SUPPORTING INFORMATION

Mechanistic insights on the reduction of glutathione disulfide by protein disulfide isomerase

Rui P. P. Neves, Pedro A. Fernandes, Maria J. Ramos*

UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências,

Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

* mjramos@fc.up.pt

Section A – Molecular modelling of the human PDI model

A peptide of 18 residues from the initial sequence was lacking, as well as the *c*-domain of hPDI. To model the initial sequence we modelled the first two residues (Met1–Leu2) with the LEaP module of the AmberTools13 package,(1) and the remaining (Arg3–Asp18) were transferred from the reduced form of human PDI (hPDI$_{red}$) – with PDB code: 4EKZ.(2) Regarding the *c*-domain, the Ala476···Val504 gap was modelled from the sequence Val488–Glu504 of yeast PDI (yPDI) – with PDB code:2B5E(3) – which includes the α-helix that can be found in this domain;(3) the remaining four residues (Lys505–Leu508) were also modelled with the LeaP module. We emphasize that the segments modelled with the LEaP module exhibit no apparent secondary structure; we expect that no significant modelling error should derive from this protocol. Regarding the use of the x-ray structure from yPDI, we note that despite the low sequence identity between PDI from different organisms, the activity and folding of hPDI and yPDI have been shown to be similar.(4) Finally, we have also modelled two gaps at the *b'*-domain: the Phe249···Thr255 gap was modelled from the same domain in the reduced form of human PDI (hPDI$_{red}$) – with PDB code: 4EKZ;(2) and the Thr319···Met324 gap was modelled from the structural equivalent Gln329···Ala334 in the yPDI.(3)

Table S1. Structural and function data for hPDI

Domain	Sequence	Length	Main function
initial sequence	Met1–Glu22	22 residues	*signalling sequence cleaved when PDI enters the ER*
domain *a*	Glu23–Gly134	112 residues	*catalytic Cys53–Gly54–His55–Cys56 motif able to catalyze thiol-disulfide exchange*
interregion *a/b*	Pro135	1 residue	–
domain *b*	Ala136–Val237	102 residues	*contributes to substrate binding*
domain *b'*	Ile238–Gly349	112 residues	*largest hydrophobic cleft to bind misfolded proteins*
linker *x*	Lys350–Trp364	15 residues	*responds to changes in the Cys397/Cys400 pair by controlling the hydrophobic area available to the substrate*

domain a'	Asp365–Asp476	112 residues	*catalytic Cys397–Gly398–His399–Cys400 motif able to catalyze thiol-disulfide exchange*
domain c	Gly477–Leu508	32 residues	*acidic tail, which may be involved in substrate binding*

The 50 ns MD simulation showed considerable changes to our initial hPDI$_{ox}$ model. When only the four thioredoxin-like domains of the model were considered, the RMSD of most conformations differed in less than 4 Å from the X-ray structure. However, and despite the differences between the X-ray structure and the results from cMD simulations, there are three points that supported the adequacy of the modelling: (1) our starting X-ray structure provided a dimerized form of hPDI; (2) hPDI presented a *U*-shaped form that exhibited a large solvent accessible area, particularly in the oxidized state of the *a'*-domain that we have studied; and (3) there is a linker sequence of 19 amino acids that allowed for rotation of the *a'*-domain, and that exhibited no secondary structure. These three points emphasized that the region that comprehended the linker- and *a'*-domains of hPDI should be very flexible in order to accommodate protein substrates, and that, in the absence of the second hPDI monomer it could interact pronouncedly with the solvent, thus inducing significant domain shifts. The comparison of the rmsd for each frame of the cMD simulation against every other frame, for all the four thioredoxin-like domains (*a*-, *b*-, *b'*-,*a'*-), and the *a*- and *b*-domains of hPDI (the latter were employed to build the QM/MM model) further emphasizes these points. While the four thioredoxin-like domains exhibited rmsd values that were very sparse and mostly higher than 4 Å, the *a*- and *b*-domains exhibited rmsd values mostly lower than 2 Å (Figure S1). Hence, the *b'xa'*-region should be in fact very flexible.

A secondary structure analysis highlighted the stability of the folding of the *ab*-region of hPDI, with the thioredoxin-like fold (β–α–β–α–β–α–β–β–α)(5, 6) conserved for both domains during the entire cMD simulation (Figure S2). Hence, the *a*- and *b*-domains in the x-ray structure are appropriate to describe a catalytic conformation of hPDI towards GSSG.
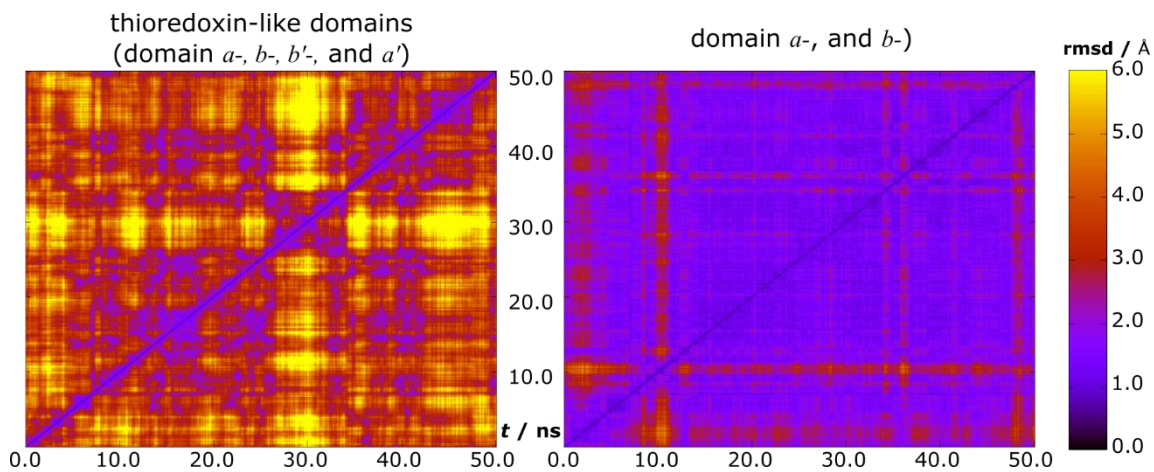
Figure S1. 2-dimensional rmsd for the four thioredoxin-folds (on the left), and the a- and b-domains (on the right) of hPDI throughout the 50 ns NPT cMD performed during this study.
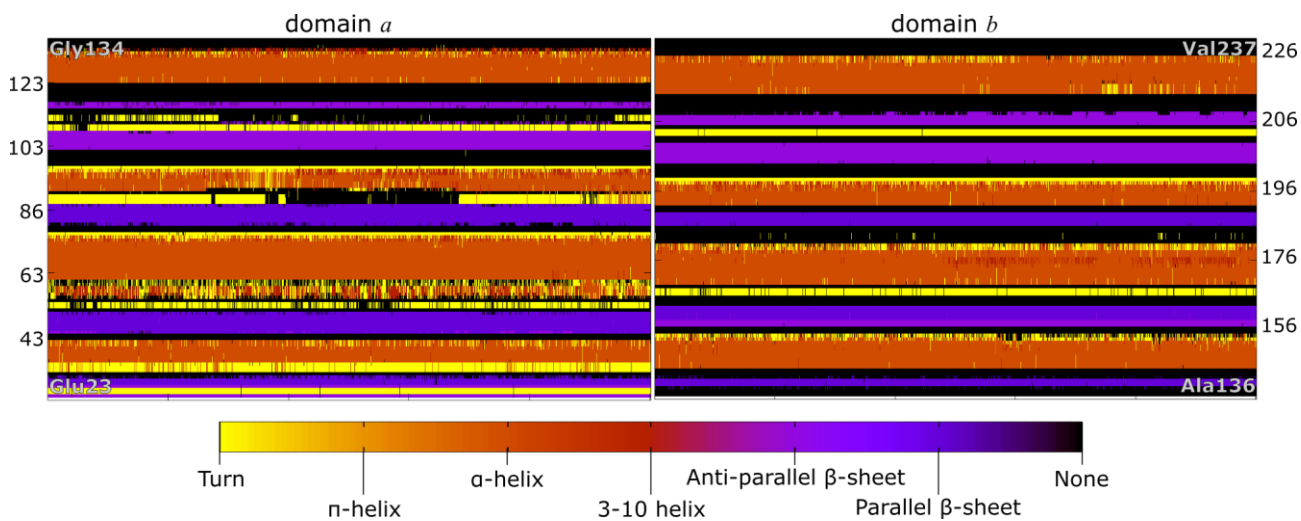

Figure S2. Secondary structure of the domain *a* and *b* of the hPDI, throughout the 50 ns cMD simulation in the NPT ensemble.

