Amino Acids (2013) 44:1009–1019
DOI 10.1007/s00726-012-1435-3

ORIGINAL ARTICLE

Triple mutated antibody scFv2F3 with high GPx activity: insights from MD, docking, MDFE, and MM-PBSA simulation

Quan Luo · Chunqiu Zhang · Lu Miao · Dongmei Zhang · Yushi Bai · Chunxi Hou · Junqiu Liu · Fei Yan · Ying Mu · Guimin Luo

Received: 21 August 2012 / Accepted: 27 November 2012 / Published online: 6 December 2012

Abstract By combining computational design and site-directed mutagenesis, we have engineered a new catalytic ability into the antibody scFv2F3 by installing a catalytic triad (Trp^{29}–Sec^{52}–Gln^{72}). The resulting abzyme, Se-scFv2F3, exhibits a high glutathione peroxidase (GPx) activity, approaching the native enzyme activity. Activity assays and a systematic computational study were performed to investigate the effect of successive replacement of residues at positions 29, 52, and 72. The results revealed that an active site Ser^{52}/Sec substitution is critical for the GPx activity of Se-scFv2F3. In addition, Phe^{29}/Trp–Val^{72}/Gln mutations enhance the reaction rate via functional cooperation with Sec^{52}. Molecular dynamics simulations showed that the designed catalytic triad is very stable and the conformational flexibility caused by Tyr^{101} occurs mainly in the loop of complementarity determining region 3. The docking studies illustrated the importance of this loop that favors the conformational shift of Tyr^{54}, Asn^{55}, and Gly^{56} to stabilize substrate binding. Molecular dynamics free energy and molecular mechanics-Poisson Boltzmann surface area calculations estimated the pK_a shifts of the catalytic residue and the binding free energies of docked complexes, suggesting that dipole–dipole interactions among Trp^{29}–Sec^{52}–Gln^{72} lead to the change of free energy that promotes the residual catalytic activity and the substrate-binding capacity. The calculated results agree well with the experimental data, which should help to clarify why Se-scFv2F3 exhibits high catalytic efficiency.

Keywords Abzyme · GPx mimic · Docking · MD simulation · MDFE simulation · MM-PBSA

Abbreviations

GPx	Glutathione peroxidase
MD	Molecular dynamics
MDFE	Molecular dynamics free energy
MM-PBSA	Molecular mechanics-Poisson Boltzmann surface area
ROS	Reactive oxygen species
GSH	Glutathione
Sec	Selenocysteine
EDTA	Etheylenediaminetetraacetic acid
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
CDR3	Complementarity determining region 3
RMSD	Root mean-square deviation

Introduction

Reactive oxygen species (ROS), including superoxide anions, hydroxyl radicals, hydrogen peroxide, hydroperoxide, nitric oxide, and singlet oxygen, are normal products

Electronic supplementary material The online version of this article (doi:10.1007/s00726-012-1435-3) contains supplementary material, which is available to authorized users.

Q. Luo (✉) · C. Zhang · L. Miao · D. Zhang · Y. Bai · C. Hou · J. Liu (✉)
State Key Laboratory of Supramolecular Structure and Materials, Institute of Theoretical Chemistry, Jilin University, Changchun 130012, People's Republic of China
e-mail: luoquan@jlu.edu.cn

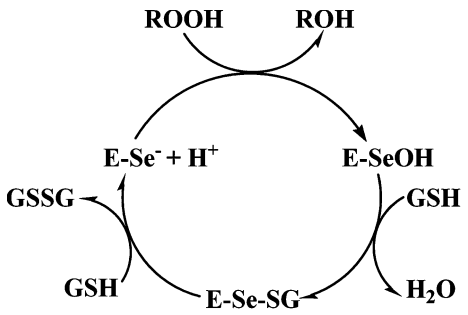
J. Liu
e-mail: junqiuliu@mail.jlu.edu.cn

F. Yan · Y. Mu · G. Luo
Key Laboratory for Molecular Enzymology and Engineering of the Ministry of Education, Jilin University, Changchun, People's Republic of China

of cellular respiration (Devasagayam et al. 2004). They display high chemical reactivity and have important roles in cell signaling and homeostasis (Thannickal and Fanburg 2000). However, an imbalance between the production of ROS and protective mechanisms may result in oxidative stress that causes serious damage to cellular membrane lipids, proteins, and nucleic acids. This oxidative stress can lead to a range of human diseases such as atherosclerosis, neurodegenerative diseases, cancer, and allergies (Matés et al. 1999; Battin and Brumaghim 2009).

