Counterintuitive structural and functional effects due to naturally occurring mutations targeting the active site of the disease-associated NQO1 enzyme*

Juan Luis Pacheco-García[1], Ernesto Anoz-Carbonell[2], Dmitry S. Loginov[3], Daniel Kavan[3], Eduardo Salido[4], Petr Man[3], Milagros Medina[2] and Angel L. Pey[5] (iD)

1 Departamento de Química Física, Universidad de Granada, Spain
2 Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) (GBsC-CSIC Joint Unit), Universidad de Zaragoza, Spain
3 Institute of Microbiology – BioCeV, Academy of Sciences of the Czech Republic, Vestec, Czech Republic
4 Center for Rare Diseases (CIBERER), Hospital Universitario de Canarias, Universidad de la Laguna, Tenerife, Spain
5 Departamento de Química Física, Unidad de Excelencia en Química Aplicada a Biomedicina y Medioambiente e Instituto de Biotecnología, Universidad de Granada, Spain

Our knowledge on the genetic diversity of the human genome is exponentially growing. However, our capacity to establish genotype–phenotype correlations on a large scale requires a combination of detailed experimental and computational work. This is a remarkable task in human proteins which are typically multifunctional and structurally complex. In addition, mutations often prevent the determination of mutant high-resolution structures by X-ray crystallography. We have characterized here the effects of five mutations in the active site of the disease-associated NQO1 protein, which are found either in cancer cell lines or in massive exome sequencing analysis in human population. Using a combination of H/D exchange, rapid-flow enzyme kinetics, binding energetics and conformational stability, we show that mutations in both sets may cause counterintuitive functional effects that are explained well by their effects on local stability regarding different functional features. Importantly, mutations predicted to be highly deleterious (even those affecting the same protein residue) may cause mild to catastrophic effects on protein function. These functional effects are not well explained by current predictive bioinformatic tools and evolutionary models that account for site conservation and physicochemical changes upon mutation. Our study also reinforces the notion that naturally occurring mutations not identified as disease-associated can be highly deleterious. Our approach, combining protein biophysics and structural biology tools, is readily accessible to broadly increase our understanding of genotype–phenotype correlations and to improve predictive computational tools aimed at distinguishing disease-prone against neutral missense variants in the human genome.

Introduction

The huge advances in DNA sequencing technologies have uncovered a tremendous genetic diversity in human population. The next step is to be able to distinguish between pathogenic and neutral variants [1–3]. To do so, it is necessary to integrate experimentation to feed and improve computational tools able to carry out accurate and large-scale prediction of mutational effects and genotype–phenotype correlations [4].

The NAD(P)H:quinone oxidoreductase 1 (NQO1) protein is associated with common diseases such as cancer, Alzheimer's, and Parkinson's disease [5]. NQO1 catalyses the FAD-dependent reduction of a large set of quinone substrates, including redox maintenance of vitamins, detoxification of xenobiotics, activation of cancer pro-drugs and regulation of the $NADH/NAD^+$ ratio [6,7]. A schematic representation of the NQO1 catalytic cycle is displayed in Fig. 1A. NQO1 is a dimeric and two-domain enzyme: the N-terminal domain (NTD, residues 1–225) contains the FAD binding site (FBS) and most of the active site residues, while the C-terminal domain (CTD, residues 225–274) completes the active site and the monomer:monomer interface (MMI) [8–12]. Dicoumarol (Dic) is a tight competitive inhibitor of NADH and the substrate [13], for which high-resolution structural information for its binding to NQO1 is available by X-ray crystallography [14]. We have recently shown that ligand binding (FAD and Dic) and mutational effects propagate to long distances in the native state ensemble of NQO1, potentially affecting different functional features in counterintuitive fashions [12,15–20]. Therefore, NQO1 represents a biomedically relevant and challenging system to compare the performance of computational and experimental methods to explain and predict genotype–phenotype on a large scale for a multi-functional protein.

The catalytic competence of NQO1 may be altered upon mutation and affect different functional features: FAD binding, the oxido-reduction reaction pathways, enzyme cooperativity and quantum (tunnelling) effects in the hydride-transfer (HT) reaction [12,19]. As mentioned above, mutational and ligand binding effects may alter the stability of functional sites across the entire protein structure (in some cases over 30 Å) [12,21–23]. In this work, we have studied the effects of five naturally occurring mutations (p.W106R, p.W106C, p.F107C, p.M155I and p.H162N) in the active site of NQO1 affecting four different residues (Fig. 1A). Inspection of a structural model of NQO1$_{dic}$ (with FAD and Dic bound) indicates that all mutations may affect the interactions with the cofactor, NAD(P)H and/or the substrate and thus, the capacity of the enzyme to achieve catalytically competent states (Fig. 1B,C). W106 is close to the FAD flavin ring (2.8 Å for flavin ring N5 to backbone N of W106) and packs against a hydrophobic pocket from the adjacent monomer (Fig. 1B,C), while F107 is also close to the FAD (2.8 Å for flavin ring O4 to backbone N of F107). Thus, mutations p.W106R and p.W106C would affect the stability of the active site (either by introducing a positive charge and a cavity, respectively, in a hydrophobic environment), while mutations at both W106 and F107 might also influence the flavin redox potential. All the mutated residues are strictly conserved in mammalian NQO1 sequences (Fig. 1C) with different degrees of conservation in the overall sequence (down to 78.5% in sequence identity vs. human NQO1). Interestingly, NQO1 sequences from reptiles (with a slightly lower identity, 66.7–70.3%) show a consistent divergence at the H162 site (actually to the active-site mutation p.H162N; Sequence IDs: XP_033014806.1; XP_044283688.1, XP_005310627.1 and XP_044843791.1). Three of these rare mutations in the human NQO1, [p.W106C (c.318G>C), p.M155I (c.465G>A) and p.H162N (c.484C>A)] are found in one or two cancer somatic cell lines in the COSMIC database (accessed by 29 June 2022; https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=NQO1#variants), whereas the two remaining rare mutations [p.W106R (c.316T>C) and p.F107C (c.320T>G)] appear in the human population with no clear association with disease (with allelic frequencies of 2.8×10^{-5} and 1.8×10^{-5}, respectively; according to The Genome Aggregation Database, gnomAD, accessed by 29 June 2022; https://gnomad.broadinstitute.org/gene/ENSG00000181019?dataset=gnomad_r2_1,). Based on different bioinformatic predictive tools, all mutations are expected to be largely deleterious except p.H162N (Table 1). Since the mutated sites are highly conserved, this set of naturally occurring mutations in the active site of NQO1 should affect function mostly due to the changes in charge, polarity and hydrophobicity [24].

Results and Discussion

All active-site variants retain the fold of the wild-type protein but some affect flavin-adenine dinucleotide content

The WT and active-site mutants were purified by IMAC to a high purity (Fig. 2A). None of the mutations had large effects on the expression levels as judged from NQO1 protein yields (between 50% and 70% of WT levels; Fig. 2A). Inspection of the near UV–visible absorption spectra indicated that all variants contained FAD bound, but the active-site
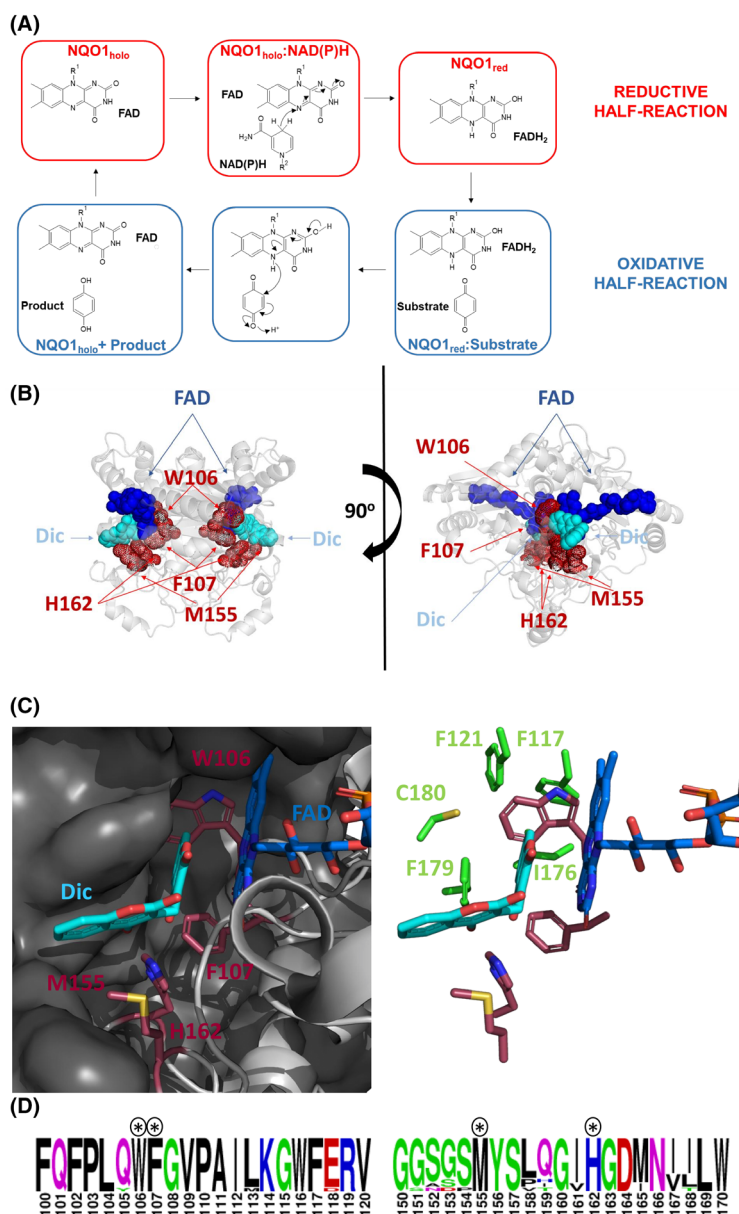
Fig. 1. Structural location and sequence conservation of the residues W106, F107, M155 and H162 in the active site of NQO1. (A) Simplified catalytic cycle of NQO1. The reductive half-reaction (in red) begins with FAD bound to NQO1 (NQO1$_{holo}$) that binds the NAD(P)H coenzyme, and likely through a direct HT reduces the FAD to FADH$_2$ (NQO1$_{red}$). In the oxidative half-reaction, the enzyme binds the substrate that becomes reduced upon oxidation of the FADH$_2$, and upon release of the product regenerates the NQO1$_{holo}$ state. Adapted from Ref. [13]. (B) The display, based on PDB 2F1O [14], shows the mutated residues (in red) and the FAD (in blue) and Dic (in cyan) molecules in human NQO1, to highlight the potential importance of mutations in the active site performance. The two views are from the same display just rotated by 90°. (B) Close view of the NQO1 active site. Mutated residues are shown in CPK coloured sticks (in red). The NQO1 monomer binding the shown FAD within the dimer is in the light grey cartoon, while the neighbouring monomer is in the dark grey surface. The right panel shows a similar display but highlights the residues (in green sticks) of the adjacent monomer that forms the hydrophobic pocket allocating W106. (C) Conservation of these four active site residues among mammalian protein sequences (highlighted with an asterisk). Segments include residues 100–120 and 150–170 of the human sequence. Thirteen sequences were used including different orders. Primates (96.7–100% identity with hNQO1; IDs: NP_000894, NP_001125152.1, XP_030650486.1 and NP_001247927.1); Artiodactyla, (89.4–92.3% identity; IDs: XP_006203814.1, XP_010977513.1, NP_001029707.1 and NP_001153085.1); others (carnivora, Chiroptera, Perissodactyla, Didelphimorphia and Rodentia; 78.1–87.5% identity; IDs:XP_038523128.1, XP_006909876.1, XP_005608432.1, XP_001378228.2 and XP_005345766.1). Sequences were retrieved using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) and the human sequence as query. Panel C was created using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

Table 1. Prediction of the pathogenicity of active-site mutations in NQO1 experimentally characterized in this work. Original scores (OS) were provided by different tools for each variant (in parenthesis) and were normalized (normalized score NS, from 0, neutral, to 1, highly deleterious) using the procedures indicated in the footnotes. The average NS for each variant is reported as NS_{pred} in the *consensus* prediction (and classified as mild, moderate or severe).