Section B – Modelling GSSG into the model of hPDI$_{ox}$

The GSSG substrate was modelled in the active site of the *a*-domain of hPDI$_{ox}$ from the x-ray structure of a homologous enzyme (glutaredoxin) bound to GSSG (PDB code: 4TR0).(7) Glutaredoxin is a thioredoxin-like enzyme that catalyzes the reduction of GSSG to GSH; like PDI, it also possesses the catalytic Cys–X–Y–Cys motif (Cys12–Pro13–Tyr14–Cys15).(5) We have aligned the heavy atoms of the backbone of the Cys–X–X–Cys motif of glutaredoxin and that of the *a*-domain of hPDI$_{ox}$, (the rmsd was as small as 0.27 Å see Figure S3) and transposed GSSG and all water molecules within 6 Å of GSSG, into PDI$_{ox}$. We have not observed clashes between these waters and hPDI.
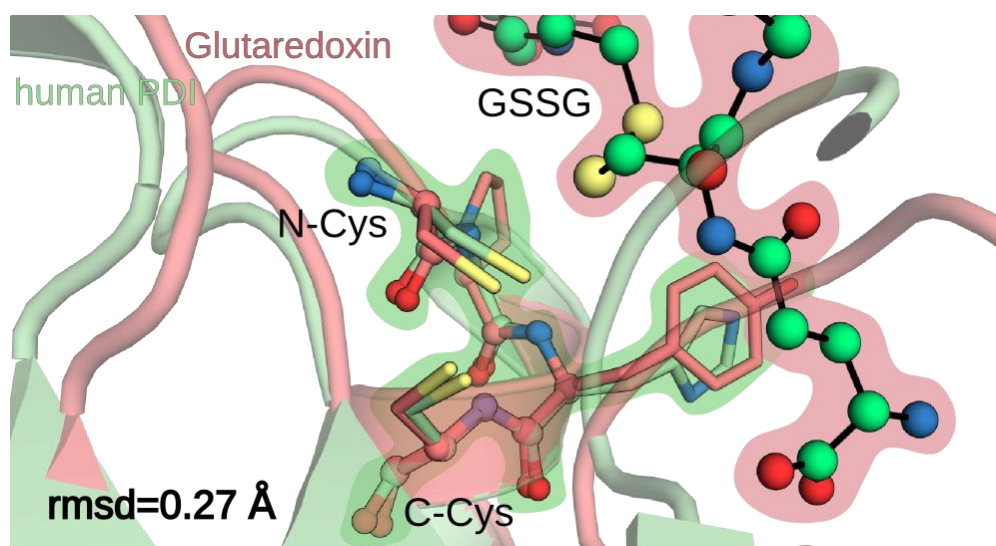

Figure S3. Result of the alignment of the backbone of the Cys12–Pro13–Tyr14–Cys15 from glutaredoxin (in pink) with that of the Cys53–Gly54–His55–Cys56 of the a-domain of hPDI (in green).

Physiological protonation states were attributed to all residues (accordingly to the results from the Propka31.py script(8, 9)), except for Cys53, Cys397 and Cys400, whose protonation states are extensively described in the literature:(2, 10-12) Cys53 was modelled as a thiolate, and the Cys397 and Cys400 pair was oxidized. All MM parameters for the enzyme were drawn from the FF99SB*ildn* force field;(13-15) the parameters for GSSG were determined for the x-ray conformation in glutaredoxin, using the Antechamber tool and the PARM99 force field,(16, 17) with Merz-Kollman (MK) charges(18) derived from a restrained electrostatic potential (RESP) fitting, from a single-point energy calculation at the HF/6-31G(d) level of theory.

To minimize the energy of our model at the molecular mechanics (MM) level of theory, we built a rectangular box solvated with TIP3P waters(19) within a radius of 12 Å from the surface of the enzyme, and neutralized the charge of the model (-39 a.u.) with 39 Na$^+$ counterions. The protocol comprehended four stages: a minimization with the whole system, except for the solvent, fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with all heavy atoms of the solute fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with the backbone of the modelled hPDI$_{ox}$ fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, and an unrestrained energy minimization of the full system. The final minimized model exhibited an rmsd of 1.03 Å relative to the x-ray model (only the heavy atoms of the backbone were considered for the calculation).

To equilibrate the solvent to the minimized x-ray conformation of the protein, we constrained all heavy atoms in the hPDI$_{ox}$:GSSG complex and x-ray water molecules with 20 kcal·mol^{-1}·Å^{-2} harmonic force constants and ran classical molecular dynamics (cMD) simulations. The system was heated during 100 ps in the *NVT* ensemble with a cMD simulation. In the first 50 ps, the temperature was linearly increased up to 310 K, and in the remaining 50 ps it was kept constant. Subsequently, we performed a 10 ns cMD simulation within an *NPT* ensemble (temperature of 310 K and pressure of 1 bar) with the solute's heavy atoms still restrained by harmonic potentials, employing the Berendsen barostat(20) and the Langevin thermostat(21). We have used the SHAKE algorithm(22) to constrain the motion in H-including bonds in order to use a 2 fs integration step, and have employed the particle-mesh-Ewald summation method (PME)(23) to account for the electrostatic interactions beyond the 10 Å *cutoff* for non-bonded interactions. We have kept the x-ray positions fixed because there is evidence that the x-ray structure, being the most common and being itself an average over the enzyme conformations in the crystal, is the most adequate to be used in single-conformation PES studies. We then took the last structure of the cMD run as the representative catalytic conformation of the solvated hPDI$_{ox}$:GSSG complex.

Section C – PES for the tests concerning the effect of the solvent

To analyze the environment around Cys53 and GSSG, we built two 2-layer QM/MM models: in the first, we included the conserved water molecules from crystallography near the active site in the DFT layer (93 atoms); in the second, we included all water molecules and the GSSG motif within 6 Å of the sulfur atoms of the Cys53-thiolate and the GSSG-disulfide in the DFT layer (141 atoms). In Figure S4, we present a representative part of the DFT layer around Cys53 and GSSG, for both test models, and the electronic energies from the linear transit scan performed along the reaction coordinate.
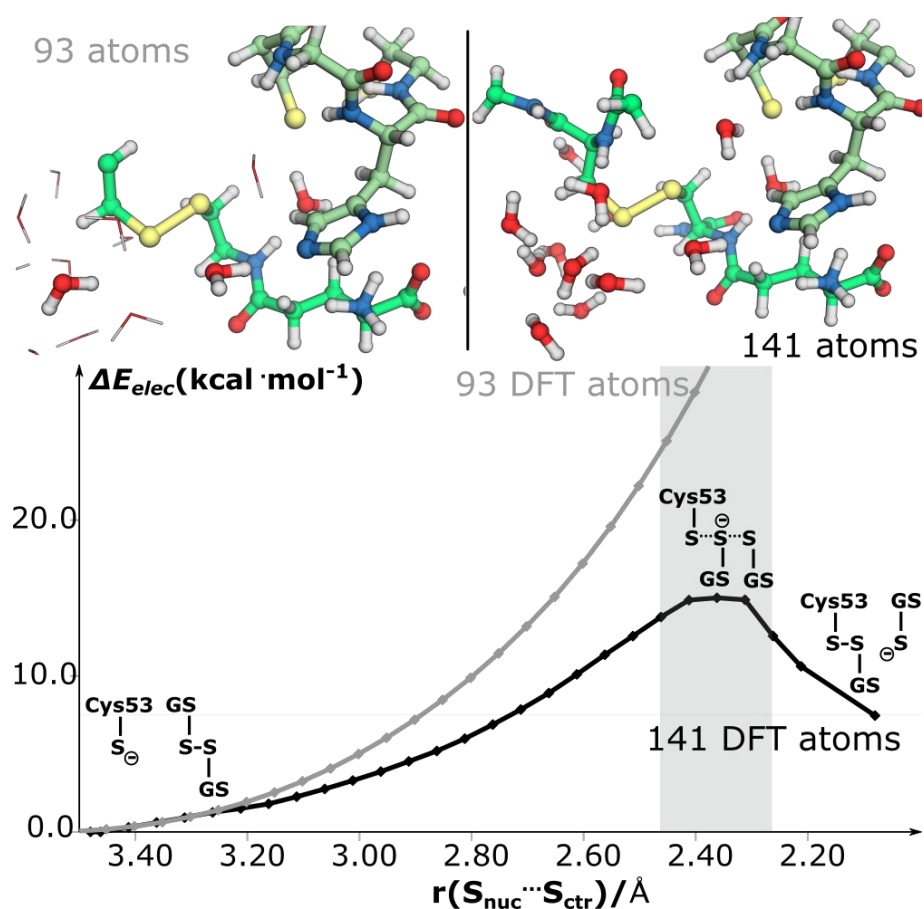

