 µL 80% glycerol for storage at -80°C. Finally, plasmid DNA was extracted from each pool with a standard miniprep kit (FroggaBio, Concord, Canada) using 1.5 ml of bacterial suspension as input and eluting in 50 µL kit elution buffer (10 mM Tris-HCl pH 8.5). After all extractions were completed, plasmid DNA concentrations were normalized to 1 ng/µL in two Greiner V-shape 96 well plate by first diluting to 10 ng/µL in 20 µL PCR grade water and then adding another 180 µL of PCR grade water. This plasmid DNA dilution was then used as the template for quality control of the DMS library.

To validate the mutant plasmid pools, we used Row-Column (RC) indexing[62] to generate barcoded position by position libraries for high-throughput sequencing in three PCR steps. We split the *FCY1* coding sequence into three overlapping regions for Illumina paired-end sequencing. First, we amplified the appropriate *FCY1* region with oligonucleotides containing flanking universal priming sites (PS) at each end using with specific oligonucleotides (F1: pool_1_for, R1: pool_1_rev, F2: pool_2_for, R2: pool_2_rev, F3: pool_3_for, R3: pool_3_rev):

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
PS-Forward oligo 10 µM	0.45

PS-Reverse oligo 10 µM	0.45
Kapa polymerase	0.1
DMS plasmid library DNA 1 ng/µL	4.5
PCR grade water	6.05
Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	20X
60 °C	15 s	
72 °C	30 s	
72 °C	1 min	

After gel verification, the amplicons were diluted 1/2,500 by two consecutive 2 µL PCR : 98µL PCR grade water dilutions. This was then used as a template for a second PCR round where plate position specific barcodes were added at both ends of the amplicon (forward: oligonucleotides row_1 to row_8, reverse: oligonucleotides col_1 to col_12).

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Row-Forward oligo 2 µM	2.25
Col-Reverse oligo 10 µM	0.45
Kapa polymerase	0.1
Diluted PS-PCR	2.25
PCR grade water	6.5
Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	18X
60 °C	15 s	
72 °C	30 s	
72 °C	1 min	

All barcoded amplicons of the same pool were then mixed in roughly equivalent quantities after gel band intensity quantification. The resulting amplicon mixes were purified using magnetic beads and used as templates for a final PCR round performed in quadruplicate to add plate barcodes at each end as well as the Illumina p5 and p7 sequences using a different combination of oligonucleotides for each pool (pool 1: plate_for_1/plate_rev_3, pool 2: plate_for_2/plate_rev_3, pool 3: plate_for_3/plate_rev_3):

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
plate_for oligo 10 µM	0.45
plate_for oligo 10 µM	0.45
Kapa polymerase	0.1
Purified RC-PCR pool 0.01 ng/µL	0.45
PCR grade water	10.1
Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	20X
61 °C	15 s	
72 °C	35 s	
72 °C	1 min	

The replicate reactions were pooled and then purified on magnetic beads and gel verified to obtain the final libraries. Sequencing was performed using the MiSeq Reagent Kit v3 on an Illumina MiSeq for 600 cycles (IBIS sequencing platform, Université Laval).

For a few codons (19, 80 and 91), the mutagenesis primers failed systematically and we therefore used equivalent degenerate oligonucleotides (alt_dms_19, alt_dms_80 and alt_dms_91) on the reverse strand with a matching forward oligonucleotide (mega_4) to generate the mutant plasmid pools. We also performed a second round of mutagenesis for some positions that showed poor coverage in the QC analysis (codons 83, 84, 107 and 108). These 7 pools were validated by Sanger sequencing.

Next-Generation Sequencing data analysis

After sequencing, samples were demultiplexed using custom python scripts (https://github.com/Landrylab/Despres_et_al_2021) according to plate and RC barcodes after quality control with FastQC[63] and trimming with Trimmomatic[64]. Data analysis scripts made use of the following packages: matplotlib[65], numpy[66], pandas[67], seaborn[68] and scipy[69]. After demultiplexing, the forward and reverse reads were merged using the PANDA-Seq software[70] and collapsed into 100% identity clusters with vsearch[71]. Reads were then aligned to reference locus sequences using the Needle software from EMBOSS[72]. A custom script was then used to parse the alignments and extract codon frequencies for each sub-library to assess variant coverage. All raw sequencing files were deposited on the NCBI SRA, accession number PRJNA782569.

CRISPR-Cas9 Knock-in of *FCY1* variants

We used a modified version of the protocol presented in Ryan et al. 2016[73] to generate all the competent cells required for the DMS yeast transformations in a single batch. Colonies from a BY4742 *fcy1Δ::HphMX4* plate were used to inoculate a 20 mL YPD + Hyg overnight pre-culture. The following day, the cells were diluted to 0.2 OD in 300 mL YPD and grown to ~1 OD. The culture was split in 50 ml falcon tubes and the cells were pelleted by centrifugation at 448 *g* for 2 minutes. The supernatant was removed and the pellet was washed with 0.5 mL LATE solution (100 mM Lithium acetate, 1mM Tris-Cl, 0.1 mM EDTA, filter sterilized) before being transferred to a 1.5 ml microcentrifuge tube. The cells were then pelleted again by centrifugation at 8,600 *g* for 1 minute, and the supernatant was removed. Finally, the cells were resuspended in 300 µl LATE solution and 300 µl 50% glycerol, mixed by a quick vortexing, and stored at -80°C until use.

To integrate the DMS pools for each codon in yeast, we co-transformed a pCAS[73] variant bearing a gRNA targeting the *Hph* resistance gene (pCAS-Hph, p_005) as well as donor DNA amplified from the codon-specific mutant plasmid pool recovered from bacteria. The modified pCAS vector was generated using the same approach as in Dionne et al. 2021[28], using oligonucleotides pCAS_hph_for and pCAS_hph_rev. The donor DNA for each codon was amplified using the following reaction:

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Oligo knock_in_for 10 µM	0.45
Oligo knock_in_rev 10 µM	0.45
Kapa polymerase	0.1
DMS plasmid library 2 ng/µL	3.5
PCR grade water	7.05
Total	15

With the following PCR cycle:

95 °C	5 mins	
98 °C	20 s	29X
60 °C	15 s	
72 °C	45 s	
72 °C	1.5 mins	

We performed one transformation/codon in a 96 well PCR plate using a similar approach to what is described in Dionne et al 2021[28]. Each well contained 20 µL competent cells, 2 µL carrier DNA (salmon testes ssDNA 10 mg/mL, boiled for 5 min and then cooled on ice) and 3.75 µL 200 ng/µL pCAS-Hph, prepared as a master mix and dispensed before adding 10 µL donor DNA. A 200 µL volume of freshly prepared PLATE solution (40% PEG 3350, 100 mM Lithium acetate, 1mM Tris-Cl, 0.1 mM EDTA, filter sterilized) was then added to each well. The plates were sealed with sterile aluminum foil, vortex mixed, and incubated at 30°C for 30 minutes before a 15 minutes heat shock at 42°C. The plates were then centrifuged at 448 *g* for 6 minutes to pellet the cells, the supernatant was removed, 100 µL YPD was added, and the plates were sealed with aluminum foil again for a 3h30 outgrowth at 30°C. The cells were then plated on YPD+G418 with glass beads and incubated at 30°C for ~64 hours.

To recover cells, the plates were soaked with 5 ml YPD, scraped with a glass rake and the cell suspension was transferred to a 15 mL falcon tube. The cell suspensions were homogenized by vortexing, and the OD was measured for each pool. Based on the cell concentration, a volume equivalent to 5 OD units of each pool was mixed to generate the three DMS mutant pools: pool 1 contained codons 2 to 67, pool 2 codons 49 to 110, and pool 3 codons 93 to 158. The OD of each pool was measured to measure their final cell concentration, and pool aliquots were archived by mixing 380 µL of cell pool with 180 µL of 80% glycerol. Pool aliquots were stored at -80°C until use. A 500 µL sample of each pool

was spun down and used as input for genomic DNA extraction with the phenol/chloroform method[74]. Next-generation sequencing libraries were generated and sequenced on illumina for a final round of quality control to ensure proper coverage in each pool. Sequencing was performed using the MiSeq Reagent Kit v3 on an Illumina MiSeq for 600 cycles (IBIS sequencing platform, Université Laval).