Prediction → Variant ↓	POLYPHEN-2 (OS)[a]	SIFT (OS)[b]	CADD (OS)[c]	MUT.ASSESSOR (OS)[a]	REVEL (OS)[a]	METAL R (OS)[a]	CONSENSUS (NS_{pred})
p.W106C (COSV57731335)	Prob.Dam. (1.00)	Deleterious (0.00)	Like. Delet. (31)	High (0.965)	Like. Disease (0.635)	Tolerated (0.198)	Severe (0.69)
p.W106R (rs769655604)	Prob.Dam. (1.00)	Deleterious (0.00)	Like. Delet. (32)	High (0.983)	Like. Disease (0.778)	Tolerated (0.203)	Severe (0.72)
p.F107C (rs760765580)	Prob.Dam. (1.00)	Deleterious (0.01)	Like. Benign (28)	Medium (0.909)	Like. Disease (0.522)	Tolerated (0.135)	Severe (0.64)
p.M155I (COSV57732645)	Prob.Dam. (0.966)	Deleterious (0.04)	Like. Benign (25)	Medium (0.574)	Like. Benign (0)	Tolerated (0.061)	Moderate (0.48)
p.H162N (COSV57731604)	Benign (0.015)	Tolerated (0.85)	Like. Benign (16)	Neutral (0.24)	Like. Benign (0)	Tolerated (0.015)	Mild (0.07)

[a]OS were used as NS from 0 (Benign, Neutral, Likely Benign or Tolerated) to 1 (Probably Damaging, High, Likely Disease or Not tolerated); [b]OS were normalized as binary (Tolerated, 0; Deleterious, 1); [c]OS were linearly normalized from 0 (Benign, Like. Benign, OS 0) to 1 (Deleterious, Like. Delet. original score 99).

mutants reduced this content compared with that of the WT protein (except p.F107C) (Fig. 2B,C). Alterations in the microenvironment of the FAD bound to some of the active-site variants are supported by differences in the near-UV CD spectra of samples saturated with FAD, particularly for the variant p.W106R (Fig. 2B). Additional analyses by circular dichroism (CD) and fluorescence spectroscopies and dynamic light scattering (DLS) supported only mild alterations in the overall conformation of NQO1 upon mutation (Fig. 2D–F). While the secondary structure was not significantly affected (Fig. 2D), the mutation p.W106R increased the fluorescence intensity by about 30% without affecting the shape of the spectra (with an average SCM for all variants of 354.03 ± 0.16 nm). Since Tyr residues are not fluorescent in this range, we suggest that this might originate from either local changes in the microenvironment of the remaining five Trp residues or in the quenching induced by FAD binding. Analysis of the hydrodynamic behaviour revealed some interesting changes, particularly the increased width of the size distribution and the average hydrodynamic radius of p.M155I and p.H162N. We must note that these results are consistent with an increase in hydrodynamic volume of about 15% and 30%, for the p.M155I and p.H162N mutants, respectively. This expanded conformation caused by the p.H162N mutation (in the holo-state) is similar in magnitude to that observed in the WT holo-protein upon withdrawal of FAD [22,25] and may imply an increased population of partially folded states in the native ensemble of holo-p.H162N.

The mutants p.W106R and p.W106C

W106 is involved in the MMI, the FBS and the Dic binding site (DBS). W106 is buried in the structure of WT NQO1 determined by X-ray crystallography ($9.9 \pm 0.6\%$ of accessible surface area, ASA, without considering ligands bound and using the six monomers of PDB 2F1O [14] and the software GETAREA, http://curie.utmb.edu/getarea.html, [26]). The mutations p.W106R and p.W106C are expected to cause large structural perturbations, the former due to the introduction of a positive charge and the latter by creating a cavity in a hydrophobic and buried region of NQO1 (see Ref. [19] for the characterization of some non-natural cavity-making mutations in NQO1).

We first compared the FAD binding affinity of p.W106R and p.W106C with that of WT NQO1 (Fig. 3A,C) by fluorescence titrations. These titrations indicated a much lower affinity in the mutant p.W106R, while the mutant p.W106C showed a similar K_d value for FAD binding than that of the WT protein. Titrations using CD spectroscopy (Fig. 3B) allowed to estimate of a K_d for FAD binding of 7400 ± 4000 nM (about 500-fold higher than that of WT NQO1). To the best of our knowledge, this mutant shows the lowest affinity for FAD ever reported for a single NQO1 missense variant, decreasing by 3.7 kcal·mol⁻¹ the apparent binding free energy.

We have determined the changes in the structural stability of the protein caused by the p.W106R and p.W106C mutations using HDX-MS (Figs S1, S2 and S4). We first looked at the effects on the entire protein. The effects are remarkably different for both
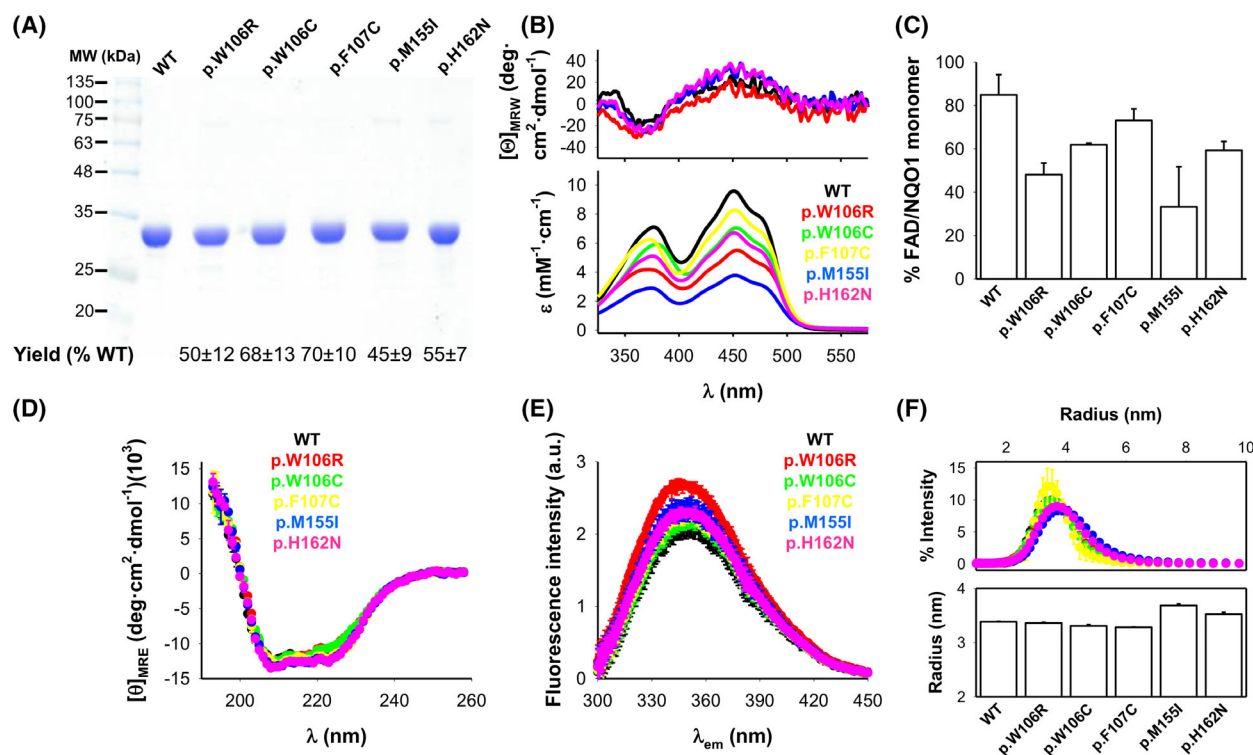
Fig. 2. Purification, conformation and bound FAD to NQO1 WT and the active-site mutants. (A) SDS/PAGE (12% acrylamide) of purified NQO1 variants. In each lane, 5 μg of protein was loaded; NQO1 protein yields (as mg NQO1 protein·L^{-1} cultures) are also shown as average ± SD from three different protein expression experiments. The amount of NQO1 was determined from UV–visible absorption measurements. (B, C) FAD bound to purified NQO1 proteins by UV–visible absorption and CD spectra. Lower panel B shows the average absorption spectra for each NQO1 variant and the upper panel shows CD spectra for selected variants. Panel C shows the average ± SD of the % of NQO1 monomer with bound FAD. (D) Far-UV CD spectra (average ± SD) for NQO1 variants with a fivefold excess of FAD; (E) emission fluorescence spectra (exc. 280 nm) of NQO1 variants in the presence of fivefold excess of FAD. (F) Hydrodynamic behaviour (by DLS) of NQO1 variants in the presence of a fivefold excess of FAD. The upper panel shows the distribution of the apparent radius corresponding to the NQO1 dimer and the lower panel shows the average hydrodynamic radius. Errors correspond to the SD from three different purifications (panel C) or replicate (panels D–F). Errors in panel B are not displayed for sake of display, but representative relative values (and SD) are shown in panel C.

mutants when we analyse them in the holo- and dic-states (Figs 4A–C, 5A–C and 6A; Figs S5A–C, S6A–C, S7A–C, S8A–C, S9A–C and S10A–C). In the NQO1$_{holo}$ state, the p.W106R mutation caused some local and mild stabilizing/destabilizing effects, whereas p.W106C mildly destabilized the entire structure (Fig. 6A). In the NQO1$_{dic}$ state, destabilization by p.W106R was much more extensive and stronger, affecting almost the whole NTD (40% of the residues of the entire protein; Figs 5B and 6A), whereas p.W106C caused virtually no effects (Fig. 6A). Since NTD is critical for enzymatic function [10,25], it is likely that mutation p.W106R will affect other functional features.