Figure S4. Electronic energy for the nucleophilic attack of the Cys53-thiolate (S_{nuc}) to the GSSG-disulfide (S_{ctr}), for two test models with a DFT layer of 93 atoms (grey) and 141 atoms (black). The grey shadowed region indicates the region where the TS for the reaction occurs.

We observe the formation of the mixed-disulfide intermediate only in the model in which we included a considerable amount of solvent and a larger portion of the substrate in the QM layer. In the smaller model, we

cannot observe a transition state, and the potential energy surface (PES) for the reaction goes far beyond the 25 kcal·mol^{-1} limit for an efficient enzymatic catalysis. These observations support the fact that the solvent plays an important role in the reaction, stabilizing the leaving negatively charged glutathione (GSH$_{lg}$) that is formed upon cleavage of the GSSG-disulfide.(24, 25) It also shows that the description of the solvent interacting with the GSH$_{lg}$ at the MM level is not good enough. This is in agreement with the general modelling rule in QM/MM calculations of including in the QM layer all atoms that directly interact with the reacting atoms.
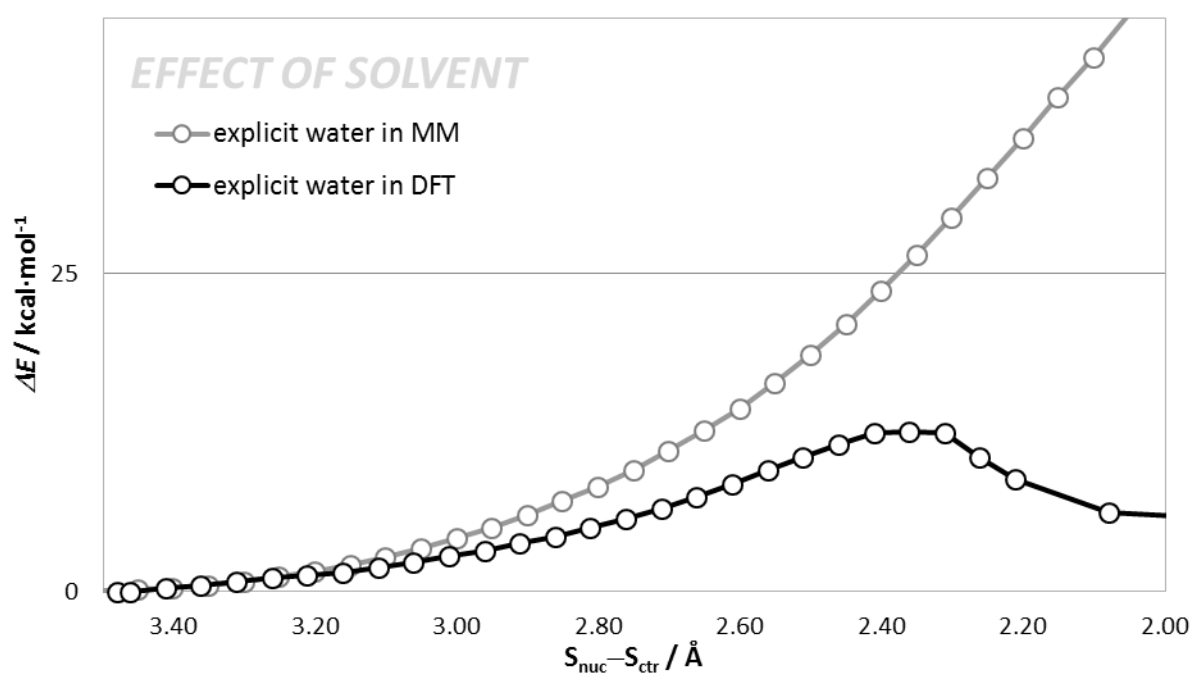

Figure S5. PES for the attack of the nucleophilic sulfur in Cys53 to the central sulfur (in GSSG), by including waters explicitly in the DFT layer (in black) or in the MM layer (grey).

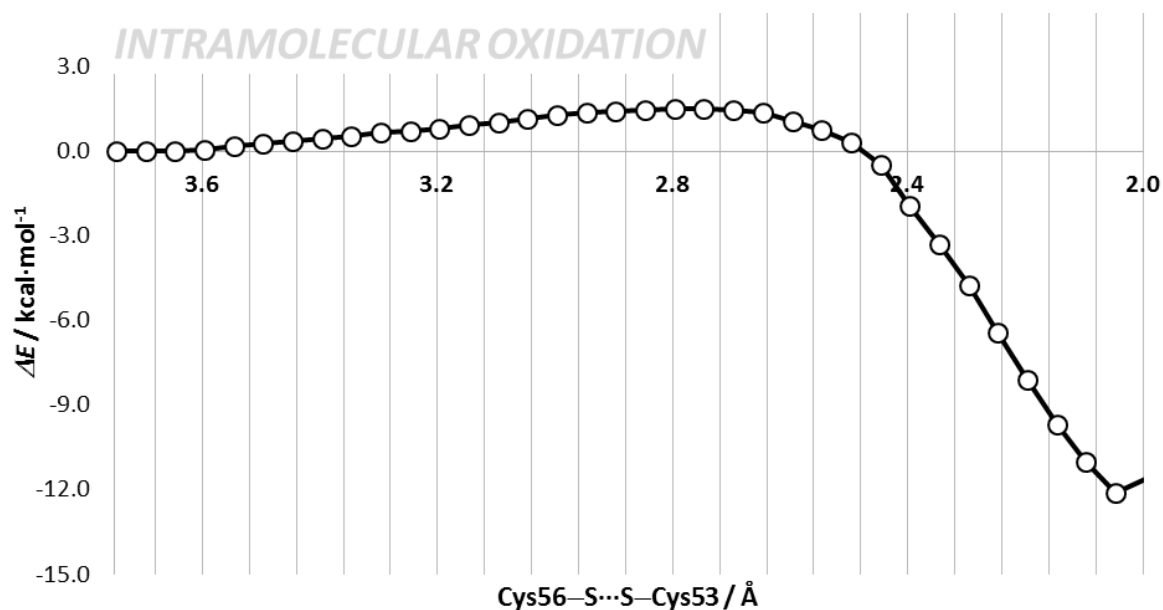# Section D – PES for the intramolecular oxidation of the *a*-domain upon inclusion of Ala50 in the DFT layer

Figure S6. PES for the intramolecular oxidation of the a-domain of hPDI upon inclusion of Ala50 and Pro51 in the DFT layer (185 atoms).