The same approach was used to build the strain BY4742 *FCY1_codon_opt* (s_004), but using the wild-type plasmid and oligonucleotides KI_short_for and KI_short_rev to generate the donor DNA, and recovering single colonies instead of scraping the entire transformation plate. Transformants were used to inoculate 1 mL YPD cultures, grown overnight, and then streaked on YPD. One colony from each streak was resuspended in 200 µL water and spotted on YPD as well as YPD+G418 and YPD+Hyg to check for pCAS loss and successful loss of the *HphMX4* cassette. The strain was then validated by PCR and Sanger sequencing.

DMS library screen and library preparation

An archived aliquot of each pool was thawed ~30 minutes on ice. From the corresponding tubes, a volume corresponding to 2 OD units was used to inoculate 20 ml SC media in duplicate. The cells were grown for 24h and then split in the 5-FC and cytosine conditions: SC + 12 µM 5-FC or SC-ura + 84 µM Cytosine. Each 12h, 2 OD units of cells were centrifuged for 5 minutes at 224 *g*, the supernatant was removed, and the pellet was resuspended in 1 ml sterile water. The cell suspension was then used to inoculate 19 mL of 1.05X media, so that the addition of the cells brought the media to 1X. The screen was stopped once pools reached ~20 generations based on OD measurements: this took two passages for the 5-FC condition, and 4 passages for the cytosine condition. Genomic DNA was extracted from 2 OD pellets (approximately $2x10^7$ cells) of the final passage with the phenol/chloroform method. The same approach was used for pool2 and the 5-FC+cytosine condition, but with SC-ura + 24 µM 5-FC + 113 µM cytosine.

To prepare NGS libraries for deep sequencing, we used a similar method as was used for the plasmid pool QC, but barcoding individual timepoints instead of codons, and with each reaction being performed in duplicate:

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Pool forward oligo 10 µM	0.45
Pool forward oligo 10 µM	0.45
Kapa polymerase	0.1
DMS screen gDNA 30 ng/µL	5
PCR grade water	5.55

Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	20X
60 °C	15 s	
72 °C	30 s	
72 °C	1 min	

After gel verification, the amplicons were diluted 1/2,500 by two consecutive 2 µL PCR : 98µL PCR grade water dilutions. This was then used as a template for a second PCR round where plate position specific barcodes were added at both ends of the amplicon (primers row_1-3 with primers col_1-6), with each reaction performed in duplicate.

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Row-Forward oligo 2 µM	2.25
Col-Reverse oligo 10 µM	0.45
Kapa polymerase	0.1
Diluted PS-PCR	2.25
PCR grade water	6.5
Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	18X
60 °C	15 s	

72 °C	30 s	
72 °C	1 min	

Replicate reactions were pooled together, purified using magnetic beads, and quantified using a NanoDrop spectrophotometer. Barcoded amplicons of the same pools were then mixed in equivalent quantities. The resulting pools diluted and used as templates for a final PCR round performed in quadruplicate to add plate barcodes at each end as well as the Illumina p5 and p7 sequences:

Reagent	Volume (µL)
Kapa HF buffer 5X	3
dNTPs 10 mM	0.45
Plate_for-Forward oligo 10 µM	0.45
Plate_rev-Reverse oligo 10 µM	0.45
Kapa polymerase	0.1
Purified RC-PCR pool 0.01 ng/µL	1
PCR grade water	9.55
Total	15

With the following PCR cycle:

98 °C	3 mins	
98 °C	30 s	20X
61 °C	15 s	
72 °C	35 s	
72 °C	1 min	

The replicate reactions were pooled, purified on magnetic beads and gel verified to obtain the final libraries. Sequencing was performed using the MiSeq Reagent Kit v3 on an Illumina MiSeq for 600 cycles (IBIS sequencing platform, Université Laval).

DMS score

Data analysis of the DMS screen was performed using custom python scripts. Briefly, samples were demultiplexed into individual fastq files containing merged reads using the same strategy as was used for the plasmid pool QC analysis. Identical reads were then collapsed using vsearch[71] and aligned on the appropriate *FCY1* region using Needle[72] to detect mutations. Reads with more than three mutations in the target DMS sequence or with mutations in multiple codons were discarded. Library raw counts were converted to relative abundance within libraries by dividing counts by the total number of DMS mutants and wild-type reads. Relative abundances were then compared between the end of the SC media passage and the end of the selection conditions (5-FC or cytosine) to obtain $\log_2$ fold-changes for all mutant codons. To convert to amino acid log2 fold-change, we considered each group of synonymous codons as replicates and calculated the median fold-change. This allowed us to obtain three (one per pool) amino-acid level matrices per condition. Finally, we averaged $\log_2$ fold-change between replicates (as shown in Figure S4). As some *FCY1* positions in pools 1 and 3 overlapped with those in pool 2 (positions 49 to 67 and 93 to 110 respectively), ~400 mutants/pool had their $\log_2$ fold-change measured in duplicate. As the fold-changes of same mutants were well correlated between pools and spanned the full range of possible values, we fit a linear regression that allowed us to adjust all scores to pool 2. For mutants in the overlapping regions, we used the average of the two log2-fold changes (one from pool 2, and the adjusted value from pool 1 or 3).

Once we had harmonized $\log_2$-fold changes along our three pools, we wanted the values for both experiments (5-FC and cytosine) to be on the same scale to facilitate comparison. To do this, we scaled values using the median log2-fold change of synonymous mutations and nonsense mutants using the following equations:

For 5-FC:

$$FC_{score} = \frac{AA_{log_2} - SYN_{median}}{STOP_{median} - SYN_{median}}$$

For cytosine

$$CY_{score} = \frac{AA_{log_2} - STOP_{median}}{SYN_{median} - STOP_{median}}$$

This scaling relates the $\log_2$ fold-change observed for a DMS mutant with the difference observed between the values of synonymous (silent) and nonsense (stop) mutants. For example, a score of 0.5 would represent a $\log_2$ fold-change value that is perfectly in the middle between the medians of synonymous and nonsense mutations. We chose to set the score according to the phenotype measured, so that in the 5-FC condition nonsense mutations would have a median FC_{score}=1 and synonymous mutation would have a median median CY_{score}=0. Conversely, in the cytosine condition, synonymous mutations would have a median CY_{score}=1 and nonsense mutations would have a median median FC_{score}=0.

Amino acid and mutation properties analysis

We used the structure from Ireton et al. 2003[16] (PDB: 1P6O) to compute residue relative solvent accessibility (RSA), distance to the active site and interface, and as baseline for the FoldX[37] ΔΔG predictions which were performed using the same approach as in Marchant et al. 2019[75]. Briefly, the biological assembly was generated and energy minimization was performed using the FoldX Repair function 10 times to ensure convergence[76]. Then, the FoldX BuildModel and AnalyseComplex functions were used to estimate mutational effects. As mentioned in the text, FoldX predictions were generally comparable between dimer bearing one or two mutated copies. RSA and temperature measurements were well correlated between the matching residues on both chains (Figure S11C and D), and the values were averaged to obtain the final value. The distance to the active site was considered to be the shortest distance between a specific amino acid and any active site residue. To determine which residues were part of the interface, we used a definition based on distance to amino acids in the other subunit, where residues whose two closest non-hydrogen atoms are separated by a distance smaller than the sum of their van der Waals radii plus 0.5 Å are considered to be at the interface (coded as 1). Residues whose alpha carbons are within 6 Å of the other chain were considered to be at the interface rim (near the interface, coded as 0.75).

We retrieved a set of 218 high confidence *FCY1* orthologs by combining sequences from Emsembl[77], MetaPhOrs[78], and the Yeast genome order browser[79]. The sequences were aligned using Muscle[80] with default parameters. After visual inspection, three sequences were removed as they were discordant with major aligned blocks, yielding a final set of 215 sequences (Supplementary Dataset 1). The orthologs were aligned with Muscle again, using the stable.py script to keep the same sequence order in the output. Surprisingly, we found a higher than expected proportion of orthologs in bacterial and archaeal genomes. Previous work had hypothesized that cytosine deaminases from the *FCY1* family were exclusive to fungal species, with bacteria usually having orthologs of *E. coli codA*[16]. We instead found over 30 orthologs in bacteria and seven in archaea, suggesting a more complex evolutionary history for this enzyme class. This alignment was then used to calculate normalized evolutionary rates using Rate4site[38] using default parameters. To determine variant occupancy, the positions of the alignment corresponding to the *S. cerevisiae* amino acids were parsed to identify the frequency of each amino acid. The fully annotated dataframe containing log2 fold-changes, 5-FC and cytosine score, as well as all annotations for mutants is provided as Supplementary table 6.