We then analysed the effects of p.W106R on the stability of different functional sites: the MMI, involved in the conformational stability of the protein and communication between monomers during the

catalytic cycle [12,13,20,27], the FBS, essential in the redox reaction and in close contact with W106 and the DBS, where the Dic inhibitor likely occupies partly the NADH/substrate binding site and/or represents a transition state analogue and interacts with W106 [13]. The results obtained paralleled those found for the entire protein. The mutation p.W106R mild–moderately affected the stability of the MMI, FBS and DBS in the NQO1$_{holo}$ state, and these effects were much stronger upon Dic binding (Fig. 6B–D). The mutation p.W106C mildly affected the MMI, FBS or DBS in the NQO1$_{holo}$ state and these effects were essentially abolished in the NQO1$_{dic}$ state (Fig. 6B–D). Thus, the effects of p.W106R and p.W106C are consistent with the large effect of the former mutation on FAD binding affinity and also predict much stronger effects for it on enzyme kinetics than those of p.W106C.
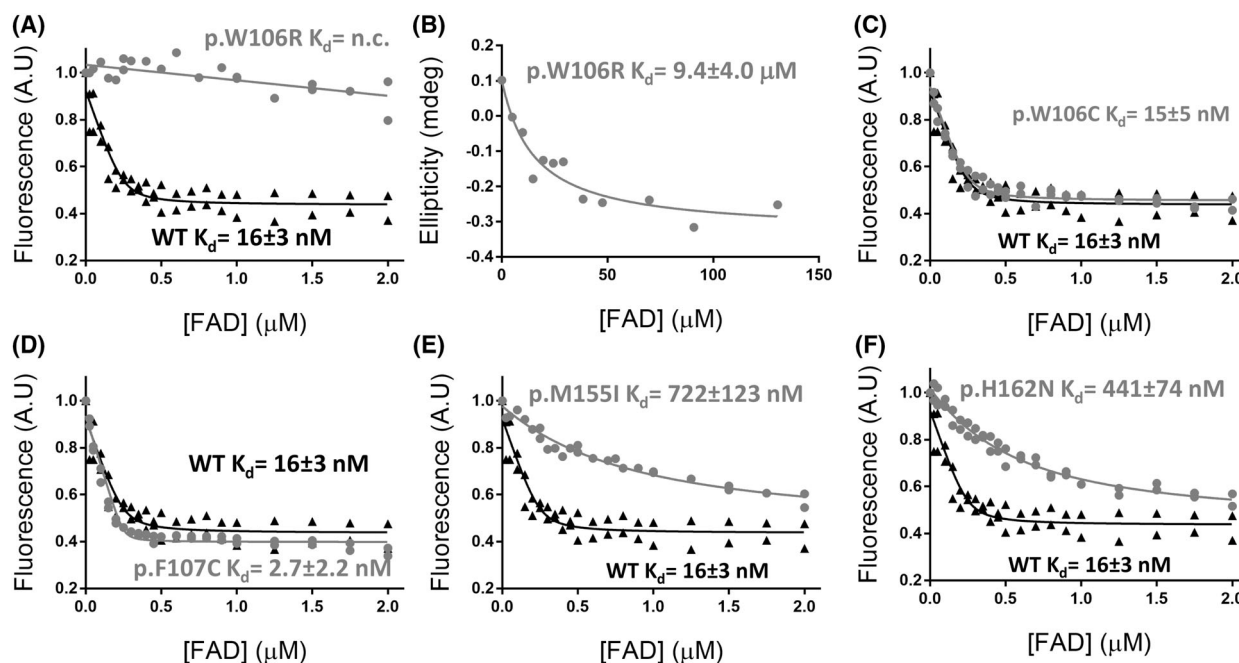
Fig. 3. FAD binding affinity of WT NQO1 and the active site mutants. Titrations of apo-proteins with FAD determined by fluorescence (A, C–F) or circular dichroism (B) spectroscopies. Titrations were carried out in 20 mM K-phosphate, pH 7.4. The temperature was 25 °C. Data came from at least two different titrations, except for panel B (single titration). Lines are best fits for a single type of independent binding sites. N.c., not converged.

The catalytic cycle of NQO1 can be divided into two steps: the reductive and oxidative half-reactions (Fig. 1A), being the former the rate-limiting in the WT enzyme [13]. In addition, pre-steady-state enzyme kinetic analysis of WT NQO1 has revealed the existence of two different pathways, termed *fast* and *slow*, for the reduction of the two FAD molecules within the $NQO1_{holo}$ dimer by NADH, as well as for oxidation of the $FADH_2$ in $NQO1_{red}$ by DCPIP [12,13,19] (Tables 2 and 3; Figs S11A, S12A, S13A, S14A, S15A and S16A). When we compared the kinetics of p.W106R and p.W106C mutants for these two-half reactions, their behaviours were clearly different (Tables 2 and 3; Figs S11B,C, S12B,C and S13B,C). The mutant p.W106R largely perturbed two aspects in the catalytic cycle: (a) The reductive half-reaction of the FAD was about 6000-fold slower for the mutant p.W106R; (b) there is an alteration in the reaction equilibrium and full reduction of the FAD cofactor cannot be achieved. Thus, DCPIP-mediated reoxidation of FAD cannot be measured and this prevents the completion of the catalytic cycle (i.e. reduction of a proper substrate) in the p.W106R mutant. In addition, the spectral properties of species B and C were different in p.W106R than to those in WT NQO1 (with somewhat higher absorbance in the mutant, as

reflected in the lower $\%\Delta A_{450}$; Fig. S12A,B). Interestingly, whereas WT enzyme kinetics showed a hyperbolic dependence on [NADH] (suggesting a transition from EX2 to EX1 mechanism as the NADH is increased sufficiently), no dependence was observed for p.W106R (Fig. 7A,B). When we analysed the mutant p.W106C, much more subtle changes were observed. The p.W106C mutant showed similar spectral properties for the reductive and oxidative half-reactions than those of the WT protein (Figs S11A,C, S12A,C, S13A, C, S14A,B, S15A,B and S16A,B). At stoichiometric concentrations of $NQO1_{holo}$:NADH and $NQO1_{red}$: DCPIP, the reductive half-reaction was slightly accelerated in p.W106C for both *fast* and *slow* pathways (up to 50%), whereas both steps were slowed down in the oxidative half-reaction (two- to sixfold) (Table 2). The NADH-dependence of k_{obs} for the reductive half-reaction in p.W106C also showed subtle changes: the k_{HT} for both *fast* and *slow* pathways was increased by 40–80%, whereas the $K_{d(NADH)}$ was increased in a similar way, leading to small or no changes in the catalytic efficiency ($k_{HT}/K_{d(NADH)}$) in both steps compared with the WT protein (Table 3 and Fig. 7C). Therefore, we must only note the effect of p.W106C on the kinetics of the oxidative half-reduction, that could be associated with the mild structural
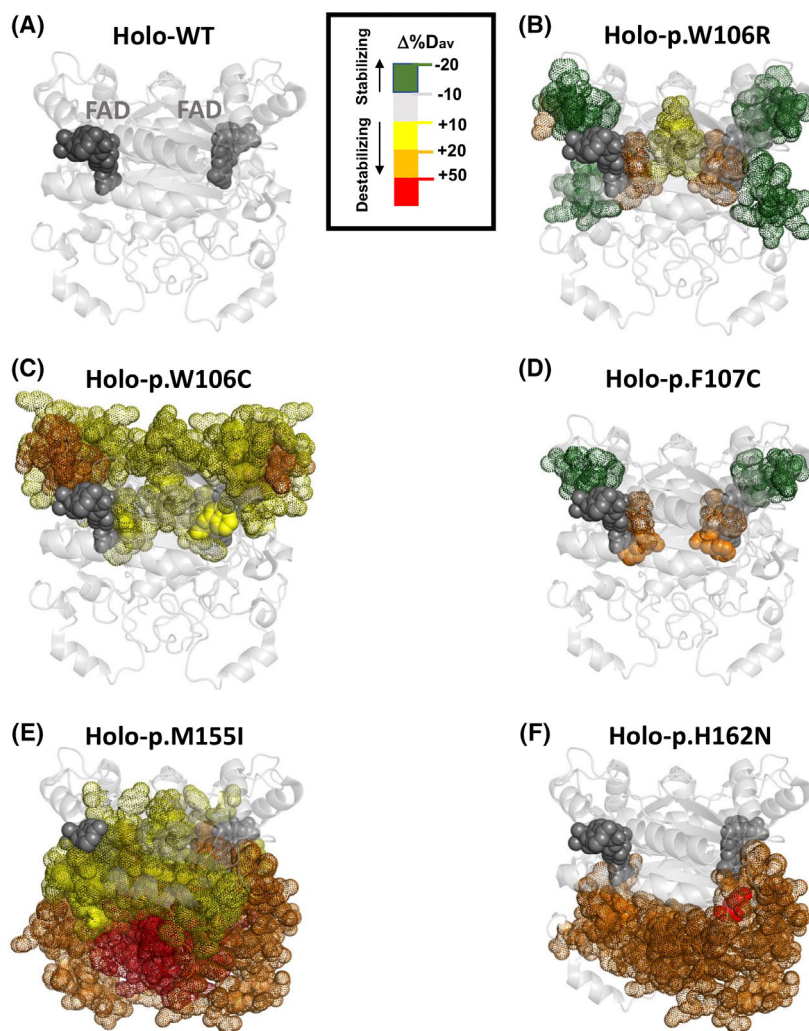
Fig. 4. Changes in local stability of NQO1$_{holo}$ variants determined by HDX-MS. Panel A shows the location of FAD (in dark grey spheres) in the WT structure. Panels B–F show the mutational effects on local stability. Colouring of B–F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [14].

destabilization of the FBS and DBS in the NQO1$_{holo}$ and NQO1$_{dic}$ states of p.W106C (Fig. 6C,D). Regarding the cooperative kinetic behaviour, negative cooperativity was slightly lower in the p.W106C mutant (~ 6) than for WT NQO1 (~ 11) as measured by an operational cooperative index (the quotient of catalytic efficiencies of the *fast/slow* pathways). This lower cooperativity might be explained by altered allosteric communication between active sites due to changes in the stability of the MMI (Fig. 6B; please see also [12,13]).