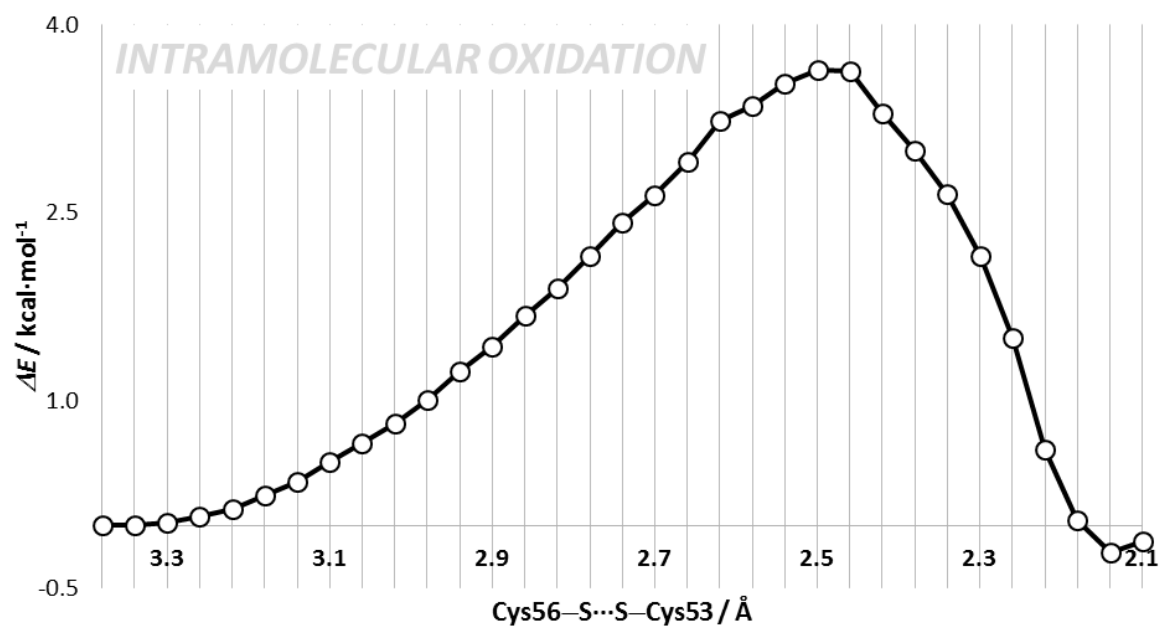# Section E – PES for the intramolecular oxidation of the *a*-domain after the exit of the GSH

Figure S7. PES for the intramolecular oxidation of the a-domain of hPDI in the absence of the hydrogen bond donors other than water molecules nearby.

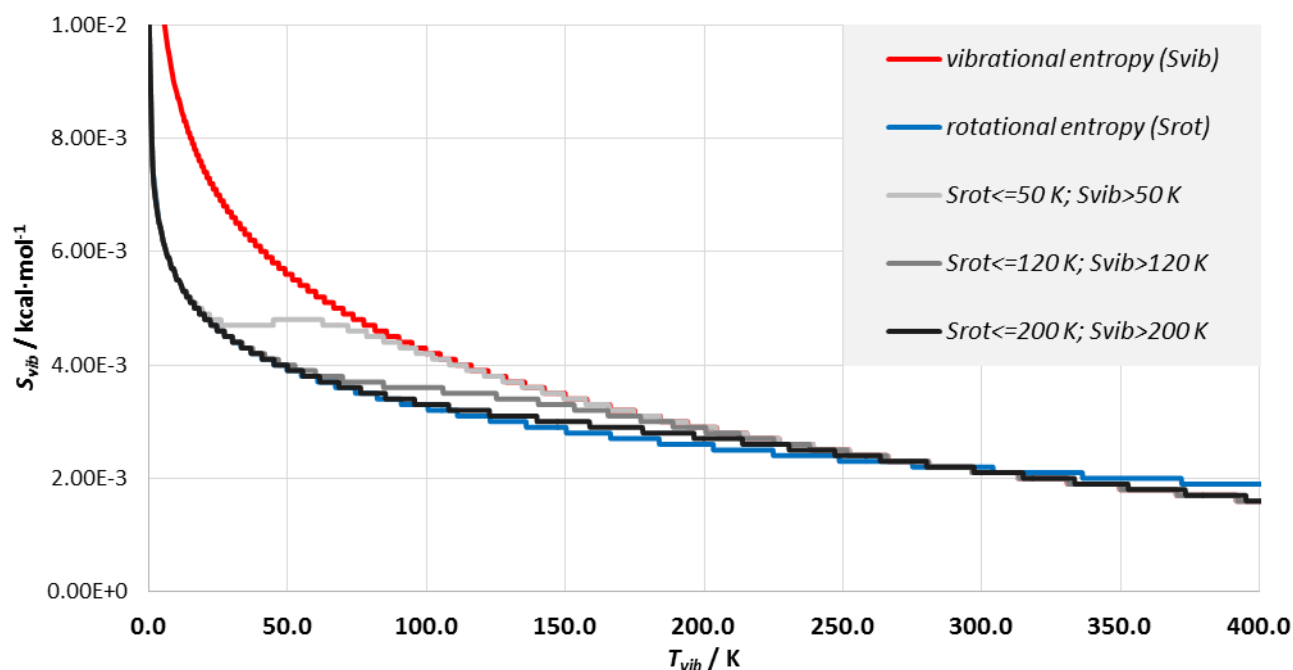## Section F – Calculation of vibrational entropy

Figure S8. Contributions for the vibrational entropy as a function of vibrational temperatures, using the harmonic oscillator approximation (in red), the free-rotor approximation (in blue), and a combination of both (in grey). In the latter we employed different thresholds (50 K, 120 K and 200 K) to change from the free-rotor to the harmonic oscillator approximation.

To calculate the vibrational entropy of each stationary point in the reaction, we first evaluated the magnitude of the lowest frequency for each state. Low vibrations usually involve motions from a large group of atoms, and can often introduce artefacts in the calculation of vibrational entropy. This is evidenced in Figure S8, where the vibrational entropy increases abruptly for short vibrational temperatures. Given the unclear interpretation of these lower vibrational temperatures, we discarded the vibrational temperatures lower than 10 K.

We also observed a few imaginary frequencies in the MM layer of the model (about 1 to 4), which should be a result of the waters kept fixed during the calculations. Hence, we compared the lowest vibrational temperature for each state of the reaction in such a way that: (1) the number of positive vibrational temperatures was the same for every state of the reaction (except for TSs in which there is one less positive vibrational temperature); (2) the lowest vibrational temperature for each state was of the same magnitude for every state. The lower vibrational temperature that we have accounted for was on average 21±2 K, which corresponds to a $T\Delta S_{vib}$ of

2.3 kcal·mol^{-1}. We accounted for 11,375 vibrational temperatures during the formation of the mixed-disulfide, and 11,377 during the intramolecular oxidation step. We emphasize that these steps comprehend a different number of atoms in the model. Table S2 sums up the results we have obtained with the harmonic oscillator approximation and the combined free-rotor/harmonic oscillator approximation for a threshold vibrational temperature of 120 K (~100 cm^{-1}, as proposed by Grimme *et al.*)(26) for the mechanism of PDI on GSSG.

Table S2. Zero-point corrections (ZPE), Gibbs energy corrections ($\Delta\Delta G$), and vibrational entropy corrections ($T\Delta S_{vib}$) for all reaction states of the catalysis of the reduction of GSSG by hPDI, calculated with the harmonic oscillator approximation (HO) and the free-rotor/harmonic oscillator approximation (FR/HO). All energies are given in kcal·mol^{-1}, and relative to the GSSG and GSH$_{ox}$-Cys53···Cys56 (pre) states.