Construction of the *FCY1* mutants validation set

The *FCY1* variant strains tested as part of the validation studies were constructed using the same strategy used to generate the DMS library, starting from BY4742 *fcy1::HphMX4*. We initially selected 80 variants for validation studies, but not all strain constructions attempts were successful, explaining why data is not presented for all of them. To generate mutated donor DNA, we used a fusion PCR strategy where oligonucleotides containing the mutations of interest are used to generate two overlapping amplicons both bearing a mutation starting from genomic DNA extracted from strain BY4742 *FCY1_codon_opt*. The first reaction, which amplified from the five-prime flanking homology arm before the start of the coding sequence up to the mutated amplicon overlap, used oligonucleotides fusion_for and mutant_rev_1 to

80. The second reaction, which amplified from the mutated overlap to the three-prime homology arm after the coding sequence, used oligonucleotides mutant_for_1 to 80 and fusion_rev. Both resulting amplicons are then diluted 1/2,500 and used as template for a second reaction using oligonucleotides fusion_for and fusion_rev to reconstitute the full length *FCY1* sequence.

We then performed one by one CRISPR knock-ins using the same reaction setup as for the DMS library, but using the fusion PCR products as donor DNA. Transformants (four/mutant) were used to inoculate 1.2 mL YPD cultures, grown overnight, and then streaked on YPD. One colony from each streak was resuspended in 200 µL water and spotted on YPD as well as YPD+G418 and YPD+Hyg to check for pCAS loss and successful loss of the *HphMX4* cassette. *FCY1* integration was verified by PCR using oligonucleotides FCY1_alt_A and fusion_rev on quick DNA extractions[81]. PCR positive strains were then validated by Sanger sequencing using fusion_for as the sequencing oligonucleotide. We succeeded in constructing 73/80 mutants. Variant growth rates were measured by performing growth curves in 384 well plates with a 80 µl sterile water border. Cells from overnight pre-cultures in SC media were spun down and resuspended in sterile water at a concentration of 0.4 OD. From these dilutions, 20 µl was used to inoculate either 60 µl 1.33X SC media + 12 µM 5-FC or 60 µl 1.33X SC-ura media + 84 µM cytosine, for a final media concentration of 1X and 0.1 OD cell density. Each strain was tested in duplicate for both conditions: measurements were taken every 15 minutes for 48 hours without agitation on a Tecan Spark (Tecan Life Sciences, Männedorf) with an active Tecool temperature control module. We excluded one mutant due to a technical problem with the plate cover, leading to the final count of 72 mutants. The data and annotations related to validation mutants is presented in supplementary table 7.

Measurement of mutant-WT dimer interaction

Although the Fcy1p-Fcy1p homodimer interaction had already been reported to be detectable using this DHFR-PCA[82], we reconstructed a full set of DHFR strains (s_005 to s_008) and validated that we could observe it. The C-terminal DHFR-tagged strains were constructed by amplifying the DHFR tags from pAG25-DHFR[1,2] and pAG32-DHFR[3] using oligonucleotides from the DHFR collection (DHFR_tag_F and DHFR_tag_R)[39]. The tagged strains were validated by PCR and Sanger sequenced.To generate a N-terminal DHFR[1,2] tagged allele of Fcy1p, we used fusion PCR to amplify the DHFR[1,2]-linker fragment from pGD009[83] (oligonucleotides Nter_DHFR_for and Nter_DHFR_rev) and then fuse it to the codon optimized sequence (oligonucleotides linker_fcy1_for and knock_in_rev) which was amplified from strain *FCY1_codon_opt* (s_004) genomic DNA. The fused amplicon was then used as donor DNA for a CRISPR transformation of *fcy1::HphMX4* with pCAS-Hph. The positive clone was validated by PCR and Sanger sequencing, yielding strain s_009 (*DHFR[1-2]-FCY1_codon_opt*). Mutant alleles tagged in N-terminal with DHFR[1,2] were generated using the same strategy, but using genomic DNA from strain s_009 as template instead. We were able to successfully construct 75/80 N-terminal DHFR[1,2] tagged strains.

To measure changes in PPIs with wild-type Fcy1p, tagged variants were crossed with BY4741 *FCY1-DHFR[3]* by mixing 5 µl of a SC-complete DHFR[1,2]-*FCY1* mutant pre-culture with 5 ul of a YPD+Hyg *FCY1-DHFR[3]* pre-culture in 2 ml YPD and incubating ~8 hours at 30 C. After incubation, 5 µl/mating was spotted on SC-lys-met + Hyg (pH buffered) media and incubated overnight at 30°C to select for diploids. Parental strains were also spotted in parallel to control for background growth. Cells from each spot were resuspended in sterile water and spotted a second time on SC-lys-met+Hyg to perform a second round of selection. From these spots, we inoculated SC-lys-met pre-cultures that were grown overnight. The next morning, the cells were spun down and resuspended in sterile water at a concentration of 0.4 OD. From these dilutions, 8 µl was used to inoculate either 72 µl 1.1X MTX media or 72 µl 1.1X DMSO media (see supplementary Table 5), for a final media concentration of 1X and a 0.1 OD cell density. Each strain was tested in duplicate for both conditions: measurements were taken every 15 minutes for 48 hours without agitation in a Tecan Spark plate as described above.

Dose-response curves and growth rates measurements in multiple 5-FC and cytosine concentration

The dose-response growth curves were generated by measuring the maximum growth rate of BY4742 in 20 different concentrations of 5-FC or cytosine in SC and SC-ura media respectively. The range of concentrations tested was specific to the compound to ensure proper coverage of the dynamic range. The growth rates were measured in triplicate, with each replicate starting from a different colony. The values used to fit the Hill equation are presented as supplementary Table 8.

We measured the growth rates of Fcy1p mutants in 96 wells plates that contained the 73 Sanger validated mutants, of which 13 were present in duplicate. The plate also included 4 replicates of *FCY1_codon_opt* and 4 replicates of BY4742, as well as two blank wells with culture media and no cells. The growth curves were started by diluting overnight pre-cultures and inoculating 180 1.11X SC or SC -ura media containing different amounts of 5-FC or cytosine. The optical densities were manually read for each plate using a Tecan Spark with an active Tecool temperature control module at regular intervals for 24h, at which point the cultures were fully resuspended and a final datapoint was acquired. The AUC measured for the wild-type strains (*FCY1_codon_opt* and BY4742) clustered together in all assay conditions (Figure S18A), and AUCs for replicate Fcy1p mutant cultures were well correlated across all conditions (ρ=0.8-0.99, n=13 mutants) except in SC media without 5-FC, which is expected to be non-selective for Fcy1p function (Figure S18B). For mutants present in duplicate, the values for each replicate were averaged, and the WT AUC was obtained by taking the median of the 8 WT measurements. Data from variant T86M was excluded from figure 5 because no cytosine growth rate from the validations was available. AUC values used to generate Figure 4D and Figure S18 are available as supplementary Table 9.

Acknowledgements

The authors thank members of the Landry lab for helpful discussions, in particular Daniel Evans-Yamamoto, Mathieu Hénault, Florian Mattenberger and Romain Durand. This work was supported by the Canadian Institutes of Health Research Foundation grant 387697 to C.R.L. and a Vanier graduate scholarship to P.C.D, as well as by the National Science and Engineering Research Council through the EvoFunPath CREATE grant (555337-2021) and by FRQNT through team grant (2022-PR-298169) and a PBEEE scholarship to A.F.C. Figure 1 was generated using Biorender.