Altogether, these observations support that p.W106R cannot achieve a competent binding state for catalysis between the nicotinamide of NADH and the isoalloxazine of FAD and this limits the overall kinetics (Appendix S1). In addition, the low binding affinity of p.W106R for FAD (Fig. 3A,B) and the large destabilization of the FBS and DBS (Fig. 6C,D) may lead to the destabilization of NQO1$_{holo}$,

particularly in the reduced state, and further contribute to the slow kinetics observed for the reductive half-reaction [28,29]. A key role of altered active site electrostatics in the catalytic impairment and structural destabilization of the active site by p.W106R is supported by calculations on the active site electrostatic surface potentials (Fig. S17) (note the R106 residue, with a strong basic character and a pK_a in the range 13–14 will be always protonated). The effects of p.W106C in the active site stability and performance were much milder. These results nicely illustrate how two different non-conservative mutations at the same residue of the active site can lead to drastically different effects on protein function.

The mutant p.F107C

F107 is involved in the MMI, FBS and DBS. F107 is buried in the structure (7.6 ± 3.5% of ASA, using the

Fig. 5. Changes in local stability of NQO1$_{dic}$ variants determined by HDX-MS. Panel A shows the location of FAD (in dark grey spheres) and Dic (in blue spheres) in the WT structure. Panels B–F show the mutational effects on local stability. Colouring of B–F follows the scale shown as an inset. The structural model used for display was PDB code 2F1O [14].
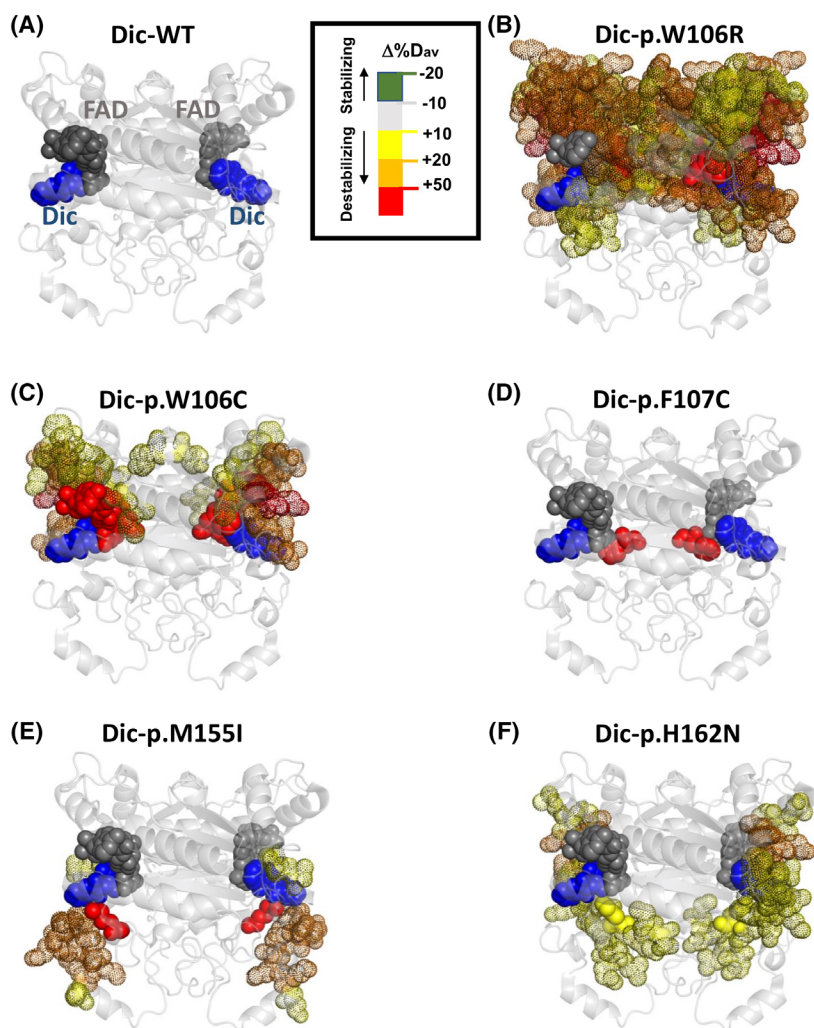
same procedure described for W106). The mutation p.F107C would cause large structural perturbation creating a cavity and introducing a more polar residue [30,31]. The mutation p.F107C did not affect FAD binding affinity (Fig. 3D; please note that the affinity of WT is close to the limit of detection of the technique, and thus we cannot state that the affinity of p.F106C is higher). These results are striking since F107 is located in the FBS and is clearly a non-conservative mutation. HDX-MS analysis revealed a slight destabilization of the FBS and DBS in the NQO1$_{holo}$ state, with no effect on the NQO1$_{dic}$ state (Figs 4–6; Figs S1, S2, S5D, S6D, S7D, S8D, S9D, S10D and S17). These stability measurements are consistent with little or no effects on FAD binding and predict mild or no effects on enzyme kinetics.

However, enzyme kinetic analysis revealed puzzling results for p.F107C. Despite causing little perturbation on the stability of the active site (possibly the most

noticeable the loss of Van der Waals interactions with the Dic inhibitor), the kinetic analysis revealed some alterations due to p.F107C (Fig. 7D; Figs S11D, S12D, S13D, S14C, S15C and S16C; Tables 2 and 3). At a 1 : 1 ratio of NQO1$_{holo}$ and NADH, this mutation reduced by twofold the k_{obs} for the *fast* pathway for the reductive half-reaction, with no effect on the constant for the *slow* pathway. Nonetheless, in contrast to the WT protein, p.F107C FAD reduction is dominated by the *fast* pathway, yielding essentially identical spectra for the B and C species, and suggesting particular large alterations in the reaction equilibrium of the *slow* reaction pathway (Fig. S12D). Additionally, under similar conditions, the oxidative half-reaction was hardly affected (Table 2; Figs S14D, S15D and S16D). Analysis of the NADH-dependence of the FAD reduction kinetics provided further support to these notions as well as some explanations for these effects. For the *fast* pathway, we observed a

Fig. 6. Mutational effects on the local stability of different functional sites determined by HDX-MS. Effects on the entire NQO1 structure (A), the MMI (B), the FBS (C) and the DBS (D). Amino acids belonging to different functional sites were determined as described [20]. Data are expressed as the % of the residues belonging to each category (A, entire protein, 274 residues; B, MMI, 76 residues; C, FBS, 21 residues; D, DBS, 15 residues; according to Ref. [20]) affected by a given mutation and ligation state (NQO1$_{holo}$ and NQO1$_{dic}$ states are referred to as H and D in the *x*-axis). The colour code shows the sign and magnitude of mutational effects (as Δ% D_{av}) in a given ligation state (H or D).
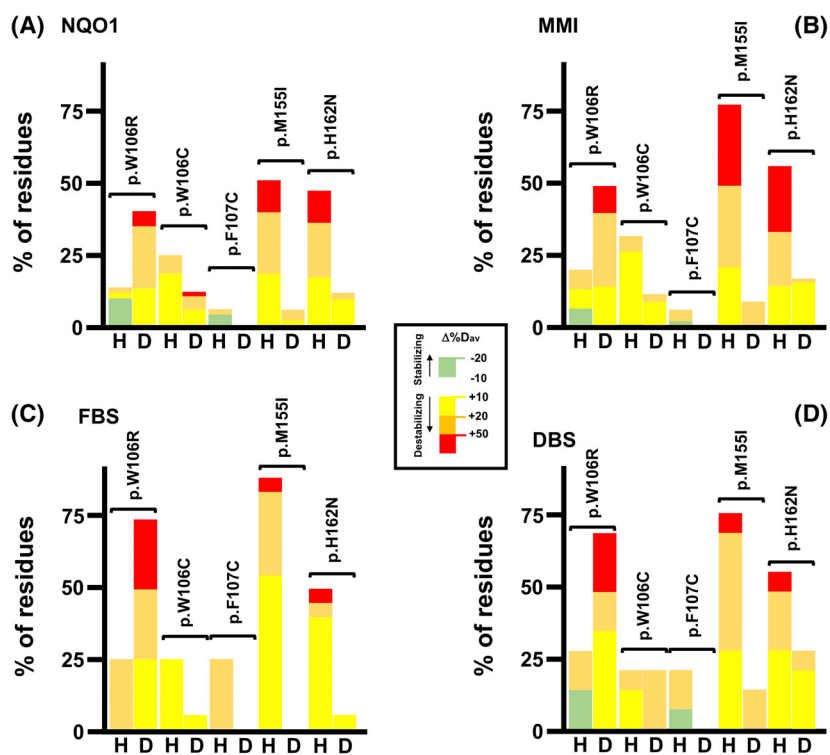
Table 2. Summary of observed rate constants (k_{obs}) for the reductive and oxidative half-reactions involving NQO1. Measurements were carried out in 20 mM HEPES-KOH, pH 7.4 at 6 °C. Evolution of the reaction was followed in the 400–1000 nm wavelength range using stopped-flow equipment with a photodiode array detector ($n > 3$, mean ± SD). 7.5 μM of NQO1$_{holo}$ protein was mixed with 7.5 μM NADH (reductive half-reaction) or 7.5 μM of NQO1$_{red}$ was mixed with 7.5 μM DCPIP (oxidative half-reaction). Primary data and fittings are shown in Figs S11, S13, S14 and S16.

NQO1 variant	Reductive half-reaction		Oxidative half-reaction	
	k_{obs} (s^{-1})		k_{obs} (s^{-1})	
	Fast	Slow	Fast	Slow
WT	78 ± 1	7.2 ± 0.3	> 500	160 ± 14
p.W106R	$1.3 ± 0.2 × 10^{-2}$	$3.3 ± 0.7 × 10^{-3}$	N.d.[a]	N.d.[a]
p.W106C	83 ± 3	12.0 ± 1.4	286 ± 34	24 ± 8
p.F107C	41 ± 2	8.8 ± 0.1	> 500	134 ± 20
p.M155I	18.2 ± 3.7	$8.3 ± 4.0 × 10^{-3}$	N.d.[a]	N.d.[a]
p.H162N	95 ± 3	11.9 ± 1.5	≫ 500	25 ± 2

[a]Full reduction of NQO1$_{holo}$ cannot be achieved in these variants to evaluate its reoxidation (N.d., not det.).

3.2-fold decrease in k_{HT} and a 1.8-fold increase in affinity for FAD, resulting in a 45% decrease in catalytic efficiency (Table 3). Interestingly, the *slow* pathway for FAD reduction was performed efficiently but, as observed for the same pathway in p.W106R, it was NADH-independent thus suggesting that the mutation is changing the rate-limiting step (analogous to the EX2 to EX1 shift in mechanism, see Appendix S1). The cooperative index for this mutant was ~ 10,

similar to that of WT, consistent with an efficient communication between active sites during the reaction (note that the MMI is hardly affected by the mutation p.F107C; Fig. 6B). However, conclusions related to these effects on the negative cooperativity are not clear since the reduction kinetics is strongly dominated by the amplitude and kinetics of the *fast* step. Although these catalytic alterations are not as dramatic as those of p.W106R, our results indicate that in some cases,

Table 3. Enzyme kinetic parameters for the reductive half-reaction of NQO1$_{holo}$ variants with NADH. Primary data and fittings are shown in Fig. 7.