A family of antioxidant selenoenzymes has evolved in mammals as a regulatory system for ROS (Liu et al. 2012). Among them, GPx is an important member that catalyzes the reduction of a variety of hydroperoxides by glutathione (GSH) to protect cells against oxidative damage (Thomas et al. 1990; Esworthy et al. 1998). Studies on the structure and catalytic mechanism of GPx reveal that a catalytic triad composed of selenocysteine (Sec), Gln and Trp is directly involved in catalysis and a perfect location of Sec in the specific GSH-binding site can greatly enhance the reaction rate (Maiorino et al. 1995). The detailed redox cycle involves the oxidation of selenol to selenenic acid, which reacts with GSH to form a selenenyl sulfide adduct, and then a second GSH attacks the sulfur of this adduct to produce a GSH disulfide (Scheme 1) (Toppo et al. 2009).

Due to the biological importance of GPx, considerable effort has been devoted to reproducing the properties of this selenoenzyme. To date, a number of artificial GPx models, ranging from small molecular compounds to macromolecular ones, have been reported to exhibit significant GPx-like functions (Bhabak and Mugesh 2010; Huang et al. 2011). Using a GSH analog (GSH-S-DN_2phBu) as an antigen, we have produced a monoclonal antibody 2F3 that is capable of binding GSH (Ren et al. 2001). The single chain fragment scFv2F3, which is derived from 2F3, can be further converted into Se-scFv2F3 by chemical modification of the reactive serine to Sec (Wu and Hilvert 1990). Like natural GPxs, Se-scFv2F3 can efficiently catalyze the reduction of hydroperoxides by GSH. Its GPx activity was found to be 3,394 U/μmol, which approaches the activity of native GPx from rabbit liver (5,780 U/μmol). Although this result is very promising, the detailed mechanism of how Se-scFv2F3 exhibits high catalytic efficiency remains unclear. The reason for this is that chemical modification has the disadvantage of unspecifically modifying other serine residues on the surface of scFv2F3 into Sec. As a result, the high GPx activity of Se-scFv2F3 may be attributed to an additive effect of several Sec residues.



Scheme 1 The catalytic cycle of GPx. *ROOH* hydroperoxide, *GSH* glutathione, *E-SeH* selenol, *E-SeOH* selenenic acid, *E-SeSG* selenenyl sulfide adduct, *GSSG* GSH disulfide

To determine which serine is the primary site for chemical modification, a theoretical model of scFv2F3 has been constructed to guide site-directed mutagenesis studies. Structural analysis suggested that Ser^{52} is an ideal site for Sec incorporation and two adjacent residues, Phe^{29} and Val^{72}, may participate in the formation of a catalytic triad after replacing them with Trp and Glu, respectively (Luo et al. 2010). In this work, we generated a series of scFv2F3 mutants with replacements at positions 29, 52, and 72 to complement the above experiment by performing activity assay on each variant. Furthermore, a systematic computational study, including MD, MDFE, docking, and MM-PBSA simulation, was carried out to help elucidate the conformational flexibility, residual activity, and substrate-binding mode/capacity of scFv2F3 mutants. These calculated results were consistent with the experimental data, which provided significant insights into the cooperative ability of the designed catalytic triad to enhance the catalytic reaction.