References

1. Fisher, M. C. *et al.* Threats Posed by the Fungal Kingdom to Humans, Wildlife, and Agriculture. *MBio* **11**, (2020).
2. Fisher, M. C., Hawkins, N. J., Sanglard, D. & Gurr, S. J. Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science* **360**, 739–742 (2018).
3. Nash, A. *et al.* MARDy: Mycology Antifungal Resistance Database. *Bioinformatics* **34**, 3233–3234 (2018).
4. Ksiezopolska, E. *et al.* Narrow mutational signatures drive acquisition of multidrug resistance in the fungal pathogen Candida glabrata. *Curr. Biol.* **0**, (2021).
5. Bryce Taylor, M. *et al.* yEvo: Experimental evolution in high school classrooms selects for novel mutations and epistatic interactions that impact clotrimazole resistance in S. cerevisiae. *bioRxiv* 2021.05.02.442375 (2021) doi:10.1101/2021.05.02.442375.
6. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat. Rev. Microbiol.* **8**, 260–271 (2010).
7. Gerstein, A. C., Lo, D. S. & Otto, S. P. Parallel genetic changes and nonparallel gene-environment interactions characterize the evolution of drug resistance in yeast. *Genetics* **192**, 241–252 (2012).
8. Yang, F. *et al.* The fitness costs and benefits of trisomy of each Candida albicans chromosome. *Genetics* **218**, (2021).
9. Kanafani, Z. A. & Perfect, J. R. Antimicrobial resistance: resistance to antifungal agents: mechanisms and clinical impact. *Clin. Infect. Dis.* **46**, 120–128 (2008).
10. Iyer, K. R., Revie, N. M., Fu, C., Robbins, N. & Cowen, L. E. Treatment strategies for cryptococcal infection: challenges, advances and future outlook. *Nat. Rev. Microbiol.* **19**, 454–466 (2021).
11. Flowers, S. A., Colón, B., Whaley, S. G., Schuler, M. A. & Rogers, P. D. Contribution of clinically derived mutations in ERG11 to azole resistance in Candida albicans.

Antimicrob. Agents Chemother. **59**, 450–460 (2015).

12. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–338 (2003).
13. World Health Organization Model List of Essential Medicines – 22nd List, 2021. Geneva: World Health Organization; 2021 (WHO/MHP/HPS/EML/2021.02). Licence: CC BY-NC-SA 3.0 IGO.
14. Edlind, T. D. & Katiyar, S. K. Mutational analysis of flucytosine resistance in Candida glabrata. *Antimicrob. Agents Chemother.* **54**, 4733–4738 (2010).
15. Erbs, P., Exinger, F. & Jund, R. Characterization of the Saccharomyces cerevisiae FCY1 gene encoding cytosine deaminase and its homologue FCA1 of Candida albicans. *Curr. Genet.* **31**, 1–6 (1997).
16. Ireton, G. C., Black, M. E. & Stoddard, B. L. The 1.14 A crystal structure of yeast cytosine deaminase: evolution of nucleotide salvage enzymes and implications for genetic chemotherapy. *Structure* **11**, 961–972 (2003).
17. Kievit, E. *et al.* Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts. *Cancer Res.* **59**, 1417–1421 (1999).
18. Erbs, P. *et al.* In vivo cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene. *Cancer Res.* **60**, 3813–3822 (2000).
19. Korkegian, A., Black, M. E., Baker, D. & Stoddard, B. L. Computational thermostabilization of an enzyme. *Science* **308**, 857–860 (2005).
20. Stolworthy, T. S. *et al.* Yeast cytosine deaminase mutants with increased thermostability impart sensitivity to 5-fluorocytosine. *J. Mol. Biol.* **377**, 854–869 (2008).
21. Warren, T. D., Patel, K., Rivera, J. L., Eshleman, J. R. & Ostermeier, M. Comprehensive mutagenesis on yeast cytosine deaminase yields improvements in 5-fluorocytosine toxicity in HT1080 cells. *AIChE J.* **66**, e16688 (2020).
22. Breslow, D. K. *et al.* A comprehensive strategy enabling high-resolution functional

analysis of the yeast genome. *Nat. Methods* **5**, 711–718 (2008).

23. Qian, W., Ma, D., Xiao, C., Wang, Z. & Zhang, J. The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Rep.* **2**, (2012).
24. Becker, J. M. *et al.* Pathway analysis of Candida albicans survival and virulence determinants in a murine infection model. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 22044–22049 (2010).
25. Hartzog, P. E., Nicholson, B. P. & McCusker, J. H. Cytosine deaminase MX cassettes as positive/negative selectable markers in Saccharomyces cerevisiae. *Yeast* **22**, 789–798 (2005).
26. Li, A., Acevedo-Rocha, C. G. & Reetz, M. T. Boosting the efficiency of site-saturation mutagenesis for a difficult-to-randomize gene by a two-step PCR strategy. *Appl. Microbiol. Biotechnol.* **102**, 6095–6103 (2018).
27. Biot-Pelletier, D. & Martin, V. J. J. Seamless site-directed mutagenesis of the Saccharomyces cerevisiae genome using CRISPR-Cas9. *J. Biol. Eng.* **10**, 6 (2016).
28. Dionne, U. *et al.* Protein context shapes the specificity of SH3 domain-mediated interactions in vivo. *Nat. Commun.* **12**, 1597 (2021).
29. Eddy, A. A. Expulsion of uracil and thymine from the yeast Saccharomyces cerevisiae: contrasting responses to changes in the proton electrochemical gradient. *Microbiology* **143 (Pt 1)**, 219–229 (1997).
30. Kurtz, J. E., Exinger, F., Erbs, P. & Jund, R. New insights into the pyrimidine salvage pathway of Saccharomyces cerevisiae: requirement of six genes for cytidine metabolism. *Curr. Genet.* **36**, 130–136 (1999).
31. Fujimura, H. Growth inhibition of Saccharomyces cerevisiae by the immunosuppressant leflunomide is due to the inhibition of uracil uptake via Fur4p. *Mol. Gen. Genet.* **260**, 102–107 (1998).
32. Després, P. C., Dubé, A. K., Nielly-Thibault, L., Yachie, N. & Landry, C. R. Double Selection Enhances the Efficiency of Target-AID and Cas9-Based Genome Editing in Yeast. *G3* **8**, 3163–3171 (2018).

33. Sadhu, M. J. *et al.* Highly parallel genome variant engineering with CRISPR-Cas9. *Nat. Genet.* **50**, 510–514 (2018).
34. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
35. Wang, J. *et al.* Role of glutamate 64 in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase. *Biochemistry* **51**, 475–486 (2012).
36. Ivankov, D. N., Finkelstein, A. V. & Kondrashov, F. A. A structural perspective of compensatory evolution. *Curr. Opin. Struct. Biol.* **26**, 104–112 (2014).
37. Schymkowitz, J. *et al.* The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–8 (2005).
38. Mayrose, I., Graur, D., Ben-Tal, N. & Pupko, T. Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol. Biol. Evol.* **21**, 1781–1791 (2004).
39. Tarassov, K. An in vivo map of the yeast protein interactome. *Science* **320**, (2008).
40. Freschi, L., Torres-Quiroz, F., Dubé, A. K. & Landry, C. R. qPCA: a scalable assay to measure the perturbation of protein--protein interactions in living cells. *Mol. Biosyst.* **9**, 36–43 (2013).
41. Wagih, O. A resource of variant effect predictions of single nucleotide variants in model organisms. *Mol. Syst. Biol.* **14**, (2018).
42. Chang, A. *et al.* BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Res.* **49**, D498–D508 (2021).
43. Hottes, A. K. *et al.* Bacterial adaptation through loss of function. *PLoS Genet.* **9**, e1003617 (2013).
44. Kvitek, D. J. & Sherlock, G. Whole genome, whole population sequencing reveals that loss of signaling networks is the major adaptive strategy in a constant environment. *PLoS Genet.* **9**, e1003972 (2013).
45. Jain, N. *et al.* Loss of allergen 1 confers a hypervirulent phenotype that resembles

mucoid switch variants of Cryptococcus neoformans. *Infect. Immun.* **77**, 128–140 (2009).