	Fast			Slow		
	k_{HT} (s^{-1})	$K_{d(NADH)}$ (μM)	$k_{HT}/K_{d(NADH)}$ (s$^{-1}\cdot$μM^{-1})	k_{HT} (s^{-1})	$K_{d(NADH)}$	$k_{HT}/K_{d(NADH)}$ (s$^{-1}\cdot$μM^{-1})
WT	281 ± 12	15.2 ± 1.9	18.5 ± 2.4	14.3 ± 1.5	8.2 ± 3.3	1.7 ± 0.7
p.W106R	$12.5 \pm 0.1 \times 10^{-3a}$	N.a.[a]	N.a.[a]	$2.6 \pm 0.4 \times 10^{-3a}$	N.a.[a]	N.a.[a]
p.W106C	397 ± 17	29 ± 3	13.9 ± 1.5	26 ± 1	10.9 ± 1.5	2.4 ± 0.4
p.F107C	87 ± 2	8.3 ± 0.7	10.5 ± 0.9	8.8 ± 0.1[a]	N.a.[a]	N.a.[a]
p.M155I	337 ± 61	64 ± 21	5.2 ± 2.0	0.104 ± 0.028	43 ± 24	$2.4 \pm 1.5 \times 10^{-3}$
p.H162N	415 ± 31	27 ± 5	15.3 ± 6.4	N.a.[b]	N.a.[b]	2.1 ± 0.9[b]

[a]k_{obs} is independent of NADH concentration, therefore data do not allow us to determine $K_{d(NADH)}$ (N.a., not applicable) and the given value corresponds to a limiting k_{HT} calculated as the average of k_{obs} at different [NADH]; [b]k_{obs} is linearly dependent on [NADH]; therefore, data do not allow to determine neither $K_{d(NADH)}$ nor k_{HT} (N.a., not applicable) and the given value corresponds to the second-order rate constant of the process calculated as the slope of the k_{obs} vs. [NADH] plot.

even mild perturbations of the structural stability of the active site may cause noticeable alterations in enzyme catalysis and kinetic mechanism. An alternative hypothesis is that this mutation is affecting dynamics relevant to the reaction mechanism. This is not well probed by HDX-MS (note that dynamics cannot be properly addressed since the EX1 regime is marginally detected by this technique; Fig. S3) and possibly the relevant dynamics are faster (in the μs to ms scale) than those sampled by HDX-MS.

The mutant p.M155I

M155 belongs to the MMI and the DBS and is next to H156 which is part of the FBS and the DBS. M155 is buried in the structure ($12.9 \pm 3.9\%$ of ASA, using the same procedure described for p.W106R). This mutation causes an increase in hydrophobicity [30] without a change in size [31].

Despite the mutation p.M155I does not appear to be highly disruptive, its effects on the local stability of the holo-state are strong, affecting about 50% of residues in NQO1, the MMI, the FBS and the DBS. It does destabilize both the NTD and CTD. Upon Dic binding, these effects are significantly reduced (Figs 4–6; Figs S1, S2, S5E, S6E, S7E, S8E, S9E, S10E and S17). It is interesting to note that this mutation reduced the FAD binding affinity by ~ 50-fold (Fig. 3E), consistent with a strong destabilization of the FBS in the NQO1$_{holo}$ state (Fig. 6C). This is one of the lowest affinities for FAD reported for a single missense variant of NQO1 ([12,19] and this work), reducing the apparent binding free energy by 2.3 kcal·mol^{-1}. The strong destabilization on the MMI, FBS and DBS in the NQO1$_{holo}$ state may also have implications in the catalytic performance of this mutant.

To further characterize the functional consequences of p.M155I, we carried out kinetic analysis for the reductive and oxidative half-reactions at stoichiometric protein to reductive/oxidative ligand (NADH/DCPIP) (Table 2; Figs S11E, S12E and S13E). For the reductive half-reaction, the two-step mechanism was observed but with large effects on k_{obs}, which were reduced by 4-fold and 900-fold in the *fast* and *slow* steps. In addition, the spectral properties of the intermediate species B and C showed markedly higher absorption intensities than those of these species in WT NQO1 (Fig. S12A,E), supporting severe alterations of equilibria in the overall reaction mechanism. Oxidation of FAD by DCPIP could not be analysed due to the inability to achieve the full NQO1$_{red}$ state upon reduction. The NADH-concentration dependence of k_{obs} for the reductive half-reaction (Fig. 7E and Table 3) confirmed substantial alterations in the kinetic properties. For the *fast* step, p.M155I reduced the affinity for NADH by fourfold, resulting in a threefold decrease in catalytic efficiency (Table 3). These effects were dramatic for the *slow* step that showed 140-fold lower k_{HT}, fivefold higher $K_{d(NADH)}$ and resulted in a 700-fold decrease in catalytic efficiency compared with WT NQO1 (Table 3). Accordingly, the cooperative index for p.M155I was about 600-fold higher than that of WT NQO1. This large negative cooperativity can be regarded as an extreme case of half-of-sites reactivity in p.M155I (i.e. only one active site is operative per dimer in a relevant time-scale, up to 1 min). Therefore, the strong effects on FAD and NADH binding, catalytic efficiency and negative cooperativity seem to primarily stem from large destabilization of the active site (FBS and DBS) and the MMI in the NQO1$_{holo}$ state (Fig. 6C–B).
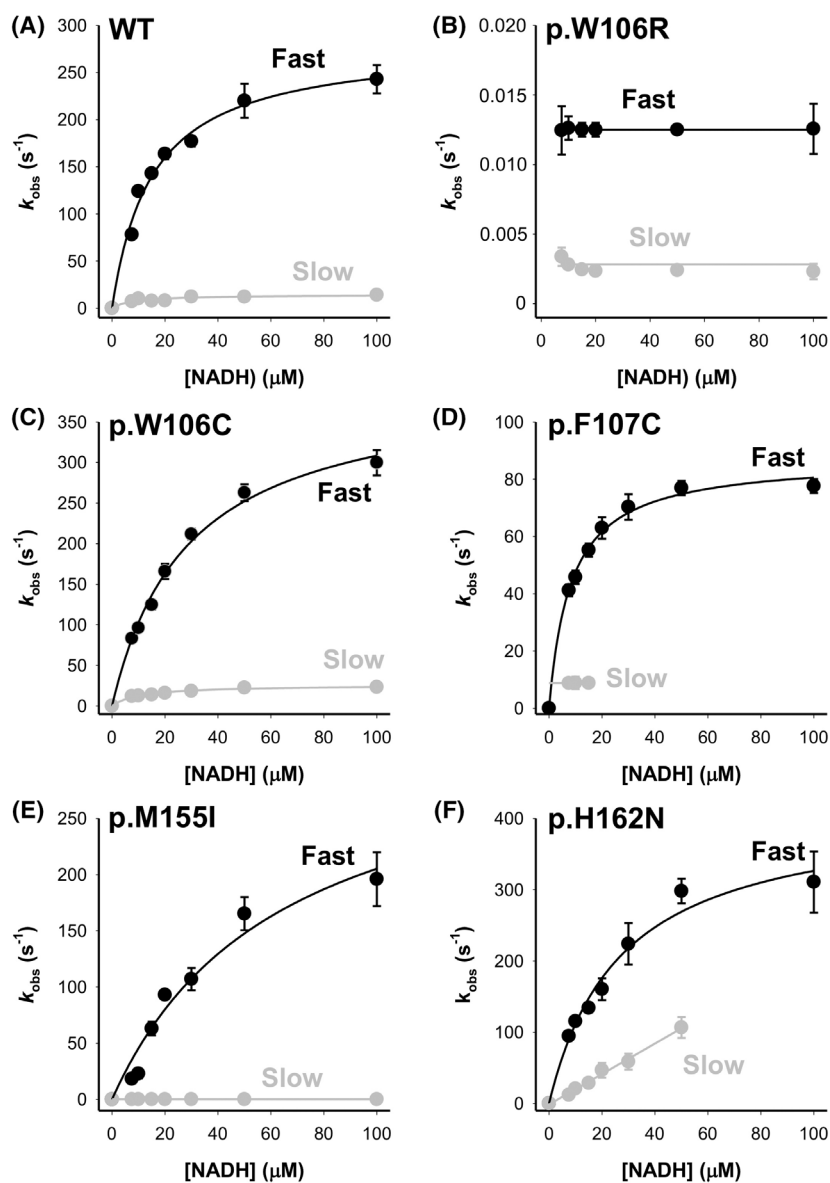
Fig. 7. Pre-steady state enzyme kinetic analyses for the reductive half-reaction of NQO1 WT and active-site mutants. Panels A–F show the dependence of the observed rate constants on NADH concentration for the different NQO1 variants. Data for the *fast* and *slow* steps are shown in black and grey, respectively. Lines are best fits to a hyperbolic function or to a straight line.

The mutant p.H162N

H162 belongs to the MMI, FBS and DBS and it is buried in the structure ($6.6 \pm 0.4\%$ of ASA, using the same procedure described for p.W106R). The mutation p.H162N causes little change in hydrophobicity [30] or size [32]. Therefore, we may initially consider this mutation as quite conservative.

The mutation p.H162N strongly reduces the FAD binding affinity (by 30-fold vs. WT NQO1; Fig. 3F) equivalent to an apparent binding free energy penalty of ~ 2.0 kcal·mol^{-1}. The effects of the mutation p.H162N on protein stability are quite similar to those observed for p.M155I (Figs 4–6; Figs S1, S2,

S5F, S6F, S7F, S8F, S9F, S10F and S18). The NQO1$_{holo}$ state was largely destabilized globally and at the MMI, FBS and DBS (about 50% of the residues are destabilized; Fig. 6), and these effects were much weaker in the NQO1$_{dic}$ state (particularly in the FBS). As proposed for the mutation p.M155I, the low stability of the FBS in the NQO1$_{holo}$ state may cause the low FAD binding affinity, and reduced stability of the MMI, FBS and DBS in the holo-state may affect catalytic performance and functional cooperativity.

Rate constants for the reductive half-reaction (at stoichiometric NQO1$_{holo}$:NADH) showed little or no effects due to p.H162N, while in the oxidative half-

reduction, the *fast* step was accelerated (beyond the temporal resolution of the technique) and the *slow* step was slowed down by sixfold (Table 2). The spectral changes associated with the two steps in the reductive (Fig. S12A,F) and oxidative (Fig. S15A,D) processes supported that the mutation p.H162N does not affect the overall reaction mechanism. When evaluating the NADH dependence of k_{obs}, k_{HT} and $K_{d(NADH)}$ for the *fast* step were slightly affected (both increased by 50–80%) leading to nearly no change in catalytic efficiency (Fig. 7F and Table 3). Interestingly, for the *slow* step, the k_{obs} values despite being larger than at equivalent NADH concentrations in the WT case, linearly increased on substrate concentration with no evidence of saturation (i.e. similar to an EX2 scenario; Appendix S1). This may result from different effects of this mutation in the *slow* pathway for FAD reduction: a large increase in the $K_{d(NADH)}$ and/or k_{off}, or a large decrease in the k_{on}.