Theory and methods

Preparation of scFv2F3 mutants and their models

The scFv2F3 model was constructed through a series of procedures involving conformational search for side chains, end repair, splice repair, and MD simulation, as described by previous work (Luo et al. 2010). The structural quality of the final model was verified using Profile-3D and PROCHECK (Lüthy et al. 1992; Laskowski et al. 1993). The scFv2F3 mutants, *Mutant I* (Ser^{52} → Sec), *Mutant II* (Ser^{52} → Sec, Phe^{29} → Trp), *Mutant III* (Ser^{52} → Sec, Val^{72} → Gln) and *Mutant IV* (Ser^{52} → Sec, Phe^{29} → Trp, and Val^{72} → Gln), were prepared to test their catalytic activity and the corresponding models, *Model I*, *Model II, Model III, and Model IV,* were constructed using Swiss-PdbViewer (Guex and Peitsch 1997). In these mutant models, Ser^{52} was replaced by Cys rather than Sec, because Se and S share similar redox properties and the parameters of Sec are unavailable to the AMBER 11 program (Maiorino et al. 1998; Ponder and Case 2003). 10-ns MD simulations were performed using standard

AMBER ff03 force field for model refinements. Each system was solvated in a 15 Å octahedral TIP3P water box and neutralized by two chlorine ions. The time step of the simulations was 2.0 fs with a cutoff of 15 Å for the non-bonded interaction, and long range electrostatic interactions were calculated with the particle mesh Ewald method. For temperature regulation, the Langevin thermostat was used to maintain the temperature of our system at 300 K. The SHAKE algorithm was employed to constrain all bonds involving hydrogen atoms (Simmerling et al. 2002; Tuccinardi et al. 2007).

Activity measurements for scFv2F3 mutants

The gene for scFv2F3 mutants was created by QuikChange Site-Directed mutagenesis (Braman et al. 1996) and the proteins were expressed and purified to measure their activities. The GPx activity was determined using a previously described method (Wilson et al. 1989). The reaction was carried out at 37 °C in 700 µl of solution containing 50 mM PBS (pH 7.0), 1 mM EDTA (etheylenediaminetetraacetic acid), 1 mM sodium azide, 1 mM GSH, 0.25 mM NADPH (reduced nicotinamide adenine dinucleotide phosphate), 1 U of glutathione reductase, and 10–50 µM of Se-scFv2F3. The Se-scFv2F3 was pre-incubated with GSH, NADPH, and glutathione reductase. The reaction was initiated by the addition of 0.5 mM H_2O_2. The activity was determined by the decrease of NADPH absorption at 340 nm using a UV–vis spectrophotometer. 50 mM PBS (pH 7.0) was used as a control sample. The activity unit (U) is defined as the amount of abzyme that consumes 1 µmol of NADPH per min. The specific activity is expressed in U/µmol.

pK_a Shifts calculations for the catalytic residue

The residual activity of the catalytic residue depends on its ionisation state and could be affected by the change of the active site microenvironment. Using a MDFE simulation, the proton pK_a shift for the sulfhydryl group of Cys^{52} in each mutant model relative to a reference model *N*-acetyl-Cys-*N*-methylamide was calculated.

Scheme 2 shows two proton binding reactions (Kollman 1993). The standard reaction free energy ΔG related to the equilibrium constant K_a is given by the following formula:

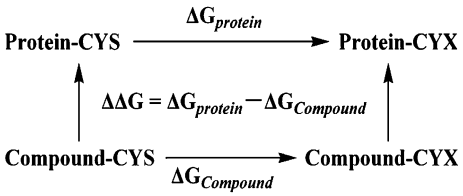
$$\Delta G = -kT \ln K_a. \quad (1)$$

pK_a can be written as:

$$pK_a = -\log_{10} K_a = \frac{1}{2.303kT}\Delta G \quad (2)$$

and

$$pK_{a,protein} = pK_{a,compound} + \frac{1}{2.303kT}\Delta\Delta G \quad (3)$$



Scheme 2 Thermodynamics cycle to analyze pK_a shifts

where $pK_{a,protein}$ and $pK_{a,compound}$ are the pK_a of the SH group of Cys in the mutant model and the reference model, respectively. $\Delta\Delta G = \Delta G_{protein} - \Delta G_{compound}$ is the double free energy difference for the thermodynamic cycle.

Two protonation states of *N*-acetyl-Cys-*N*-methylamide (ACE–CYS–NME and ACE–CYX–NME) were created. Based on the assumption that only the charges local to the sulfhydryl proton change on deprotonation, we set the charges of hydrogen and sulfur atom to match AMBER/ff03 force field values, and then folded the charges in all the other atoms into the beta-carbon atom (Scheme 3). In this way, two protonation states of Cys^{52} were also created for the mutant models.