46. Wang, Z. & Zhang, J. Abundant indispensable redundancies in cellular metabolic networks. *Genome Biol. Evol.* **1**, 23–33 (2009).
47. Goodwin, M. L. & Drew, R. H. Antifungal serum concentration monitoring: an update. *J. Antimicrob. Chemother.* **61**, 17–25 (2008).
48. Oliver, J. D. *et al.* F901318 represents a novel class of antifungal drug that inhibits dihydroorotate dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12809–12814 (2016).
49. Hoenigl, M. *et al.* The Antifungal Pipeline: Fosmanogepix, Ibrexafungerp, Olorofim, Opelconazole, and Rezafungin. *Drugs* (2021) doi:10.1007/s40265-021-01611-0.
50. Verweij, P. E., Te Dorsthorst, D. T. A., Janssen, W. H. P., Meis, J. F. G. M. & Mouton, J. W. In vitro activities at pH 5.0 and pH 7.0 and in vivo efficacy of flucytosine against Aspergillus fumigatus. *Antimicrob. Agents Chemother.* **52**, 4483–4485 (2008).
51. Gsaller, F. *et al.* Mechanistic Basis of pH-Dependent 5-Flucytosine Resistance in Aspergillus fumigatus. *Antimicrob. Agents Chemother.* **62**, (2018).
52. Stiffler, M. A., Hekstra, D. R. & Ranganathan, R. Evolvability as a function of purifying selection in TEM-1 β-lactamase. *Cell* **160**, 882–892 (2015).
53. Esposito, D. *et al.* MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biol.* **20**, 223 (2019).
54. McManus, B. A., Moran, G. P., Higgins, J. A., Sullivan, D. J. & Coleman, D. C. A Ser29Leu substitution in the cytosine deaminase Fca1p is responsible for clade-specific flucytosine resistance in Candida dubliniensis. *Antimicrob. Agents Chemother.* **53**, 4678–4685 (2009).
55. Chang, Y. C. *et al.* Moderate levels of 5-fluorocytosine cause the emergence of high frequency resistance in cryptococci. *Nat. Commun.* **12**, 3418 (2021).
56. Brachmann, C. B. *et al.* Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene

disruption and other applications. *Yeast* **14**, 115–132 (1998).

57. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, (2007).
58. Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
59. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* **15**, 1541–1553 (1999).
60. DeLuna, A., Springer, M., Kirschner, M. W. & Kishony, R. Need-based up-regulation of protein levels in response to deletion of their duplicate genes. *PLoS Biol.* **8**, e1000347 (2010).
61. Casadaban, M. J. & Cohen, S. N. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J. Mol. Biol.* **138**, 179–207 (1980).
62. Yachie, N. *et al.* Pooled-matrix protein interaction screens using Barcode Fusion Genetics. *Mol. Syst. Biol.* **12**, 863 (2016).
63. Andrews, S. *FastQC: A quality control analysis tool for high throughput sequencing data*. (Github).
64. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
65. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9**, 90–95 (2007).
66. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
67. Reback, J. *et al. pandas-dev/pandas: Pandas 1.3.4*. (2021). doi:10.5281/zenodo.5574486.
68. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
69. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

70. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformaticshttps://doi.org/10.1186/1471-2105-13-31 (2012).
71. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
72. Rice, P., Longden, L. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet.https://doi.org/10.1016/S0168-9525(00)02024-2 (2000).
73. Ryan, O. W., Poddar, S. & Cate, J. H. D. Crispr–cas9 genome engineering in Saccharomyces cerevisiae cells. *Cold Spring Harb. Protoc.* **2016**, (2016).
74. Amberg, D. C., Burke, D. J., Burke, D. & Strathern, J. N. *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. (CSHL Press, 2005).
75. Marchant, A. *et al.* The role of structural pleiotropy and regulatory evolution in the retention of heteromers of paralogs. *Elife* **8**, (2019).
76. Usmanova, D. R. *et al.* Self-consistency test reveals systematic bias in programs for prediction change of stability upon mutation. *Bioinformatics* **34**, 3653–3658 (2018).
77. Howe, K. L. *et al.* Ensembl 2021. *Nucleic Acids Res.* **49**, D884–D891 (2021).
78. Chorostecki, U., Molina, M., Pryszcz, L. P. & Gabaldón, T. MetaPhOrs 2.0: integrative, phylogeny-based inference of orthology and paralogy across the tree of life. *Nucleic Acids Res.* **48**, W553–W557 (2020).
79. Byrne, K. P. & Wolfe, K. H. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* **15**, 1456–1461 (2005).
80. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
81. Lõoke, M., Kristjuhan, K. & Kristjuhan, A. Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques* **50**, 325–328 (2011).
82. Schlecht, U., Miranda, M., Suresh, S., Davis, R. W. & St Onge, R. P. Multiplex assay for condition-dependent changes in protein-protein interactions. *Proc. Natl. Acad. Sci. U. S.*

A. **109**, 9213–9218 (2012).

83. Diss, G. & Lehner, B. The genetic landscape of a physical interaction. *Elife* **7**, (2018).

84. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

1224 # Figure Legends

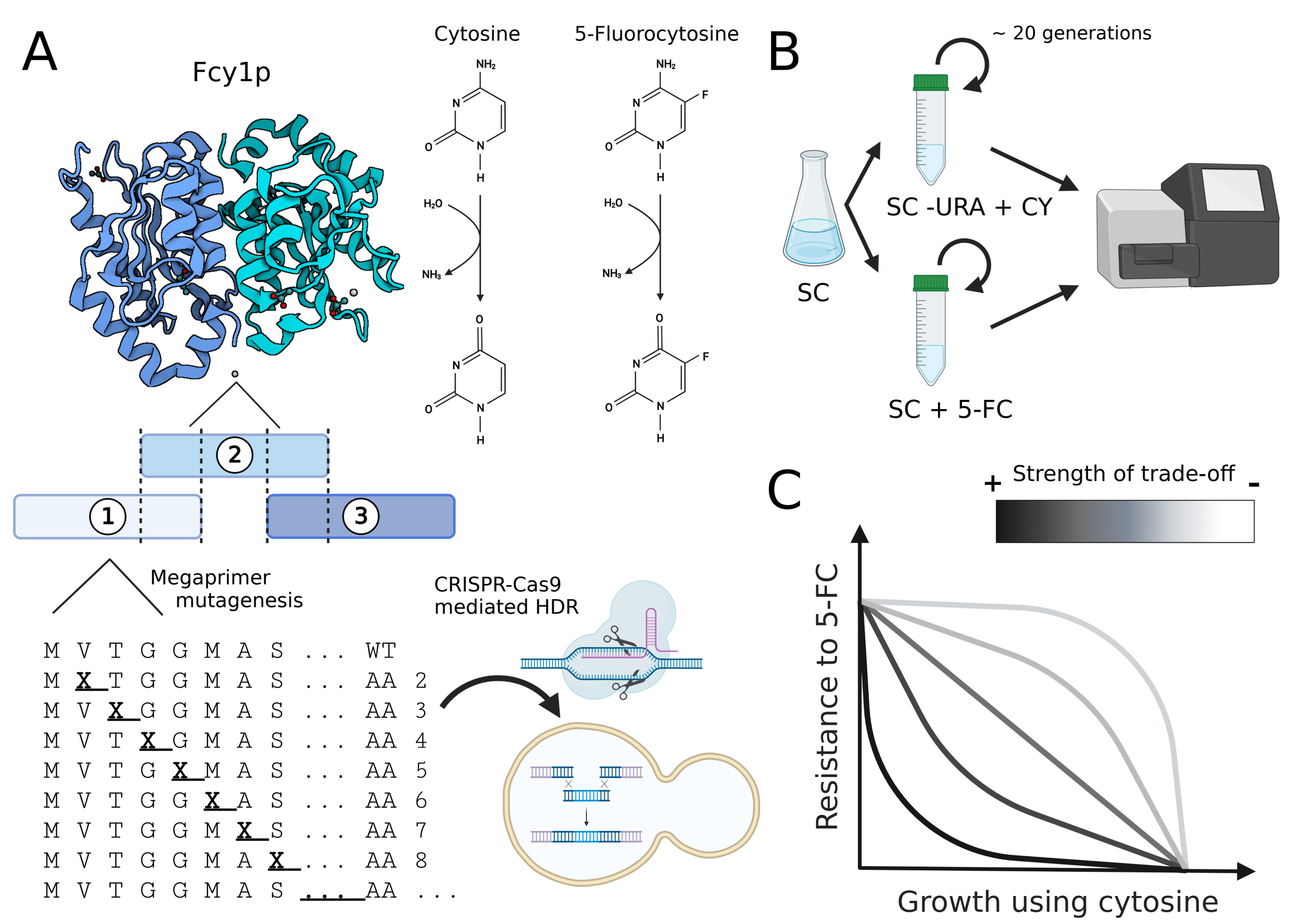
Figure 1: Fcy1p as a model for resistance-function trade-offs. A) Structure of the Fcy1p homodimer (pdb: 1p6o[16]) and two reactions it catalyzes with opposed effects on fitness. To explore Fcy1p's functional landscape, each codon was mutated one by one to all other possibilities to generate a pool of mutants. The mutants were split into three partly overlapping fragments, and studied in three pools: pool 1 covers from position 2-67, pool 2 from 43-110 and pool 3 from 93 to 158. The resulting libraries of *FCY1* variants were then integrated at the native *Saccharomyces cerevisiae* locus using CRISPR-Cas9 genome editing, therefore preserving its endogenous transcriptional regulation. **B)** Workflow for the Fcy1p DMS pooled growth assays. Mutant libraries were expanded by a round of culture in non-selective media before being split into 5-FC and cytosine media where they competed for around ~20 generations. Variant abundance before and after competition was measured by high-throughput sequencing. **C)** Trade-off of different intensities between resistance to 5-FC and growth using cytosine. In a weak trade-off scenario (light grey), resistance mutations have almost no effect on canonical enzyme function. A strong trade-off (dark) instead implies that mutations that lead to resistance necessarily lead to loss of cytosine deamination.