Genotype–phenotype correlations using multifeatured experimental characterization of active-site mutants and bioinformatic predictions

The results obtained from our detailed experimental characterization of the NQO1 active-site mutants are summarized in Table 4. There, mutational effects on 13 features regarding protein functionality and stability are semiquantitatively ranked. Overall, we may conclude that the mutations p.W106R, p.M155I and p.H162N are largely detrimental to NQO1 function and stability, whereas the effects of p.W106C and p.F107C are milder. Interestingly, mutational effects predicted by different widely used bioinformatic tools did not agree well with the experimental outcome, predicting (within some degree of disagreement between tools) than all mutations except p.H162N should be largely detrimental. Therefore, experimental and computational prediction of mutational seemed to correlate poorly.

To analyse more quantitatively this apparently poor correlation, we determined scores for the experimental results (13 features compiled in Table 4) and those provided by computational predictions (using six different tools, Table 1). All 19 scores were normalized in such a way that a score of 0 indicated WT-like behaviour and a score of 1 indicated a largely deleterious effect. Thus, we evaluated average scores for experimental analyses (NS_{exp}) and predictions (NS_{pred}) (Tables 1 and 4). These scores were aimed to capture either the average effects (with the same weight for each of the features; note that expression of all variants was quite successful and thus, none seemed to largely affect NQO1 foldability) observed experimentally on NQO1 functional and stability and the average performance of the predictions. When we simply plotted the two scores (Fig. 8), we observed a very weak correlation between the two scores, as expected. This simple exercise also confirmed that the functional effects of the mutations p.W106C and p.F107C (two largely non-conservative mutations) are much milder than those predicted by *in silico* tools, while the opposite behaviour was found for p.H162N (in principle a more conservative mutation). Overall, these results are

Table 4. Summary of the effects of active site mutations. For each protein variant and feature, these were clustered for semiquantitative comparison in four categories: (++++) mildly improved vs. WT; (+++) WT-like; (++) mildly moderately impaired vs. WT; (+) largely impaired vs. WT. NS_{exp} is the average of experimental normalized scores, considering ++++ as −0.5, +++, as 0, ++ as 0.5 and +as 1.

	NQO1 variant					
	WT	p.W106R	p.W106C	p.F107C	p.M155I	p.H162N
Protein conformation	+++	+++	+++	+++	++	++
FAD bound	+++	++	++	+++	++	++
FAD affinity	+++	+	+++	++++	+	+
Global stability by HDX-Holo	+++	+++	++	+++	+	+
Global stability by HDX-Dic	+++	+	++	+++	+++	++
FBS stability by HDX	+++	+	++	++	+	+
DBS stability by HDX	+++	+	++	+++	+	+
MMI stability by HDX	+++	+	++	+++	+	+
Fast FAD reduction	+++	+	+++	++	+++	+++
Slow FAD reduction	+++	+	+++	++	++	++
FAD oxidation	+++	+	++	+++	+	+++
Reaction mechanism	+++	++	+++	++	+	+++
Cooperativity	+++	+	+++	++	+++	+++
NS_{exp}	0	0.77	0.27	0.15	0.65	0.54
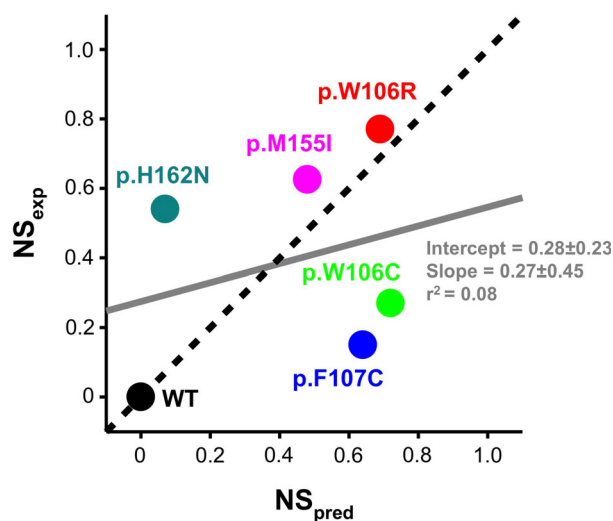
Fig. 8. Lack of correlation between mutational effects from experiments (NS_{exp}) and predictive tools (NS_{pred}) for the active-site mutants. Predicted scores (NS_{pred}) are the average from normalized scores using six different predictive tools (Table 1). Experimental scores (NS_{exp}) are the average of normalized scores derived from the characterization of 13 functional and stability features (Table 4). The dashed black line represents a perfect linear correlation between scores, whereas the solid grey line shows the actual (poor) linear correlation.

somewhat counterintuitive, since active site residues evolve slowly and are strongly constrained in terms of physico-chemical properties of the residue such as polarity, charge and hydrophobicity [24,33].

Conclusions

State-of-the-art DNA sequencing technologies allow to compare the genomes of hundreds of thousands of individuals and have revealed an astonishingly large genetic diversity in the human genome. In this context, classical tools to establish genotype–phenotype correlations and predict pathogenicity of missense variants *in silico* are still not sufficiently accurate, particularly when individual mutations are analysed (actually, the output of many of these tools is binary, for example, pathogenic or damaging vs. benign or neutral). In this work, we carried out comprehensive structure–function studies of five naturally occurring missense mutations affecting residues belonging to the active site of the cancer-associated NQO1 enzyme and predicted to be pathogenic in most of the cases (Table 1). Experimental characterization of mutational effects shows widely different and counterintuitive consequences (Table 4). For instance, two non-conservative mutations affecting the same residue (p.W106R and p.W106C) have

largely different consequences (the former is catastrophic while the latter is much milder). In addition, quite conservative mutations (p.M155I and H162N) display much larger effects than some non-conservative changes (p.W106C). Remarkably, mutations with large functional consequences are shown to target the stability of different functional states (p.W106R, the $NQO1_{dic}$ state vs. p.M155I and p.H162N, the $NQO1_{holo}$ state). Although in most cases, structure–function correlations can be nicely drawn, some mutations appear to affect enzyme function with little structural impact (p.F107C). Our work highlights the necessity to incorporate additional information to predictive tools, such as structural effects on different ligation states (i.e. the flavin, substrate) and several protein functional features as well as some notion of structural plasticity (as the ability of the protein conformational landscape to adapt to different types of mutations in different or similar structural locations).

Materials and methods

Protein expression and purification

Mutations were introduced by site-directed mutagenesis in the wild-type (WT) NQO1 cDNA cloned into the pET-15b vector (pET-15b-NQO1) by GenScript (Leiden, the Netherlands). Codons were optimized for expression in *Escherichia coli* and mutagenesis was confirmed by sequencing the entire cDNA. The plasmids were transformed in *E. coli* BL21(DE3) cells (Agilent Technologies, Santa Clara, CA, USA) for protein expression. These constructs contain a hexa-his N-terminal tag for purification.

For protein purifications, a preculture (40 mL) was prepared from a single clone for each variant and grown for 16 h at 37 °C in LBA (Luria-Bertani medium with 0.1 mg·mL^{-1} ampicillin) and diluted into 2.4–4.8 L of LBA. After 3 h with shaking (200 r.p.m.) at 37 °C, NQO1 expression was induced by the addition of 0.5 mM IPTG for 6 h at 25 °C. Cells were harvested by centrifugation at 8000 *g* and frozen overnight at −80 °C. NQO1 proteins were purified using immobilized nickel affinity chromatography (IMAC) columns (Cytiva, Marlborough, MA, USA) as described [13]. Isolated dimeric fractions of NQO1 variants were exchanged to HEPES-KOH buffer 50 mM pH 7.4 using PD-10 columns (Cytiva). The UV–visible spectra of purified NQO1 proteins were registered in a Cary 50 spectrophotometer (Agilent Technologies, Waldbronn, Germany) and used to quantify the protein content as described in [11]. For the samples used in pre-steady state kinetic analyses, NQO1 proteins were incubated with 1 mM FAD and the excess of FAD was removed using PD-10 columns, obtaining a saturation fraction (FAD:NQO1

monomer) higher than 90% based on UV–visible spectra. Apo-proteins were obtained by treatment with 2 M urea and 2 M KBr as described [20], obtaining samples with < 2% saturation fraction of FAD based on UV–visible spectra. Samples were stored at −80 °C upon flash freezing in liquid N$_2$. Protein purity and integrity were checked by polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate (SDS/PAGE). Each NQO1 variant was expressed at least three times.

Protein conformation

After centrifugation for 20 min at 21 000 g at 4 °C, the FAD bound to NQO1 proteins as purified was determined from the UV–visible spectrum using $\varepsilon_{(280)} = 47\,900$ M^{-1}·cm^{-1} for NQO1, $\varepsilon_{(450)} = 22\,000$ M^{-1}·cm^{-1} and $\varepsilon_{(280)} = 11\,300$ M^{-1}·cm^{-1} for FAD [11]. Spectra were collected in a Cary 50 spectrophotometer using 0.3 cm path-length quartz cuvettes and ~ 30 μM of NQO1 in the protein subunit. Data were reported as mean ± SD from three different purifications.

Far-UV CD spectroscopy was performed at 25 °C in K-phosphate 20 mM at pH 7.4 using 5 μM protein (in NQO1 monomer) with 25 μM FAD. Spectra were collected at 25 °C in a Jasco J-710 spectropolarimeter (Tokyo, Japan) in the 195–260 nm range, at 100 nm·min^{-1}, using 1 nm bandwidth, 1 s response time and 1 mm path length cuvette. Each spectrum was the average of six scans, and each sample was prepared in triplicate. Mean residue ellipticities ($[\Theta]_{MRW}$) were calculated using Eqn (1):

$$[\Theta]_{MRW} = \frac{MRW \cdot \Theta_{obs}}{10 \cdot l \cdot c}, \qquad (1)$$

where MRW was equal to the molecular weight of the NQO1 monomer (31 691 g·mol^{-1}) divided by $(N-1)$, being $N = 280$ the number of residues in the monomer, Θ_{obs} was the ellipticity (in mdeg), l was the path length (in cm) and c the concentration of protein in mg·mL^{-1}. Spectra were reported as mean ± SD from three replicates.