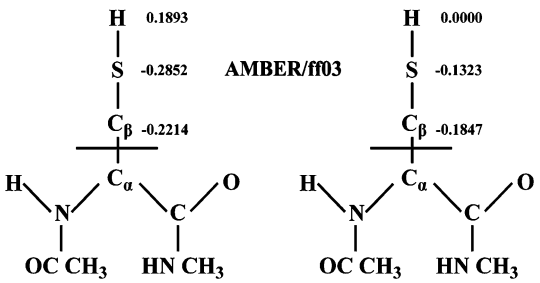
The proton binding can be described as the introduction of a new particle into the protein model. The system gradually transforms from an initial state A (protonated Cys) to final state B (ionized Cys) (Hummer and Szabo 1996). The energy function U can be modified and expressed as a linear function of a progress variable

$$U(\lambda) = (1-\lambda)U_A + \lambda U_B \quad (4)$$

where λ is the 'coupling parameter', and U_A and U_B are the energy functions of state A and state B, respectively. Varying λ from 0 to 1, the free energy change ΔG can be estimated by numerical integration of $\partial G/\partial\lambda$

$$\Delta G = \int_0^1 \langle \partial G/\partial\lambda \rangle_\lambda d\lambda \quad (5)$$

However, using some given weights (w_i) and quadrature points (λ_i), Gaussian quadrature formulas of higher order are more efficient at calculating ΔG



Scheme 3 The reference model (*N*-acetyl-Cys-*N*-methylamide) with partial charges in the protonated state (*left*) and ionized state (*right*)

$$\Delta G \approx \sum_{i=1}^{n} w_i \langle \partial G / \partial \lambda \rangle_{\lambda_i} \quad (6)$$

The thermodynamic integration calculations were started using the ff03 force field and GB continuum solvent model (Simonson et al. 2004; Bashford and Case 2000). First, a quick minimization stage consisting of 750 steps, of which the first 500 were used for the steepest descent and the last 250 for the conjugate gradient, was done to remove bad contacts. Then, free energy simulations were performed at 300 K using five λ values (0.0, 0.11270, 0.5000, 0.88729 and 1.0000), and 3 ns simulations were run for each λ value to equilibrate the system. The temperature was controlled by Langevin dynamics method (Pastor et al. 1988) and the time step was 2.0 fs with non-cutoff for the non-bonded interaction. SHAKE was applied to keep all bonds involving hydrogen atoms rigid.

Molecular docking and binding free energy calculations

The structure of GSH was optimized at the B3LYP/6-31G level by Gaussian 03 program (Frisch et al. 2003). InsightII/Affinity was used to fulfill the molecular docking procedure between GSH and the scFv2F3 mutant models (Luo et al. 2008). The binding free energy (ΔG_{bind}) of each docked complex was analyzed by MM-PBSA method (Kollman et al. 2000). This method combines the molecular mechanics, Poisson Boltzmann electrostatics, surface accessible calculations, and normal mode analyses for the entropy. The binding free energy of enzyme–substrate, ΔG_{bind}, is expressed as

$$\Delta G_{\text{bind}} = \Delta G_{\text{bind,vacuum}} + \Delta G_{\text{solv,complex}} - \Delta G_{\text{solv,enzyme}} - \Delta G_{\text{solv,substrate}} \quad (7)$$

where $\Delta G_{\text{bind,vacuum}}$ is the free energy of association of enzyme and substrate in the gas phase, and $\Delta G_{\text{solv,enzyme}}$, $\Delta G_{\text{solv,substrate}}$, and $\Delta G_{\text{solv,complex}}$ are the solvation free energies of the abzyme models, substrate GSH, and abzyme–GSH complex, respectively. $\Delta G_{\text{bind,vacuum}}$ can be described by the following formula:

$$\Delta G_{\text{bind,vacuum}} = \Delta E_{\text{MM}} - T\Delta S \quad (8)$$

where ΔE_{MM} is the difference in the sum of bond, angle, dihedral, electrostatic, and van der Waals (vdW) energies between products and reactants. The solvation free energy comprises electrostatics and non-polar free energy:

$$\Delta G_{\text{solv}} = \Delta G_{\text{PB}} + \Delta G_{\text{nonpolar}} \quad (9)$$