Reaction states	HO approximation			FR/HO approximation			N_{vib}
	ZPE	$\Delta\Delta G$	$T\Delta S_{vib}$	ZPE	$\Delta\Delta G$	$T\Delta S_{vib}$	
GSSG	0.00	0.00	0.00	0.00	0.00	0.00	
TS$_{redox1}$	-0.25	-0.50	0.37	-0.25	-0.50	0.40	
GSH$_{lg}$···GSH$_{ox}$-Cys53	-0.82	-1.76	1.64	-0.82	-1.76	1.61	11,375
TS$_{rot-Cys56}$	-1.95	-3.70	2.45	-1.95	-3.14	1.89	
Cys56···WAT	-1.13	-2.64	2.23	-1.13	-2.32	1.95	
TS$_{deprot}$	-3.83	-4.27	0.34	-3.83	-4.14	0.22	
GSH$_{ox}$-Cys53	1.44	0.56	1.40	1.44	0.56	1.40	
GSH$_{ox}$-Cys53···Cys56 (pre)	0.00	0.00	0.00	0.00	0.00	0.00	
GSH$_{ox}$-Cys53···Cys56	-0.56	-1.76	1.83	-0.56	-1.26	1.30	11,377
TS$_{redox2}$	-1.76	-3.33	2.17	-1.76	-2.57	1.36	
GSH$_{red}$···Cys53-Cys56	-1.82	-3.58	2.54	-1.82	-2.76	1.67	

The difference between the two approximations is quite small; however, the harmonic approximation provides larger contributions from the vibrational entropy. We highlight that this aspect is also a result of the neglecting of vibrational temperatures lower than 10 K, since these are present in the region where the harmonic oscillator and free-rotor approximations differ the most.

We calculated also the vibrational entropy for blocks of vibrational temperatures with different sizes (starting from the highest to the lowest vibrational temperatures). In this way, we could evaluate where the contributions to the vibrational entropy were more pronounced. These contributions are presented in Figure S9.
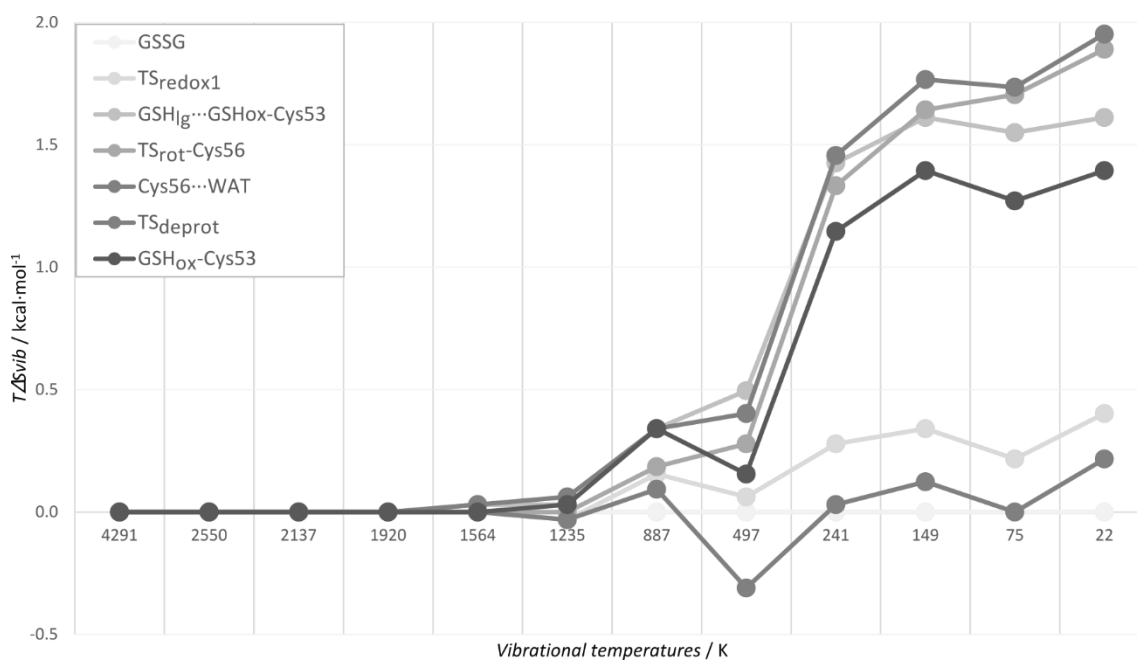
Figure S9. Vibrational entropy contributions ($T\Delta S_{vib}$) for every stage of the formation of the mixed-disulfide, calculated with the harmonic oscillator approximation for different sets of vibrational temperatures. The *x*-axis represents the lower vibrational temperature in each set.

Larger variations of the vibrational entropy were observed when vibrational temperatures between 241 K and 497 K were accounted for in the calculation, whereas the inclusion of lower frequencies did not alter the $T\Delta S_{vib}$ profile significantly. On the other hand, the neglecting of vibrational temperatures below 241 K could lead to a wrong shape of the entropy profile for the reaction. Nevertheless, the fact is that the changes in vibrational entropy observed were not very significant, and fitted well in the 3 kcal·mol[-1] error typically attributed to the study of enzyme catalysis with computational methods.

Section G – Gibbs energy profile for the catalytic mechanism of hPDI

Table S3. Gibbs energies for the catalytic cycle of the a-domain of our model of hPDI. All energies are presented in kcal·mol^{-1}; energies represented in curve brackets are given in Ha. GSSG, GSH$_{lg}$···GSH$_{ox}$-Cys53, Cys56···WAT, GSH$_{ox}$-Cys53, GSH$_{ox}$-Cys53···Cys56 (pre), GSH$_{ox}$-Cys53···Cys56 and GSH$_{red}$···Cys53-Cys56 correspond to a minimum at the mPW1N/6-31G(d):PARM99SB level of theory. TS$_{redox1}$, TS$_{rot\text{-}Cys56}$, TS$_{deprot}$ and TS$_{redox2}$ correspond to transition states at the mPW1N/6-31G(d):PARM99SB level of theory.

	Optimization and ΔG corrections	Single-point energy calculation		
	mPW1N/6-31G(d):PARM99SB	*mPW1N/6-31G(d)*	*mPW1N/6-311+G(2d,2p)*	*BB1K/6-311+G(2d,2p)*
Formation of the mixed-disulfide	GSSG	0.0 (-5670.2737)	0.0 (-5672.0045)	0.0 (-5671.4412)
	TS$_{redox1}$	16.0	17.8	18.2
	GSH$_{lg}$···GSH$_{ox}$-Cys53	9.1	9.2	8.8
	TS$_{rot\text{-}Cys56}$	10.2	10.5	9.8
	Cys56···WAT	3.3	3.6	2.2
	TS$_{deprot}$	7.4	8.2	8.7
	GSH$_{ox}$-Cys53	-9.0	-8.4	-10.6
Cleavage of the mixed-disulfide	GSH$_{ox}$-Cys53···Cys56 (pre)	0.0 (-5670.8610)	0.0 (-5672.5825)	0.0 (-5672.0191)
	GSH$_{ox}$-Cys53···Cys56	3.3	1.5	0.6
	TS$_{redox2}$	5.2	7.2	7.2
	GSH$_{red}$···Cys53-Cys56	-6.5	-6.0	-4.9

Table S4. Thermodynamic quantities (kcal·mol-1) for the complete oxidation of the a-domain of hPDI by GSSG. Energies were determined at the mPW1N/6-311+G(2d,2p):FF99SB//mPW1N/6-31G(d):FF99SB, and ZPE and thermal corrections were determined at the mPW1N/6-31G(d):FF99SB level of theory (for a temperature of 310 K and a pressure of 1 bar). Reference energies in Ha are included in brackets for the GSSG and GSHox-Cys53···Cys56(pre) states.