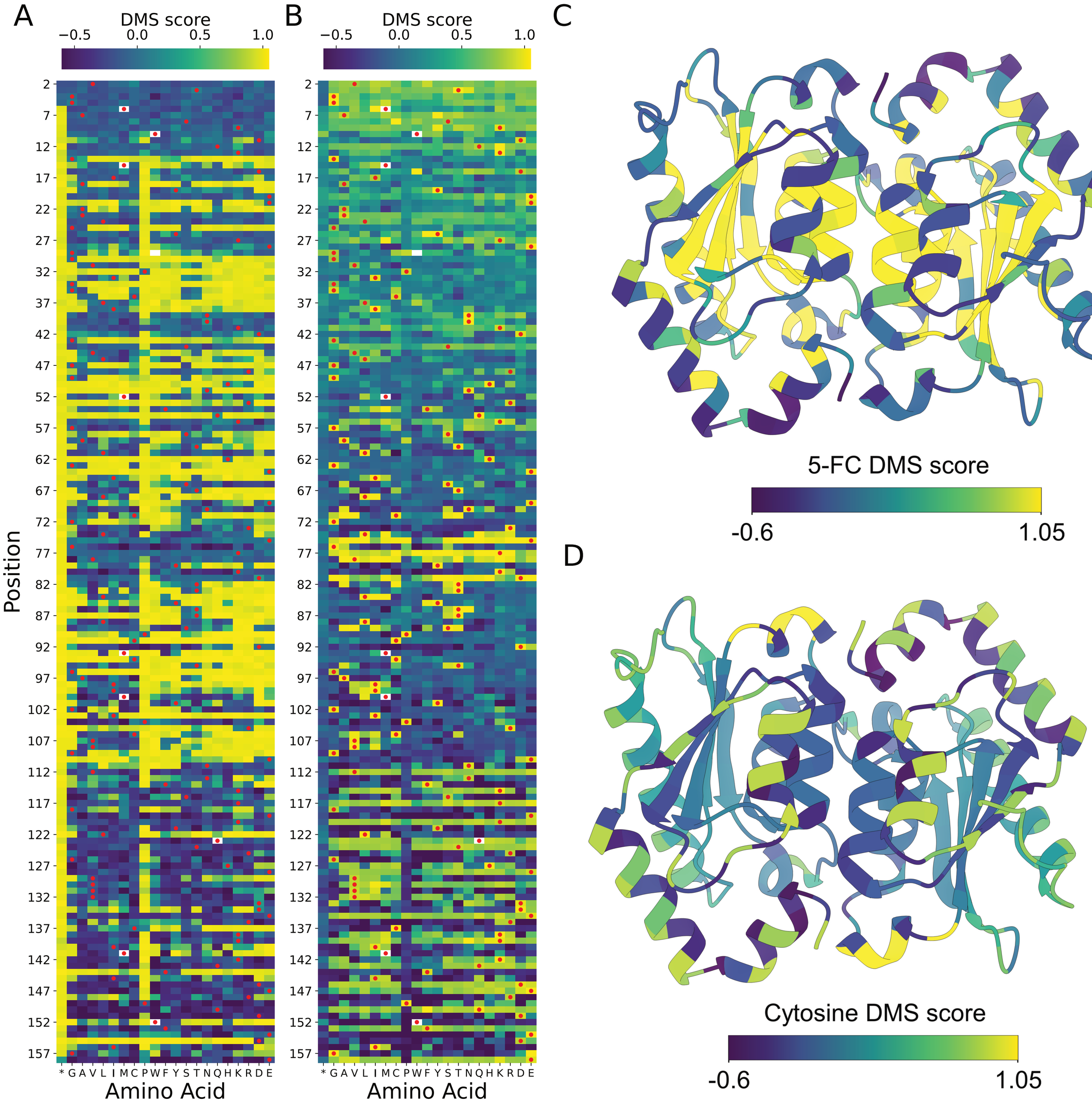
Figure 2: The functional landscape of Fcy1p identifies all resistance mutations. A) Fcy1p functional landscape in 5-FC media and **B)** in cytosine media. Red dots represent the wild-type amino acid at each position. The score represents the DMS mutant $\log_2$ fold-change, scaled by the median of scores for synonymous and nonsense mutations (see methods). The leftmost column on both heatmaps represents the scores for stop-codons (*). **C)** Position-wise median5-FC and **D)** cytosine DMS scores of missense mutants, mapped onto the Fcy1p homodimer (PDB: 1p6o[16]). The color scaling across the four panels is based on the 2th and 98th percentile of the 5-FC scores distribution.