The snapshots of each system were sampled from the last 500 ps single trajectory with an interval of 2 ps. The electrostatic contribution to the solvation free energy is calculated using the PBSA program. A 0.5 Å grid size was used and the dielectric constant for the solute and solvent were set to 1 and 80, respectively. The non-polar contribution to the solvation free energy is computed with the MOLSURF program. The solute entropic contribution was estimated using the NMODE module. All simulations were performed using AMBER 11. (Li et al. 2007; Fujiwara and Amisaki 2008; Lu et al. 2012).

Results and discussion

Construction of scFv2F3 model

Based on the results of sequence alignment and analysis (Fig. S1), the V_{H} domain model of scFv2F3 (amino acid from Gln^{1} to Ser^{118}) was built using three templates (PDB ID: 1NQB, 1LMK, and 2GKI). The overall quality of scFv2F3 structure was examined by Profile-3D and Procheck (Fig. S2). Figure 1 shows the structure of scFv2F3 model, which is comprised of two beta sheets packed closely against each other in a compressed antiparallel beta barrel. The predicted modification site Ser^{52} is positioned after strand β5. Two absolutely conserved cysteine residues (Cys^{22} and Cys^{96}), located in the β3 and β8, respectively, can form a disulfide linkage to stabilize the conserved framework regions. Analysis and characterization of the active site of scFv2F3 model shows that only six residues, including Phe^{29}, Thr^{30}, Met^{34}, Ile^{51}, Ser^{52}, and Val^{72}, are positioned in the binding cleft. Among them, Phe^{29}, Ser^{52}, and Val^{72} are in close proximity to each other, with distances ranging from 3.05 to 4.65 Å. Such distances are favorable for the formation of a hydrogen-bonding network after substitution of the three residues with Sec^{52}, Trp^{29}, and Gln^{72}, respectively.

Determination of GPx activity

The GPx activity of each variant is summarized in Table 1. In the absence of a catalytic residue Sec, scFv2F3 is essentially inactive. The Ser^{52}/Sec^{52} mutation, however, causes a significant increase in GPx activity. A comparison of the activity of *Mutant I* (2,359 U/μmol) and the chemically modified scFv2F3 (3,394 U/μmol) reveals that Sec^{52} contributes to almost 70 % of GPx activity relative to all Sec residues. This means that Ser^{52} is the predominant modification site and the high residual activity of Sec^{52} may be attributed to its favorable position in an active site to attack substrate. On the basis of *Mutant I*, a further substitution of either Phe^{29} (Phe → Trp) or Val^{72} (Val → Gln) yields a more active abzyme, with GPx activities of 2,555 and 2,637 U/μmol, respectively. This is

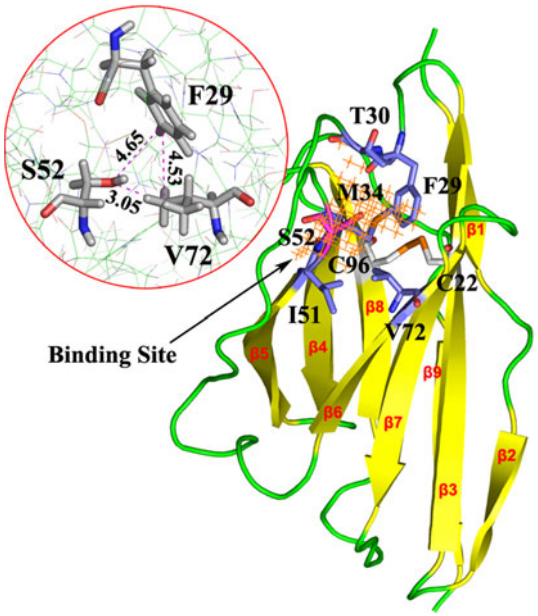


Fig. 1 The structure of scFv2F3 model and close-up view of the preferred mutation sites Phe[29], Ser[52], and Val[72]. Ser[52] is the key modification site and a disulfide bond is formed between Cys[22] and Cys[96]. The *grid points* represent the binding site of scFv2F3 model. The distances are marked in *magenta dotted lines* (unit: Å)