	Reaction stages	Thermodynamic quantities in kcal·mol^{-1}		
		ΔG	ΔH	$-T\Delta S$
Formation of mixed-disulfide	GSSG	0.0 (-5705.4522)	0.0 (-5701.8395)	0.0 (-3.6127)
	TS$_{redox1}$	18.7	19.1	-0.4
	GSH$_{lg}$···GSH$_{ox}$-Cys53	10.7	12.4	-1.7
	TS$_{rot-Cys56}$	12.5	14.4	-1.9
	Cys56···WAT	6.5	8.5	-2.0
	TS$_{deprot}$	10.1	10.3	-0.2
	GSH$_{ox}$-Cys53	-6.3	-4.9	-1.4
Cleavage of mixed-disulfide	GSH$_{ox}$-Cys53···Cys56 (pre)	0.0 (-5672.5825)	0.0 (-5663.3143)	0.0 (-3.6139)
	GSH$_{ox}$-Cys53···Cys56	1.1	2.4	-1.3
	TS$_{redox2}$	7.2	8.6	-1.4
	GSH$_{red}$···Cys53-Cys56	-6.4	-4.7	-1.7

Section H – Details on reference Gibbs energies determined from experimental considerations

The comparison between experimental kinetic rates and theoretical activation free energies in this step should take into consideration that the formation of the mixed-disulfide intermediate follows a second order kinetics (it depends on the concentration of both reduced hPDI and GSSG). Hence, our value is only comparable to experiments in the limit where GSSG is in large excess, relatively to PDI. In the endoplasmic reticulum (ER), the concentration of hPDI is in the range 200–500 μM,(27) and the redox GSH/GSSG buffer is circa 6 mM in GSH and 2 mM in GSSG; (28) furthermore, from the hPDI available in the ER, only 60% corresponds to the system used in the calculations (a-domain reduced).(29) Since there is approximately ten times more GSSG than $hPDI_{ox}$ (corresponding to a saturation of about 99.3% for $hPDI_{ox}$)(30), a first-order kinetics for the formation of the mixed-disulfide intermediate should be valid.
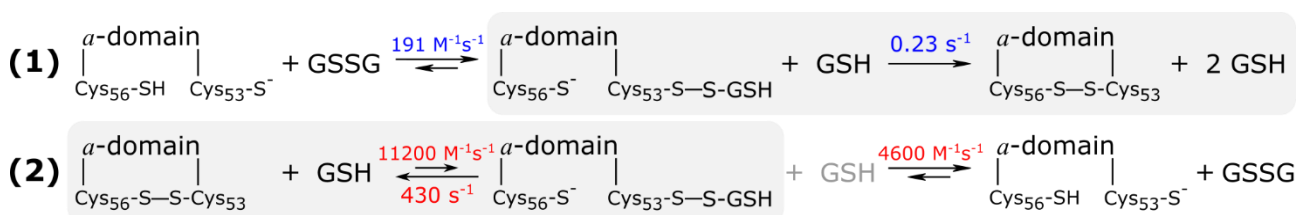
Scheme S1. Oxidation (1) and reduction (2) reactions of the a-domain of hPDI with corresponding rate constants13. The reactions marked in shade correspond to the intramolecular rearrangement of the a-domain.

Section I – Molecular dynamics simulations with the mixed-disulfide intermediate and reduced GSH

All reference coordinates were taken from the mixed-disulfide intermediate from the ONIOM calculations for the *ab*-region, and from the minimized x-ray coordinates of the hPDI$_{ox}$:GSSG complex. Physiological protonation states were attributed to all residues, except for Cys53, and Glu47, Cys397 and Cys400; Cys53 was modelled as an oxidized Cys (CYX), Glu47 was modelled in the neutral form (GLH), Cys397 and Cys400 were modelled as oxidized Cys forming a disulfide bond. All MM parameters for the enzyme were drawn from the FF99SB*ildn* force field.(13-15) Reference coordinates for the neutral GSH (GSH), the negatively charge GSH (GSM) and the oxidized GSH (GSX) were retrieved from the mixed-disulfide determined with the ONIOM method. The MM parameters for these ligands were determined using the Antechamber tool and the PARM99 force field,(16, 17) with Merz-Kollman (MK) charges(18) derived from a restrained electrostatic potential (RESP) fitting, from a single-point energy calculation at the HF/6-31G(d) level of theory. All parameter and coordinate files are included.

To minimize the energy of our model at the molecular mechanics (MM) level of theory, we built a rectangular box solvated with TIP3P waters(19) within a radius of 12 Å from the surface of the enzyme, and neutralized the charge of the model with counterions. The protocol comprehended four stages: a minimization with the whole system, except for the solvent, fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with all heavy atoms of the solute fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with the backbone of the modelled hPDI$_{ox}$ fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, and an unrestrained energy minimization of the full system. The final minimized model exhibited an rmsd of 1.03 Å relative to the x-ray model (only the heavy atoms of the backbone were considered for the calculation).

To evaluate the residence time of reduced GSH at the active site of the *ab*-domain we performed two cMD simulation: with GSH in the neutral form (GSH), and with GSH in the negatively charged form (GSM). During the cMD simulations, all heavy atoms in the hPDI$_{ox}$–GSH$_{ox}$ were constrained with 20 kcal·mol^{-1}·Å^{-2} harmonic force constants. The system was heated during 100 ps in the *NVT* ensemble with a cMD simulation. In the first 50 ps the temperature was linearly increased up to 310 K, and in the remaining 50 ps it was kept constant.

Subsequently, we performed a 20 ns cMD simulation within an *NPT* ensemble (temperature of 310 K and pressure of 1 bar) with the solute's heavy atoms still restrained by harmonic potentials, employing the Berendsen barostat(20) and the Langevin thermostat(21). We have used the SHAKE algorithm(22) to constrain the motion in H-including bonds in order to use a 2 fs integration step, and have employed the particle-mesh-Ewald summation method (PME)(23) to account for the electrostatic interactions beyond the 10 Å *cutoff* for non-bonded interactions.
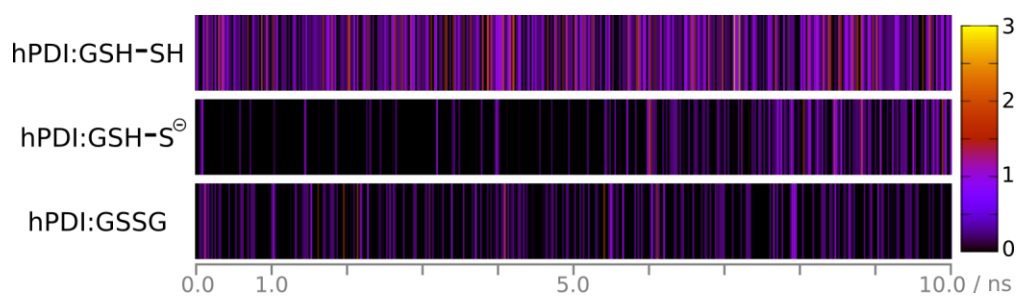

Figure S10. Number of hydrogen bonds from water molecules to the sulfur atoms of the mixed-disulfide intermediate during a 10 cMD simulation with all the atoms from the protein kept fixed.

Section J – Umbrella sampling MD simulations of the rotation of the Cys56- and GSH$_{lg}$-thiols

We built two models of the mixed disulfide intermediate to address the rotation of the Cys56- and the GSH$_{lg}$-thiols, employing the model of the final mixed-disulfide intermediate from the ONIOM calculations. In the first, Cys53 was modelled as an oxidized Cys (CYX) in a disulfide bond with GSH$_{ox}$ (GSX), Glu47 was negatively charged (GLU), and the GSH$_{lg}$ was modelled in the neutral form (GSH); while in the second, Cys53 was modelled as an oxidized Cys (CYX) in a disulfide bond with GSH$_{ox}$ (GSX), Glu47 was in the neutral form (GLH), and the GSH$_{lg}$ was modelled in the neutral form (GSH). All MM parameters for the enzyme were drawn from the FF99SB*ildn* force field.(13-15) Reference coordinates for the neutral GSH (GSH), and the oxidized GSH (GSX) were retrieved from the mixed-disulfide determined with the ONIOM method. The MM parameters for these ligands were determined using the Antechamber tool and the PARM99 force field,(16, 17) with Merz-Kollman (MK) charges(18) derived from a restrained electrostatic potential (RESP) fitting, from a single-point energy calculation at the HF/6-31G(d) level of theory. All parameter and coordinate files are included.

To minimize the energy of our model at the molecular mechanics (MM) level of theory, we built a rectangular box solvated with TIP3P waters(19) within a radius of 12 Å from the surface of the enzyme, and neutralized the charge of the model (-11) with 11 Na$^+$ counterions. The protocol comprehended four stages: a minimization with the whole system, except for the solvent, fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with all heavy atoms of the solute fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, one with the backbone of the domain *a* and all heavy atoms in domain *b* fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints, and an energy minimization of the heavy atoms in domain *b* fixed with 50 kcal·mol^{-1}·Å^{-2} harmonic constraints.