Near-UV CD spectroscopy was performed at 25 °C in K-phosphate 20 mM at pH 7.4 using 20 μM protein as purified (in NQO1 monomer) with 50 μM FAD as described above for far-UV CD measurements. Measurements were carried out in the 250–600 nm range, at 100 nm·min^{-1}, using 1 nm bandwidth, 1 s response time and a 5-mm path length cuvette. Each spectrum was the average of 10 scans, and the appropriate blank in the absence of protein was acquired and subtracted.

Fluorescence spectra were acquired in a Varian Cary Eclipse spectrofluorometer (Agilent Technologies) using 1 cm path length cuvettes and 1 μM (in monomer) of protein in the presence of 5 μM FAD in 20 mM K-phosphate pH 7.4. The excitation wavelength was 280 nm and the emission fluorescence was collected between 300 and 450 nm. Excitation and emission slits were 5 nm. All spectra were acquired at 25 °C at a 120 nm·min^{-1} scan rate and 10 scans were registered and averaged. Blanks without protein were routinely measured and subtracted. The spectral center of mass (SCM) was determined using Eqn (2):

$$SCM = \frac{\sum\limits_{330\,nm}^{380\,nm} I_\lambda \cdot \lambda}{\sum\limits_{330\,nm}^{380\,nm} I_\lambda}. \qquad (2)$$

where I_λ is the emission intensity at a given emission wavelength λ. Data were reported as mean ± SD from three replicates.

Dynamic light scattering was carried out in a Zetasizer μV instrument (Malvern Panalytical, Malvern, UK) using 1.5 mm path length cuvettes and 5 μM (in monomer) of protein with 25 μM FAD in 20 mM K-phosphate pH 7.4 at 25 °C. Thirty measurements with an acquisition time of 10 s were acquired for each DLS analysis, averaged and used to determine the hydrodynamic radius assuming spherical scattering particles (using the Stokes-Einstein approach). Data were reported as mean ± SD from three replicates. DLS data were processed and analysed using the ZETASIZER software (Malvern Panalytical).

Flavin-adenine dinucleotide binding affinity

Fluorescence titrations were carried out at 25 °C using 1 × 0.3 cm path-length cuvettes in a Varian Cary Eclipse spectrofluorometer (Agilent Technologies). Experiments were carried out in 20 mM K-phosphate, pH 7.4, essentially as described in [15]. Briefly, 20 μL of a 12.5 μM NQO1$_{apo}$ stock solution (in subunit) was mixed with 0–500 μL of FAD 10 μM, and the corresponding volume of buffer was added to yield a 1 mL of final volume. Samples were incubated at 25 °C in the dark for at least 10 min before measurements. Spectra were acquired in the 340–360 nm range upon excitation at 280 nm (slits 5 nm), and spectra were averaged over 10 scans registered at a scan rate of 200 nm·min^{-1}.

For W106R, which exhibited a very low binding affinity for FAD, the titration was carried out using near-UV CD spectroscopy. Spectra were collected in a Jasco J-710 spectropolarimeter at 25 °C in 50 mM K-HEPES pH 7.4 using 9.4 μM (in monomer) of apoprotein in the absence or presence of FAD (0–130 μM). Spectra were collected in the 300–600 nm range at 200 nm·min^{-1}, using 2 nm bandwidth, 2 s time response and 5 mm path-length cuvettes. Each spectrum was the average of 8 scans, and each sample was appropriately corrected for blanks containing the buffer and the corresponding FAD concentration.

Flavin-adenine dinucleotide binding titrations (following fluorescence intensities at 350 nm or CD ellipticity at

375 nm) were fitted using a single and identical type of binding sites as described in [15] using the GRAPHPAD PRISM 7 sofware (DotMatics, Boston, MA, USA).

Hydrogen/deuterium exchange mass spectrometry

Amide hydrogen/deuterium exchange (HDX) of NQO1 was studied for the WT and mutant variants in the NQO1$_{holo}$ and NQO1$_{dic}$ states as described previously [20] with some modifications. Briefly, to start the exchange reaction 20 μM protein solution was diluted 10 times with a D$_2$O-based 20 mM HEPES-NaOH, 100 mM NaCl, 0.5 mM TCEP [tris (2-carboxyethyl)phosphine], pD 7.4. The exchange was terminated after 10, 50, 250, 1250 and 6250 s by mixing (1 : 1) with 0.5 M Glycine-HCl, pH 2.3 and the samples were flash frozen in liquid nitrogen. Time points 10, 250 and 6250 s were replicated. Custom-made co-immobilized nepenthesin-2/pepsin column was used for online proteolysis which was driven by 0.4% formic acid (FA) in water pumped at 200 μL·min^{-1}. The solvent was delivered by 1260 Infinity II Quaternary pump (Agilent Technologies). Peptides were trapped and desalted on a SecurityGuard™ pre-column (ULTRA Cartridges UHPLC Fully Porous Polar C18, 2.1 mm; Phenomenex, Torrance, CA, USA) using the same solvent. The duration of digestion and desalting was 3 min. Following desalting, the peptides were separated on an analytical column (LUNA® Omega Polar C18 Column, 100 Å, 1.6 μm, 100 mm × 1.0 mm; Phenomenex) using 1290 Infinity II LC System (Agilent Technologies). Linear gradient 5–45% B in 6 min was followed by a quick step to 99% B lasting 5 min. Flow during the separation was 40 μL·min^{-1}. Solvent A was 0.1% FA/2% acetonitrile (ACN) in water, B was 0.1% FA/98% ACN in water. Digestion, desalting, and separation steps were performed at 0 °C and pH 2.3 to minimize deuterium back-exchange. LC system was directly coupled to an ESI source of 15T FT-ICR mass spectrometer (solariX XR; Bruker Daltonics, Bremen, Germany) operating in a broad-band MS mode. Correction of back-exchange was done as described in [34] using fully deuterated control prepared for each NQO1 variant [20]. The data were further processed in DATAANALYSIS 5.3, exported and assembled into a project under in-house developed program DEUTEX [35]. Peptides were identified using separate data-dependent LC–MS/MS runs carried out using an identical LC setup connected to ESI-timsTOF Pro with PASEF. MASCOT (v 2.4; Matrix Science, London, UK) was used for data searching against a custom-built database containing sequences of the proteases, NQO1 variants and common contaminants. Decoy search was enabled with a false-discovery ratio < 1% and an ion score cut-off of 20. All data were deposited to ProteomeXchange Consortium via the PRIDE database [PXD036417] [36].

To evaluate the effect of mutations, the difference in kinetics of deuterium incorporation (% D vs. time, Figs S1 and S2) of mutants and the WT protein was calculated for a given ligation state and each protein segment experimentally determined. Analysis of exchange behaviour showed that the EX2 mechanism dominates HDX in most of the peptides, variants and ligation states (Fig. S3), thus supporting that changes in HDX are associated with those in the local thermodynamic stability of the segments [20]. The average of the two most different time points (mutant-WT) was used to determine the $\Delta\%D_{av}$ values. This parameter allows to readily compare the HDX kinetics between two given NQO1 states/variants in different protein segments [12,19,20].

Enzyme kinetics for the reductive and oxidative half-reactions

For enzyme kinetic analyses of the reductive and oxidative half-reactions, we followed the procedures described for the WT protein under anaerobic conditions using a stopped-flow spectrophotometer as described [13]. Briefly, the reductive half-reaction was measured by mixing the NQO1$_{holo}$ protein with a solution of NADH, yielding final concentrations of 7.5 and 7.5–100 μM, respectively. The oxidative half-reaction was monitored after mixing NQO1$_{red}$ samples (NQO1$_{red}$ was obtained by previous mixing of NADH to the holo-NQO1, both at 7.5 μM) with an equimolar concentration of 2,6-Dichlorophenolindophenol (DCPIP). Reactions were performed in 20 mM HEPES-KOH, pH 7.4.

Multiple wavelength absorption data in the flavin absorption region were collected and processed as described [13]. Time-dependent spectral deconvolution was performed by global fitting analysis and numerical integration using previously described procedures [13] and allowed to determine observed rate constants (k_{obs}) for these steps as well as spectroscopic properties of these species (A, B and C). Despite practical limitations prevented these measurements to reach pseudo-first-order conditions [19], hyperbolic dependences of k_{obs} vs. NADH concentrations were fitted using Eqn (3):

$$k_{obs} = \frac{k_{HT} \cdot [\text{NADH}]}{K_{d(\text{NADH})} + [\text{NADH}]}, \qquad (3)$$

where k_{HT} is the limiting rate constant for HT and $K_{d(\text{NADH})}$ is the apparent equilibrium dissociation constant for NADH to a given active site. Fittings were carried out using SIGMAPLOT v.9.0 (SYSTAT Software Inc., Chicago, IL, USA).

Bioinformatic analysis

Six different readily available algorithms for the prediction of mutational effects were used. These are based on various

features such as evolutionary conservation, chemical nature of the amino acid change and structural consequences. These tools are briefly described in this section.

We used the Ensembl Variant Effect Predictor (VEP) (https://www.ensembl.org/Tools/VEP) [37] to obtain predictions for single nucleotide variants. Most of these predictions can be assessed at the database dbNSFP [38].

POLYPHEN-2 (http://genetics.bwh.harvard.edu/pph2/) predicts the impact of amino acid substitutions on the structure and function of a human protein using physical and comparative considerations [39]. Its score (0–1) yields the probability of the variation being damaging and contemplates three classes: benign, possibly damaging and probably damaging.

SIFT (https://sift.bii.a-star.edu.sg/) is based on sequence homology and the physical properties of amino acids [40]. It aligns protein sequences in numerous species and calculates normalized probabilities for all possible substitutions from the alignment. Its score also ranges 0–1. The amino acid substitution is predicted as damaging if the score is ≤ 0.05, and as tolerated if the score is > 0.05.

CADD (https://cadd.gs.washington.edu/) is a metapredictor that takes into account many diverse annotations into a single score (C score) [41]. The higher the raw C score (between 1 and 99), the more likely is the change to be deleterious.

MUTATION ASSESOR (http://mutationassessor.org) estimates the functional impact of a missense variant based on evolutionary conservation of the affected amino acid in protein homologues [42]. A conservation score is combined with a specificity score to determine a functional impact score (0–1). Variants classed as neutral or low are predicted to have low or no impact on protein function, whereas variants classed as medium or high are predicted to result in altered function.

REVEL (integrated in VEP) [43] is a meta-predictor based on a number of individual tools: MUTPRED, FATHMM, VEST, POLYPHEN, SIFT, PROVEAN, MUTATIONASSESSOR, MUTATIONTASTER, LRT, GERP, SIPHY, PHYLOP, and PHASTCONS. The score (between 0 and 1) classifies the variations between likely benign (score < 0.5) and likely disease-causing (score ≥ 0.5).