due to the proton donating potential of Trp and Gln that facilitate the dissociation of Sec[52], leading to a much easier nucleophilic attack on substrate. In contrast, *Mutant III* shows a greater improvement in GPx activity than *Mutant II*, suggesting that the NH group of Gln favors the formation of a more acidic SeH group of Sec[52]. As for *Mutant IV*, the triple variant exhibits a dramatically enhanced GPx activity as compared to *Mutant I*. GPx activity for *Mutant IV* is 4,235 U/μmol, an increase of 1,876 U/μmol from *Mutant I*. The increase in activity is larger than the addition of the respective contribution of Phe[29]/Trp[29] and Val[72]/Gln[72] mutations. This indicates that the designed catalytic triad Sec[52], Trp[29], and Gln[72] functions synergistically and is directly involved in catalysis.

The stability and flexibility of Se-scFv2F3 model

10-ns MD simulations of each mutant model were initiated at 300 K to collect a conformational assemble of structures. The results revealed that the flaps of *Model IV* spontaneously rearranged from one form (A-form) to another (B-form) and exhibited significantly increased fluctuations after 5.5 ns (See Fig. S3). Figure 2a shows the difference between A-form and B-form conformations, where the CDR3 loop (amino acid residue 99–107) in the B-form structure has its tip twisted relative to its conformation in the A-form structure. This conformational change is similar to the largest conformational changes induced by antigen

Table 1 The GPx activity of all mutants of scFv-2F3

Mutants	Residue 1	Residue 2	Residue 3	GPx activity (U/μmol)
scFv2F3	Phe[29]	Ser[52]	Val[72]	ND
Mutant I	Phe[29]	Sec[52]	Val[72]	2,359
Mutant II	Trp[29]	Sec[52]	Val[72]	2,555
Mutant III	Phe[29]	Sec[52]	Gln[72]	2,637
Mutant IV	Trp[29]	Sec[52]	Gln[72]	4,235
Se-scFv2F3				3,394
Native GPx				5,780

Residues 1, 2, and 3 are the mutation sites

ND not-detectable GPx activitym, *Se-scFv2F3* selenium-containing abzyme created by chemical modification

binding (Stanfield and Wilson 1994), suggesting that the plasticity of the CDR3 loop may be an important source of diversity of antigen-binding sites. Root mean square fluctuation (RMSF) analysis also shows that the residues within the CDR3 loop exhibit larger fluctuations than those in other regions (See Fig. S4), revealing the high flexibility of this loop. A notable consequence of the conformational rearrangement is that Tyr[101] in B-form structure is in a radically different position than in A-form structure. In the A-form structure, Tyr[101] is packed against Tyr[32] and Tyr[102] to fill the space adjacent to N-terminal (Fig. 2b). Conversely, Tyr[101] in B-form structure lies adjacent to the CDR2 loop; however, it also makes close contact with two tyrosine residues: the side chain of Tyr[101] interacts with Tyr[33] and Tyr[50] (Fig. 2c). The similarity of environment of Tyr[101] in both structures reveals that Tyr[101] may be of central importance in determining which conformation the CDR3 loop adopts.

Although the conformation of the CDR3 loop changes dramatically, the designed catalytic triad is not disrupted during the rearrangement event. Figure S3 shows the root mean-square deviation (RMSD) of residue Trp[29], Cys[52], and Gln[72] versus simulation time. The results indicate that there is no significant structural fluctuation for the designed catalytic triad during rearrangement. The Cα RMSD values remains at ~0.4 Å through the whole simulation, which indicates that Trp[29], Cys[52], and Gln[72] are able to provide a stable scaffold for its side chains to contact each other. In addition, the dihedral angle distribution of the side chains of the three residues were monitored (Fig. 3). The chi1 and chi2 of Trp[29] remain stable with dihedral angles of about 175° and 250°, respectively. The chi1 of Cys[52] starts about 200° during the first 200 ps and then increases sharply to about 300°, indicating that the simulation procedure improves the unfavorable conformation of Cys[52]. The chi1 and chi2 of Gln[72] stay approximately constant during the entire simulation, but the chi3 undergoes multiple