After that, we performed cMD simulation: the system was heated during 100 ps in the *NVT* ensemble with the heavy atoms in domain *b* fixed with 20 kcal·mol^{-1}·Å^{-2} harmonic constraints (in the first 50 ps the temperature was linearly increased up to 310 K, and in the remaining 50 ps it was kept constant); subsequently, we conducted a 2 ns cMD simulation within an *NPT* ensemble (temperature of 310 K and pressure of 1 bar) with the solute's heavy atoms (except couterions) still restrained by harmonic potentials, employing the Berendsen barostat(20) and the Langevin thermostat(21). We have used the SHAKE algorithm(22) to constrain the motion in H-including bonds in order to use a 2 fs integration step, and have employed the particle-mesh-Ewald summation

method (PME)(23) to account for the electrostatic interactions beyond the 10 Å *cutoff* for non-bonded interactions. The last structure from the 2 ns *NPT* cMD simulation was used as the starting structure for every umbrella window in subsequent umbrella sampling MD simulations.

During umbrella sampling MD simulations, we searched the CC–SH torsion through 365 umbrella windows every 1° in the range of -182° to 182°, to better sample the region around the -180° to 180º limits of the torsion angle. In each simulation, we confined the CC–SH torsion in Cys56 and GSH_{lg} to a harmonic well with a width of 180°, employing harmonic restraints of 0.12184 kcal·mol^{-1}.deg^{-2}. Additionally, throughout all cMD simulations the heavy atoms of domain *b* were constrained fixed 20 kcal·mol^{-1}·Å^{-2} harmonic force constants. In this way, we tried to mimic some of the constraints that could be found in our QM/MM model, and account for the absence of the *b'* and *a'* domains of PDI. Each simulation comprehended three stages: energy minimization with 2000 cycles (500 of steepest descent, and then with the conjugate gradient algorithm); *NPT* equilibration of 50 ps (at 310 K and 1 bar) with a 1 fs integration step, with the Berendsen barostat(20) and the Langevin thermostat(21), a 10.0 Å cutoff for the direct calculation of long-range interactions and the particle-mesh-Ewald summation method (PME)(23) to account for electrostatics beyond the cutoff; *NPT* production of 1 ns under the same conditions of the *NPT* equilibration, dumping the CC–SH torsion and the system's energy every 50 fs.

We also applied a set of additional constraints that were system dependent:

(1) In the CC–SH torsion of Cys56, we have also kept all heavy atoms in GSH_{lg} constrained with 20 kcal·mol^{-1}·Å^{-2} harmonic force constants, to keep its pose similarly to the one in our QM/MM model;

(2) In the CC–SH torsion of GSH_{lg}, we have compared a simulation without restraints in GSH_{lg}, and constraining the position of the terminal carbon-carbon bond in the torsion with a 20 kcal·mol^{-1}·Å^{-2} harmonic force constant. As GSH_{lg} interacts weakly with the active site of PDI, we wanted to compare the barrier for the rotation of the CC–SH torsion when it is free to diffuse into the water environment and when it is still bound to the active site of PDI.

The data dumped in the MD simulations was processed with the weighted-histogram analysis method (WHAM).(31) The data was computed from over 365 bins for a periodic system over the range of -182° to 182°,

within a tolerance of 0.01 for the free-energy between two consecutive bins. The associated errors for the free energy differences were computed from 10 Monte Carlo bootstraps.

The results of the umbrella sampling MD simulations are shown in Figure S11 and Figure S12 for the CC–SH torsion in Cys56 and GSH$_{lg}$.
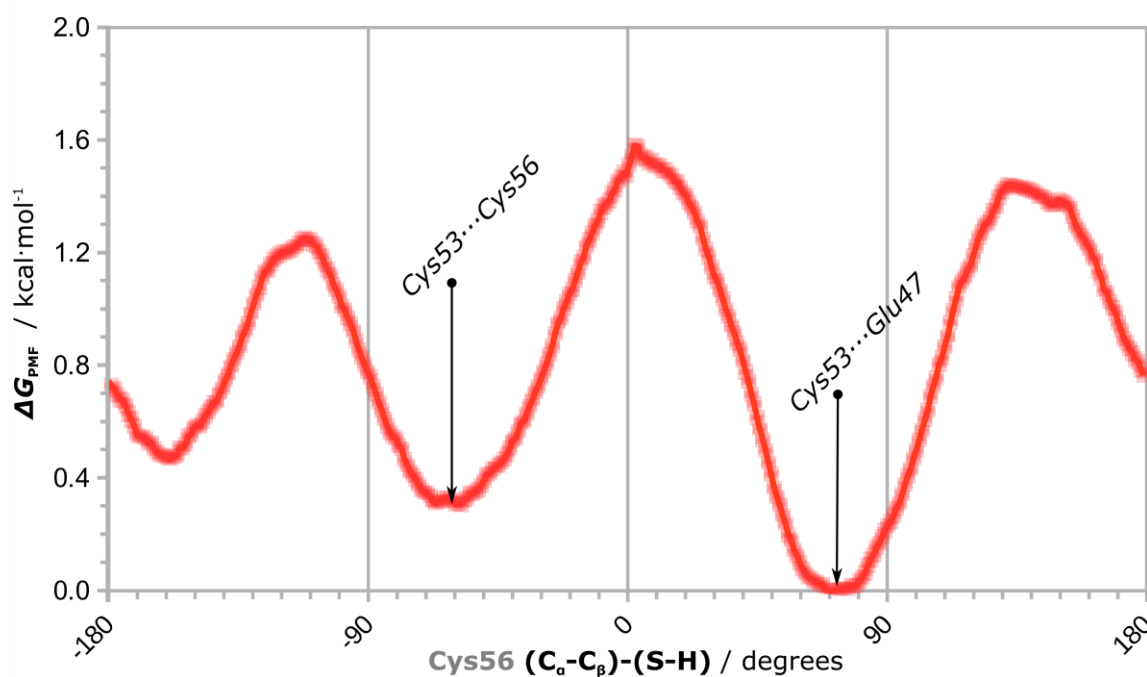

Figure S11. Gibbs energy profile for the PMF obtained from umbrella sampling MD simulations of the rotation of the CC–SH dihedral of Cys56 at the MM level, with the FF99SBildn (protein) and GAFF (glutathione) force fields. Cys53···Cys56 and Cys53···Glu47 represent the points in which the Cys56-thiol was facing Cys53 or Glu47.