METAL R (integrated into VEP) ([38]) uses logistic regression to integrate nine independent scores and allele frequency information to yield a score (between 0 and 1), with lower scores considered as tolerated and higher scores more likely to be damaging.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author contributions

JLP-G performed experiments and analysed data. EA-C performed experiments and analysed data. DSL performed experiments and analysed data. DK performed experiments and analysed data. ES analysed data and contributed reagents. PM planned experiments and analysed data. MM planned experiments and analysed data. ALP planned experiments, analysed data and wrote the paper.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.16677.

Data availability statement

All HDXMS data were deposited to ProteomeXchange Consortium via the PRIDE database [PXD036417]. No new high-resolution structures were generated in this manuscript. All data in a processed form are contained in the manuscript. Any other raw data can be retrieved upon reasonable request to the corresponding author (ALP).

References

1 Arnedo-Pac C, Lopez-Bigas N, Muiños F. Predicting disease variants using biodiversity and machine learning. *Nat Biotechnol.* 2022;**40**:27–8.

2 Høie MH, Cagiada M, Beck Frederiksen AH, Stein A, Lindorff-Larsen K. Predicting and interpreting

large-scale mutagenesis data using analyses of protein stability and conservation. *Cell Rep*. 2022;**38**:110207.

3 Shendure J, Akey JM. The origins, determinants, and consequences of human mutations. *Science*. 2015;**349**:1478–83.

4 McInnes G, Sharo AG, Koleske ML, Brown JEH, Norstad M, Adhikari AN, et al. Opportunities and challenges for the computational interpretation of rare variation in clinically important genes. *Am J Hum Genet*. 2021;**108**:535–48.

5 Beaver SK, Mesa-Torres N, Pey AL, Timson DJ. NQO1: a target for the treatment of cancer and neurological diseases, and a model to understand loss of function disease mechanisms. *Biochim Biophys Acta Proteins Proteom*. 2019;**1867**:663–76.

6 Salido E, Timson DJ, Betancor-Fernández I, Palomino-Morales R, Anoz-Carbonell E, Pacheco-García JL, et al. Targeting HIF-1α function in cancer through the chaperone action of NQO1: implications of genetic diversity of NQO1. *J Pers Med*. 2022;**12**:747.

7 Ross D, Siegel D. The diverse functionality of NQO1 and its roles in redox control. *Redox Biol*. 2021;**41**:101950.

8 Faig M, Bianchet MA, Talalay P, Chen S, Winski S, Ross D, et al. Structures of recombinant human and mouse NAD(P)H:quinone oxidoreductases: species comparison and structural changes with substrate binding and release. *Proc Natl Acad Sci USA*. 2000;**97**:3177–82.

9 Li R, Bianchet MA, Talalay P, Amzel LM. The three-dimensional structure of NAD(P)H:quinone reductase, a flavoprotein involved in cancer chemoprotection and chemotherapy: mechanism of the two-electron reduction. *Proc Natl Acad Sci USA*. 1995;**92**:8846–50.

10 Lienhart W-D, Gudipati V, Uhl MK, Binter A, Pulido SA, Saf R, et al. Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H: quinone oxidoreductase(1.). *FEBS J*. 2014;**281**:4691–704.

11 Medina-Carmona E, Fuchs JE, Gavira JA, Mesa-Torres N, Neira JL, Salido E, et al. Enhanced vulnerability of human proteins towards disease-associated inactivation through divergent evolution. *Hum Mol Genet*. 2017;**26**:3531–44.

12 Pacheco-Garcia JL, Anoz-Carbonell E, Vankova P, Kannan A, Palomino-Morales R, Mesa-Torres N, et al. Structural basis of the pleiotropic and specific phenotypic consequences of missense mutations in the multifunctional NAD(P)H:quinone oxidoreductase 1 and their pharmacological rescue. *Redox Biol*. 2021;**46**:102112.

13 Anoz-Carbonell E, Timson DJ, Pey AL, Medina M. The catalytic cycle of the antioxidant and cancer-associated human NQO1 enzyme: hydride transfer, conformational dynamics and functional cooperativity. *Antioxidants*. 2020;**9**:1–22.

14 Asher G, Dym O, Tsvetkov P, Adler J, Shaul Y. The crystal structure of NAD(P)H quinone oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry*. 2006;**45**:6372–8.

15 Pacheco-García JL, Cano-Muñoz M, Sánchez-Ramos I, Salido E, Pey AL. Naturally-occurring rare mutations cause mild to catastrophic effects in the multifunctional and cancer-associated NQO1 protein. *J Pers Med*. 2020;**10**:1–31.

16 Pey AL. Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int J Biol Macromol*. 2018;**118**:1912–23.

17 Medina-Carmona E, Palomino-Morales RJ, Fuchs JE, Padín-Gonzalez E, Mesa-Torres N, Salido E, et al. Conformational dynamics is key to understanding loss-of-function of NQO1 cancer-associated polymorphisms and its correction by pharmacological ligands. *Sci Rep*. 2016;**6**:20331.

18 Medina-Carmona E, Betancor-Fernández I, Santos J, Mesa-Torres N, Grottelli S, Batlle C, et al. Insight into the specificity and severity of pathogenic mechanisms associated with missense mutations through experimental and structural perturbation analyses. *Hum Mol Genet*. 2019;**28**:1–15.

19 Pacheco-Garcia JL, Loginov DS, Anoz-Carbonell E, Vankova P, Palomino-Morales R, Salido E, et al. Allosteric communication in the multifunctional and redox NQO1 protein studied by cavity-making mutations. *Antioxidants*. 2022;**11**:1110.

20 Vankova P, Salido E, Timson DJ, Man P, Pey AL. A dynamic Core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. *Biomolecules*. 2019;**9**:728.

21 Medina-Carmona E, Rizzuti B, Martín-Escolano R, Pacheco-García JL, Mesa-Torres N, Neira JL, et al. Phosphorylation compromises FAD binding and intracellular stability of wild-type and cancer-associated NQO1: insights into flavo-proteome stability. *Int J Biol Macromol*. 2019;**125**:1275–88.

22 Muñoz IG, Morel B, Medina-Carmona E, Pey AL. A mechanism for cancer-associated inactivation of NQO1 due to P187S and its reactivation by the consensus mutation H80R. *FEBS Lett*. 2017;**591**:2826–35.

23 Pey AL, Megarity CF, Timson DJ. NAD(P)H quinone oxidoreductase (NQO1): an enzyme which needs just enough mobility, in just the right places. *Biosci Rep*. 2019;**39**:BSR20180459.

24 Mayorov A, Dal Peraro M, Abriata LA. Active site-induced evolutionary constraints follow fold polarity principles in soluble globular enzymes. *Mol Biol Evol*. 2019;**36**:1728–33.

25 Medina-Carmona E, Neira JL, Salido E, Fuchs JE, Palomino-Morales R, Timson DJ, et al. Site-to-site

interdomain communication may mediate different loss-of-function mechanisms in a cancer-associated NQO1 polymorphism. *Sci Rep*. 2017;**7**:44532.

26 Fraczkiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem*. 1998;**19**:319–33.

27 Pey AL, Megarity CF, Timson DJ. FAD binding overcomes defects in activity and stability displayed by cancer-associated variants of human NQO1. *Biochim Biophys Acta*. 2014;**1842**:2163–73.

28 Frago S, Goñi G, Herguedas B, Peregrina JR, Serrano A, Perez-Dorado I, et al. Tuning of the FMN binding and oxido-reduction properties by neighboring side chains in *Anabaena* flavodoxin. *Arch Biochem Biophys*. 2007;**467**:206–17.

29 Nogués I, Campos LA, Sancho J, Gómez-Moreno C, Mayhew SG, Medina M. Role of neighboring FMN side chains in the modulation of flavin reduction potentials and in the energetics of the FMN:apoprotein interaction in *Anabaena* flavodoxin. *Biochemistry*. 2004;**43**:15111–21.

30 Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;**157**:105–32.

31 Counterman AE, Clemmer DE. Volumes of individual amino acid residues in gas-phase peptide ions. *J Am Chem Soc*. 1999;**121**:4031–9.

32 Perkins SJ. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur J Biochem*. 1986;**157**:169–80.

33 Echave J, Spielman SJ, Wilke CO. Causes of evolutionary rate variation among protein sites. *Nat Rev Genet*. 2016;**17**:109–21.

34 Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Cambridge: Cambridge University Press; 1993.

35 Trcka F, Durech M, Man P, Hernychova L, Muller P, Vojtesek B. The assembly and intermolecular properties of the Hsp70-Tomm34-Hsp90 molecular chaperone complex. *J Biol Chem*. 2014;**289**:9887–901.

36 Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;**50**:D543–52.

37 McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics*. 2010;**26**:2069–70.

38 Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutat*. 2016;**37**:235–41.

39 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;**7**:248–9.

40 Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc*. 2016;**11**:1–9.

41 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;**46**:310–5.

42 Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res*. 2011;**39**:e118.

43 Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an Ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;**99**: 877–85.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Plot SM1. Scenarios for model 3.
Fig. S1. Rainbow heatmaps showing the deuteration profile for WT and active site mutants in their $NQO1_{holo}$ (A) and $NQO1_{dic}$ (B) states.
Fig. S2. Differential rainbow heatmaps for the deuteration profiles of WT and active site mutants in their $NQO1_{holo}$ and $NQO1_{dic}$ states.
Fig. S3. Analysis of EX1 exchange mechanism through peak width analysis.
Fig. S4. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for the p.W106R and p.W106C mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).
Fig. S5. Effect of active-site mutations on the stability of the MMI in the $NQO1_{holo}$ state determined by HDX-MS.
Fig. S6. Effect of active-site mutations on the stability of the MMI in the $NQO1_{dic}$ state determined by HDX-MS.
Fig. S7. Effect of active-site mutations on the stability of the FBS in the $NQO1_{holo}$ state determined by HDX-MS.
Fig. S8. Effect of active-site mutations on the stability of the FBS in the $NQO1_{dic}$ state determined by HDX-MS.

Fig. S9. Effect of active-site mutations on the stability of the DBS in the NQO1$_{holo}$ state determined by HDX-MS.

Fig. S10. Effect of active-site mutations on the stability of the DBS in the NQO1$_{dic}$ state determined by HDX-MS.

Fig. S11. Time-dependent spectra of the NQO1 flavin reduction by NADH.

Fig. S12. Deconvolution of spectral species (A→B→C) observed during flavin reduction with NADH.

Fig. S13. Kinetics of NQO1 flavin reduction by NADH.

Fig. S14. Time-dependent spectra of the NQO1 flavin oxidation by DCPIP.

Fig. S15. Spectral deconvolution of intermediate species (A→B→C) observed during NQO1 flavin oxidation by DCPIP.

Fig. S16. Kinetics of NQO1 oxidation by DCPIP.

Fig. S17. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.F107C and p.M155I mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).

Fig. S18. Data for HDX regarding selected segments in which kinetics of deuterium incorporation for p.W106R and p.W106C mutants differ from that of WT NQO1 ($|\Delta\%D_{av}| > 10\%$).