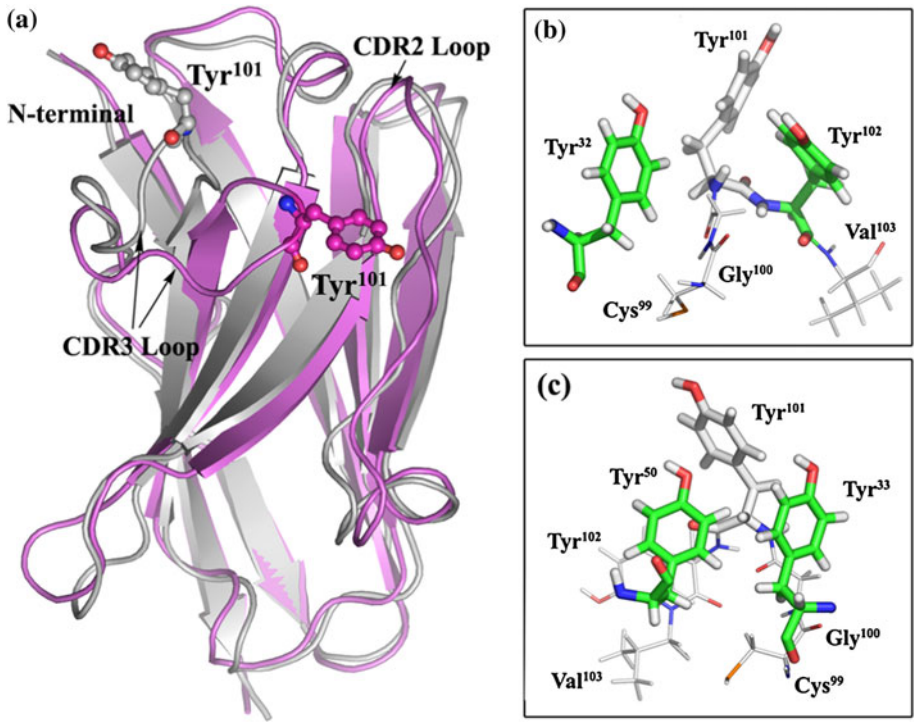


Fig. 2 **a** Structural superposition for two conformations of Se-scFv2F3 isolated from the MD trajectory: A-form (*white*), B-form (*magenta*). The difference between two kinds of conformations is in the CDR3 loop. **b** Environment of the Tyr^{101} in the CDR3 loop of A-form structure. **c** Environment of the Tyr^{101} in the CDR3 loop of B-form structure (color figure online)

conversions between about 280° and 75°. The corresponding conformations of the side chain of Gln^{72} are shown in Fig. 4a and b, respectively. According to the trajectory analysis, the former angle distribution presented for about 80 % of the trajectory appears to be the major components of the overall assemble. This finding is consistent with two conformations of the glutamine in the crystal structure of native GPx, in which the conformation of Gln that stabilizes the Cys using its amide oxygen (Fig. 4c) is more prevalent than another conformation using the amide nitrogen to stabilize Cys (Fig. 4d).

Structural comparison of Se-scFv2F3 and native GPx reveals an enzyme-like active site within the antibody scaffold. As shown in Fig. 5a, with the exception of Gln^{72} positioned at strand β6, the members of the catalytic triad are located in loops: Trp^{29} located between β3 and β4 and Cys^{52} positioned in a very sharp turn after strand β5. This is consistent with the local structural features in native GPx4 (Fig. 5b). Table S1 lists the detailed distance measurements. The distances between the sulfur atom of Cys^{52} and its potential binding partners (Trp^{29} and Gln^{72}) are 3.38 and 3.19 Å, which are shorter than the average distances (3.91 and 3.65 Å) in native GPx, suggesting that Trp^{29}, Cys^{52}, and Gln^{72} are in close proximity so as to form a catalytic triad. Arg^{152}, which is near the catalytic triad in GPx4 (Scheerer et al. 2007), is conserved in all GPx-isoforms and participates in the formation of a positively charged surface to attract substrate molecules to active site. Interestingly, this residue was also detected at a similar position in Se-scFv2F3. The distances between the guanidyl N of Arg98 to the backbone atoms N and O of Trp29 are 7.79 and 8.33 Å, close to the average distances of that in the native GPxs (7.44 and 7.45 Å).