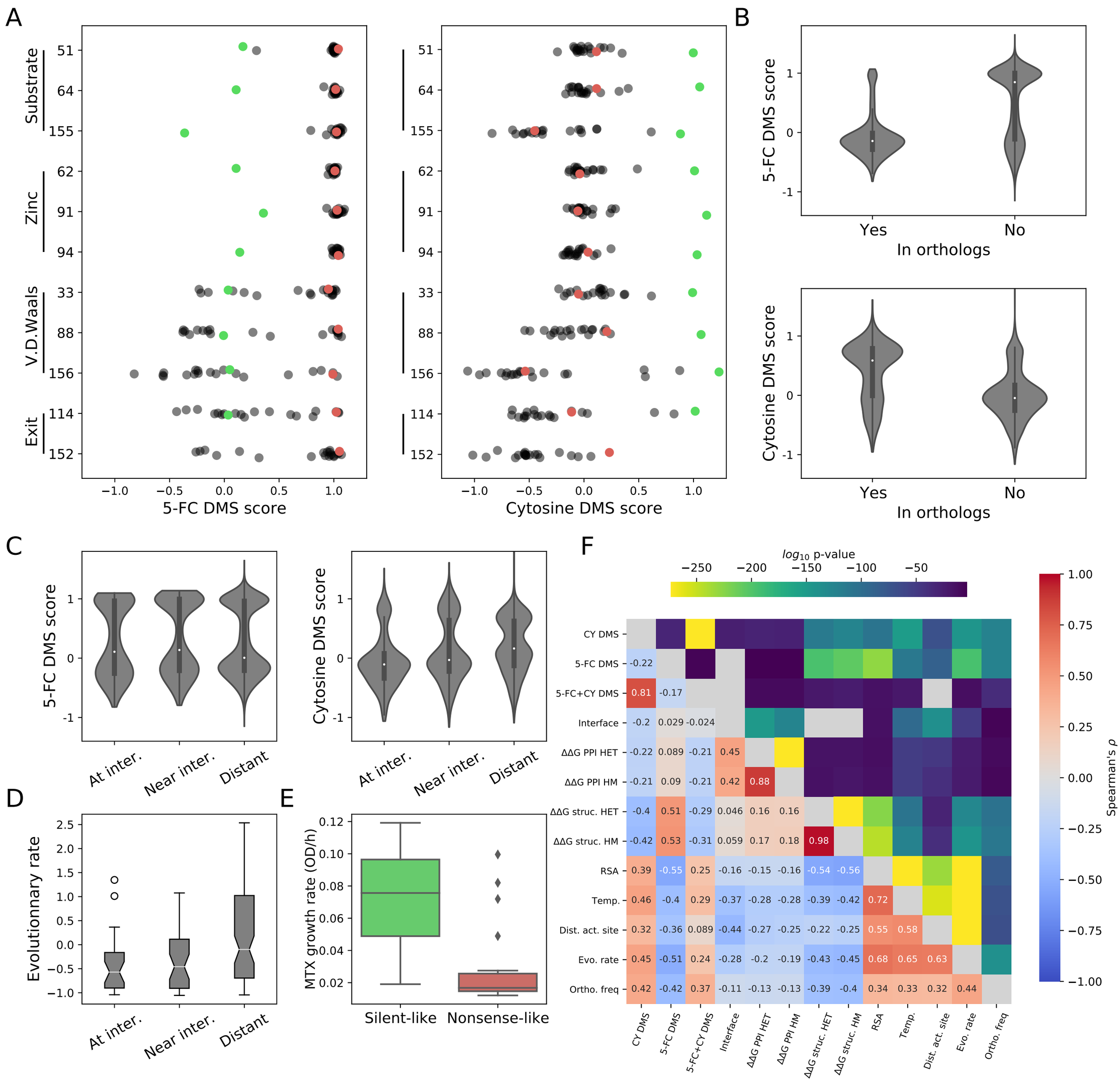
Figure 3: Multiple pathways to 5-FC resistance exist in Fcy1p. A) 5-FC and cytosine DMS scores of mutations at the substrate binding and catalytic residues (substrate), zinc cofactor binding residues (Zinc), residues involved in binding stabilization through Van der Waals interactions (V.d.Waals) and residues involved in guiding the exit of the product (exit). The score for the silent and nonsense variants for each position are shown in green and red respectively **B)** 5-FC and cytosine DMS scores of missense variants as a function of their presence in orthologous sequences (n=1080 observed, n=1902 unobserved). The difference between medians was tested using Mood's median test: for 5-FC scores, $p=8.2x10^{-136}$, and for cytosine, $p=2.0x10^{-71}$. **C)** 5-FC and cytosine DMS score distributions for residues as a function of their distance to the dimer interface. n=532 variants at the interface, n=722 variants near the interface, and n=1728 distant variants. **D)** Evolutionary rates of residues at, near or distant to the dimer interface. **E)** The stability of the complex measured by DHFR-PCA for validation mutants clustering with silent or nonsense mutations (n=10 silent-like mutants, n=18 nonsense-like mutants). **F)** Spearman's rank correlation matrix of mutation properties and DMS scores where: Interface is whether the mutation occurs at, near or distant to the interface, ΔΔG PPI HET is the change in interface stability in a WT-mutant heteromer, ΔΔG PPI HM is the change in interface stability in a mutant-mutant homomer, ΔΔG struc. HET is the change in protein structure stability in a WT-mutant heteromer, ΔΔG struc. HM is the change in protein structure stability in a mutant-mutant homomer, RSA is the residue relative solvent accessibility, Temperature (also called β-factor) is a measure of flexibility in the crystal structure, Dist. act. site is the shortest distance between a residue and one of the active site residues, Evo. rate is the evolutionary rate as computed by Rate4Site and Ortho. freq. is the variant frequency in the orthologous sequences. All ΔΔG values were computed *in silico* using FoldX. Correlations that were not significant (Bonferroni-adjusted p-value > 0.05) are shown as grey squares in the upper part of the matrix.

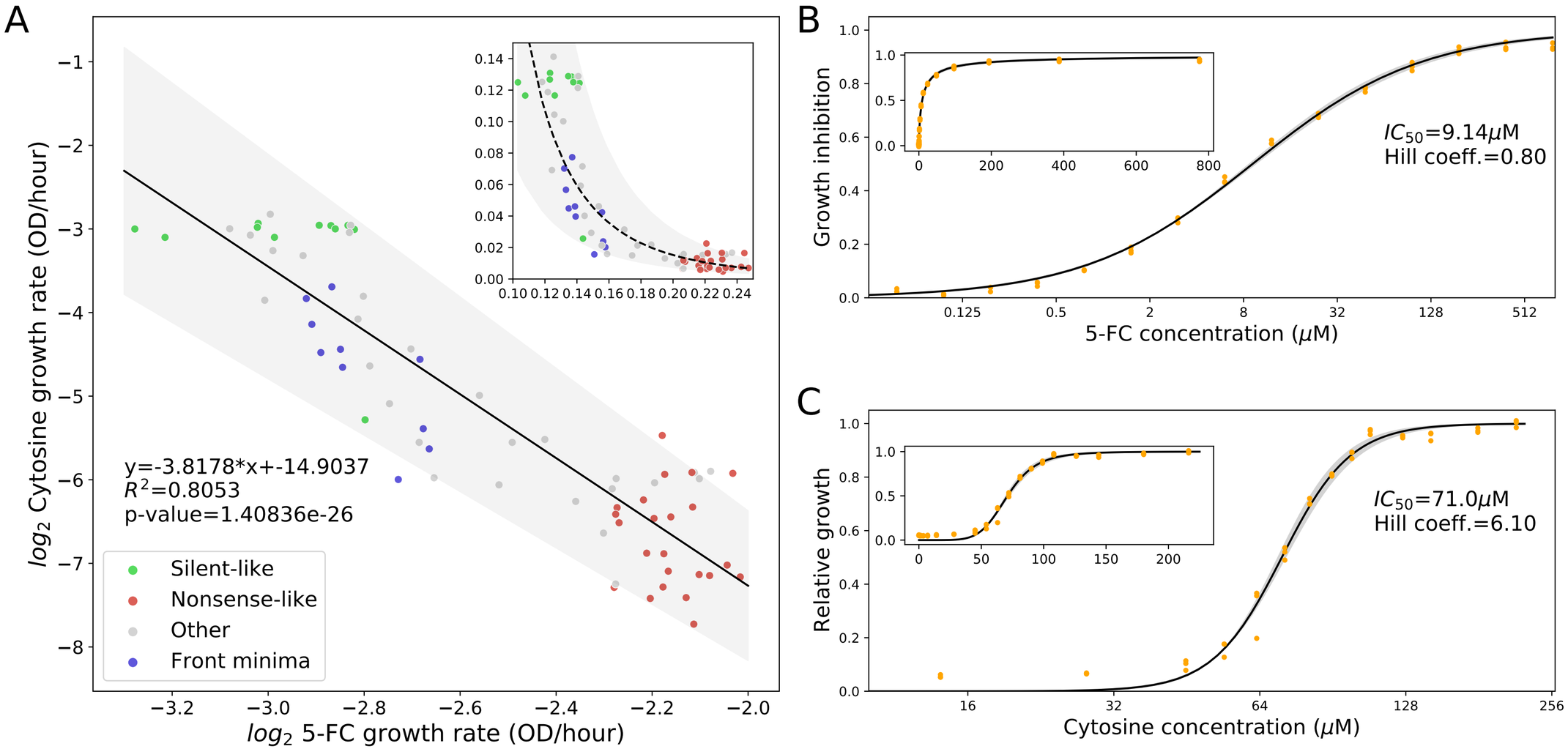
Figure 4: Fcy1p mutants are constrained to a small region of phenotypic space. A) 5-FC and cytosine growth rates of 72 validation mutants in screen conditions. The mutants were assigned to different clusters based on DMS scores. The grey area represents the 95% confidence interval around the log-log linear regression slope fit. The inset shows the values on a linear scale. **B)** Dose-response curve of wild-type *FCY1* to 5-FC and **C)** to cytosine. The Hill equation parameters were estimated using a non-linear least-square fit on relative growth rates, and the grey area represents the 95% confidence interval on IC_{50} and the Hill coefficient. The inset shows the data on a linear scale.