The Gibbs energy barrier for the rotation of the Cys56-thiol, breaking the hydrogen bond with Cys53 and forming a hydrogen bond with Glu47, was calculated as 1.113(±0.003) (rotating in one direction) and 1.274(±0.003) kcal·mol^{-1} (rotating in the opposite direction) with umbrella sampling. This result agrees well with the Gibbs energy of 1.8 kcal·mol^{-1} that we have determined with the ONIOM methodology. Nevertheless, the very low statistical errors in the umbrella sampling calculations represent the convergence of the calculations within the subset of the phase space explored during the 365 ns of simulation. It is reasonable to consider that larger statistical uncertainties might arise from longer scale fluctuations in the enzyme structure.
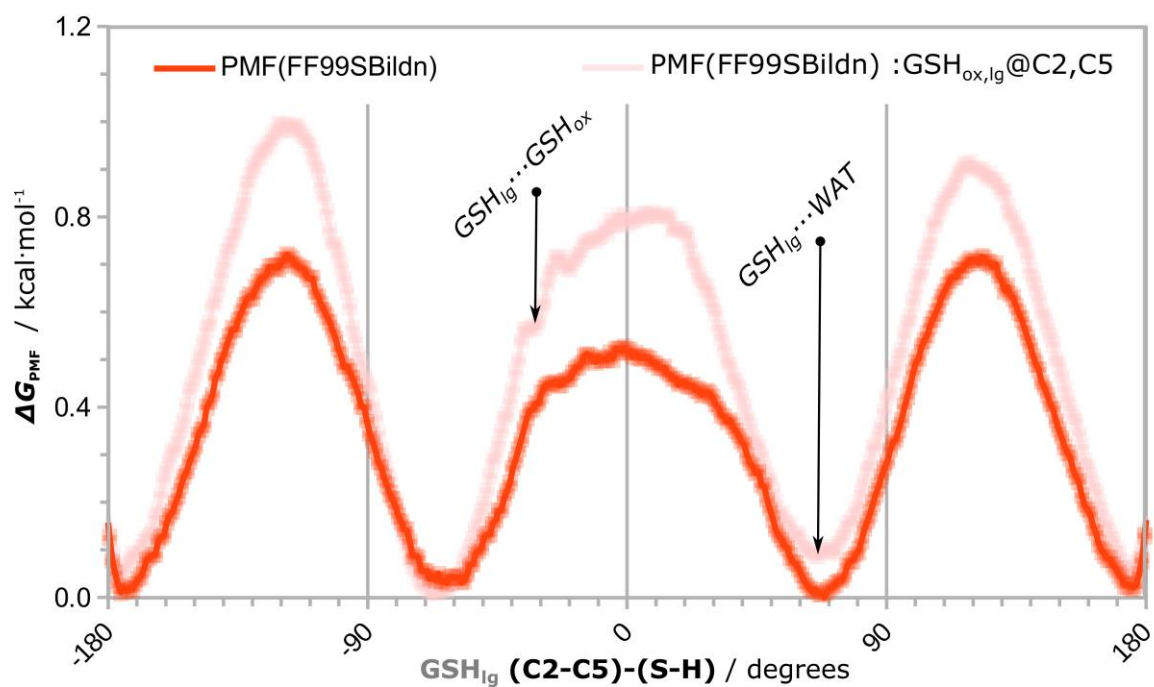
Figure S12. Gibbs energy profile for the PMF obtained from umbrella sampling MD simulations with the FF99SBildn (protein) and GAFF (glutathione) force fields for the rotation of the CC–SH dihedral of GSH$_{lg}$: (in red) with restraints on the *b*-domain of the enzyme, and (in pink) with restraints on the *b*-domain of the enzyme and on the CC bond (here named C2C5) from the CCSH torsion of GSH$_{lg}$ and the CCSS torsion of GSH$_{ox}$. GSH$_{lg}$···GSH$_{ox}$ and GSH$_{lg}$···WAT represent the points in which the GSH$_{lg}$-thiol was facing GSH$_{ox}$ or WAT.

The umbrella sampling MD simulations for the rotation of the GSH thiol show that the interaction between GSH and the active site of the *a*-domain is weak. The Gibbs barrier for the rotation of the GSH-thiol is lower than 1 kcal·mol^{-1} (between 0.487±0.009 and 0.716±0.008 kcal·mol^{-1}). However, during this simulation, GSH left the active site of the *a*-domain. As a result, we also ran umbrella sampling MD simulations with harmonic restraints of 20 kcal·mol^{-1}·Å^{-2} on the carbon-carbon bond of the CCSH torsion of the leaving group GSH and the CCSS torsion of the oxidized GSH. We did this to study the interaction between the leaving group GSH and the mixed-disulfide, upon rotation of the GSH-thiol. In this last case, we observe that the configuration in which GSH-thiol is facing the mixed-disulfide is not a minimum energy point. However, the lowest minimum energy point is close to this state (0.688±0.008 kcal·mol^{-1} below). The remaining minimum energy points present barriers of 0.718±0.007 (anticlockwise) and 0.947±0.10 (clockwise) kcal·mol^{-1} relatively to the lowest minimum energy point. Regarding the results with the ONIOM methodology, the Gibbs barrier to rotate the GSH-thiol from the solvent to the mixed disulfide was estimated as 2.5 kcal·mol^{-1}, which is higher than any of these values but still

equally low, and the difference in energy between the two states (GSH-thiol facing water and GSH-thiol facing the mixed-disulfide) was 1.1 kcal·mol-1, which is very similar to the result obtained with umbrella sampling (0.601±0.007 and 0.688±0.008 kcal·mol^{-1}, relatively to the two closest minima). We think that this also supports that intramolecular oxidation of PDI should also depend on the concentration of GSH and H-bond donors in solution.

References

1. Case DA, *et al.* (2012) AMBER 13 (University of California, San Francisco).
2. Wang C, *et al.* (2013) Structural Insights into the Redox-Regulated Dynamic Conformations of Human Protein Disulfide Isomerase. *Antioxid Redox Sign* 19(1):44-53.
3. Tian G, Xiang S, Noiva R, Lennarz WJ, Schindelin H (2006) The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* 124(1):61-73.
4. Westphal V, Darby NJ, Winther JR (1999) Functional properties of the two redox-active sites in yeast protein disulphide isomerase in vitro and in vivo. *J Mol Biol* 286(4):1229-1239.
5. Carvalho AP, Fernandes PA, Ramos MJ (2006) Similarities and differences in the thioredoxin superfamily. *Prog Biophys Mol Bio* 91(3):229-248.
6. Kozlov G, Maattanen P, Thomas DY, Gehring K (2010) A structural overview of the PDI family of proteins. *Febs J* 277(19):3924-3936.
7. Lee EH, Kim HY, Hwang KY (2014) The GSH- and GSSG-bound structures of glutaredoxin from Clostridium oremlandii. *Arch Biochem Biophys* 564:20-25.
8. Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH (2011) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *J Chem Theor Comp* 7(2):525-537.
9. Sondergaard CR, Olsson MHM, Rostkowski M, Jensen JH (2011) Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK_a Values. *J Chem Theor Comp* 7(7):2284-2295.
10. Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE (1996) Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear C-13/N-15 NMR spectroscopy. *Biochem* 35(24):7684-7691.
11. Kemmink J, *et al.* (1999) The structure in solution of the b domain of protein disulfide isomerase. *J Biomol NMR* 13(4):357-368.
12. Dijkstra K, *et al.* (1999) Assignment of H-1, C-13 and N-15 resonances of the a ' domain of protein disulfide isomerase. *J Biomol NMR* 14(2):195-196.
13. Sorin EJ Pande VS (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 88(4):2472-2493.
14. Hornak V, *et al.* (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65(3):712-725.
15. Lindorff-Larsen K, *et al.* (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78(8):1950-1958.
16. Cornell WD, *et al.* (1995) A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J Am Chem Soc* 117(19):5179-5197.
17. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21(12):1049-1074.
18. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J Phys Chem* 97(40):10269-10280.
19. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79(2):926-935.
20. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-Dynamics with Coupling to an External Bath. *J Chem Phys* 81(8):3684-3690.
21. Langevin P (1908) Sur la théorie du mouvement brownien. *C R Acad Sc* 146:530-532.
22. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23(3):327-341.
23. Ewald PP (1921) The calculation of optical and electrostatic grid potential. *Ann Phys-Berlin* 64(3):253-287.
24. Fernandes PA Ramos MJ (2004) Theoretical Insights into the Mechanism for Thiol/Disulfide Exchange. *Chem-Eur J* 10(1):257-266.

25. Neves RPP, Fernandes PA, Varandas AJC, Ramos MJ (2014) Benchmarking of Density Functionals for the Accurate Description of Thiol-Disulfide Exchange. *J Chem Theor Comp* 10(11):4842-4856.

26. Grimme S (2012) Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem-Eur J* 18(32):9955-9964.

27. Laurindo FRM, Pescatore LA, Fernandes DD (2012) Protein disulfide isomerase in redox cell signaling and homeostasis. *Free Radical Bio Med* 52(9):1954-1969.

28. Lappi AK Ruddock LW (2011) Reexamination of the Role of Interplay between Glutathione and Protein Disulfide Isomerase. *J Mol Biol* 409(2):238-249.

29. Appenzeller-Herzog C Ellgaard L (2008) In vivo reduction-oxidation state of protein disulfide isomerase: The two active sites independently occur in the reduced and oxidized forms. *Antioxid Redox Sign* 10(1):55-64.

30. Raturi A, Vacratsis PO, Seslija D, Lee L, Mutus B (2005) A direct, continuous, sensitive assay for protein disulphide-isomerase based on fluorescence self-quenching. *Biochem J* 391:351-357.

31. Grossfield A (WHAM: the weighted histogram analysis method), version 2.0.9.