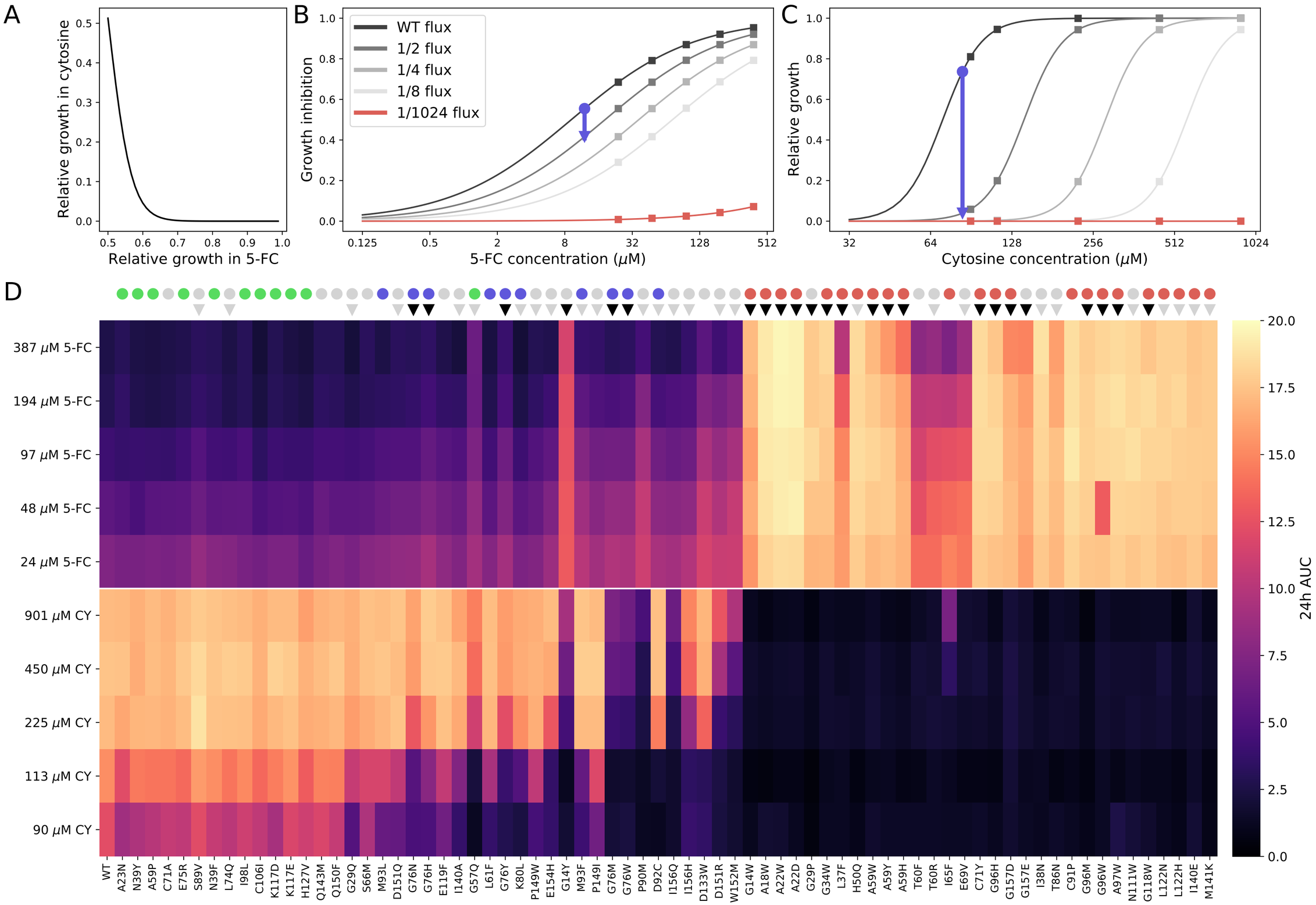
Figure 5 Dose-response parameters predict the properties of Fcy1p mutants. A) Relationship between growth in 5-FC and cytosine when concentrations drop beneath the IC_{50} based on the dose-response curves. Growth in cytosine media is abolished so quickly by lower flux that the same change has little effect on 5-FC resistance. **B)** Effect of reduced Fcy1p flux on growth inhibition by 5-FC and **C)** growth in cytosine media based on the dose-response curve parameters. The blue line represents the phenotype of a Fcy1p variant which reduces pathway flux by 50% in validation conditions (12 μM 5-FC or 84 μM cytosine). **D)** Response of validation mutants to higher concentrations of 5-FC and cytosine. The mutants were ordered based on their position along the trade-off front, starting from the top left in Figure 4A (no 5-FC resistance, full growth in cytosine). The colored dots above represent the cluster to which the mutants were assigned based on the DMS data. Mutants annotated with grey triangles are predicted to destabilize the structure by 1-5 Kcal/mol and those with black triangles by >5 Kcal/mol. Data for the wild-type (WT) is the median of 4 replicates of *FCY1_codon_opt* and 4 replicates of BY4742.


A
Fcy1p
Cytosine
5-Fluorocytosine
NH2
N
O
H
F
H2O
NH3
1
2
3
Megaprimer mutagenesis
M V T G G M A S ... WT
M X T G G M A S ... AA 2
M V X G G M A S ... AA 3
M V T X G M A S ... AA 4
M V T G X M A S ... AA 5
M V T G G X A S ... AA 6
M V T G G M X S ... AA 7
M V T G G M A X ... AA 8
M V T G G M A S ... AA ...
CRISPR-Cas9 mediated HDR
B
~ 20 generations
SC
SC -URA + CY
SC + 5-FC
C
Strength of trade-off
+
-
Resistance to 5-FC
Growth using cytosine


A
DMS score
−0.5
0.0
0.5
1.0
B
DMS score
−0.5
0.0
0.5
1.0
Position
2
7
12
17
22
27
32
37
42
47
52
57
62
67
72
77
82
87
92
97
102
107
112
117
122
127
132
137
142
147
152
157
* G A V L I M C P W F Y S T N Q H K R D E
Amino Acid
* G A V L I M C P W F Y S T N Q H K R D E
Amino Acid
C
5-FC DMS score
-0.6
1.05
D
Cytosine DMS score
-0.6
1.05


A
Substrate
Zinc
V.D.Waals
Exit
51
64
155
62
91
94
33
88
156
114
152
−1.0
−0.5
0.0
0.5
1.0
5-FC DMS score
Cytosine DMS score
B
5-FC DMS score
Cytosine DMS score
Yes
No
In orthologs
C
At inter.
Near inter.
Distant
D
Evolutionnary rate
E
MTX growth rate (OD/h)
Silent-like
Nonsense-like
F
log_{10} p-value
−250
−200
−150
−100
−50
CY DMS
5-FC DMS
5-FC+CY DMS
Interface
ΔΔG PPI HET
ΔΔG PPI HM
ΔΔG struc. HET
ΔΔG struc. HM
RSA
Temp.
Dist. act. site
Evo. rate
Ortho. freq
Spearman's ρ
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00


A
y=-3.8178*x+-14.9037
R^2=0.8053
p-value=1.40836e-26
Silent-like
Nonsense-like
Other
Front minima
log_2 Cytosine growth rate (OD/hour)
log_2 5-FC growth rate (OD/hour)
B
IC_{50}=9.14μM
Hill coeff.=0.80
Growth inhibition
5-FC concentration (μM)
C
IC_{50}=71.0μM
Hill coeff.=6.10
Relative growth
Cytosine concentration (μM)


A
Relative growth in cytosine
Relative growth in 5-FC
B
Growth inhibition
5-FC concentration (μM)
WT flux
1/2 flux
1/4 flux
1/8 flux
1/1024 flux
C
Relative growth
Cytosine concentration (μM)
D
387 μM 5-FC
194 μM 5-FC
97 μM 5-FC
48 μM 5-FC
24 μM 5-FC
901 μM CY
450 μM CY
225 μM CY
113 μM CY
90 μM CY
WT A23N N39Y A59P C71A E75R S89V N39F L74Q I98L C106I K117D K117E H127V Q143M Q150F G29Q S66M M93L D151Q G76N G76H E119F I140A G57Q L61F G76Y K80L P149W E154H G14Y M93F P149I G76M G76W P90M D92C I156Q I156H D133W D151R W152M G14W A18W A22W A22D G29P G34W L37F H50Q A59W A59Y A59H T60F T60R I65F E69V C71Y G96H G157D G157E I38N T86N C91P G96M G96W A97W N111W G118W L122N L122H I140E M141K
24h AUC