The residual activity of the catalytic residue

Five sets of 3-ns free energy simulation were run for each system at λ value of 0.0, 0.11270, 0.50, 0.88729 and 1.0 for determination of the pK_a values of sulfhydryl group (Fig. S5–S9). As previously discussed (Simonson et al. 2004), implicit solvent models frequently give quite good correlations between calculated solvation free energies and experimental pK_a values, thus the GB dielectric continuum solvent was used to simulate the effects of the aqueous solution.

Table 2 lists the derivative of the free energy versus λ, $\partial G/\partial\lambda$. Using five λ values and the corresponding weight ($w_i = 0.2777$, $\lambda = 0.11270$ or 0.88729; $w_i = 0.44444$, $\lambda = 0.50$), the free energy difference (ΔG) of the reference model was estimated to be −5.28981 kcal/mol. The ΔG values of mutant models were slightly lower, with $\Delta G = -5.53444$ kcal/mol for *Model I*, −5.68381 kcal/mol for *Model II*, −5.69049 kcal/mol for *Model III*, and −5.88009 kcal/mol for *Model IV*. These results suggest

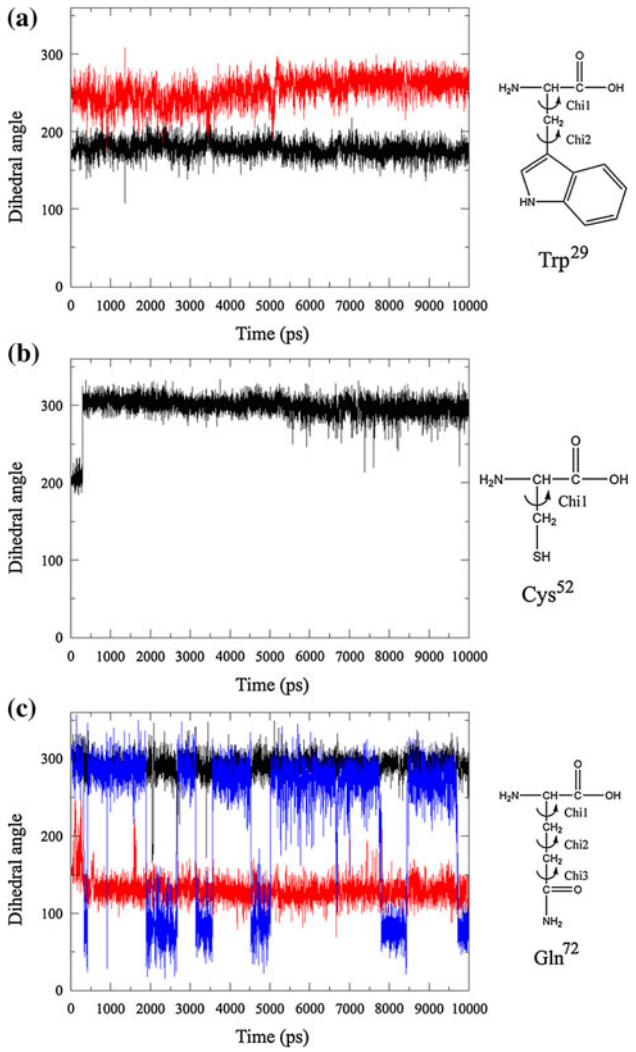


Fig. 3 The dihedral angle distribution of Trp^{29} (**a**), Cys^{52} (**b**), and Gln^{72} (**c**) during 10 ns MD simulation. chi1: *black*; chi2: *red*; chi3: *blue* (color figure online)

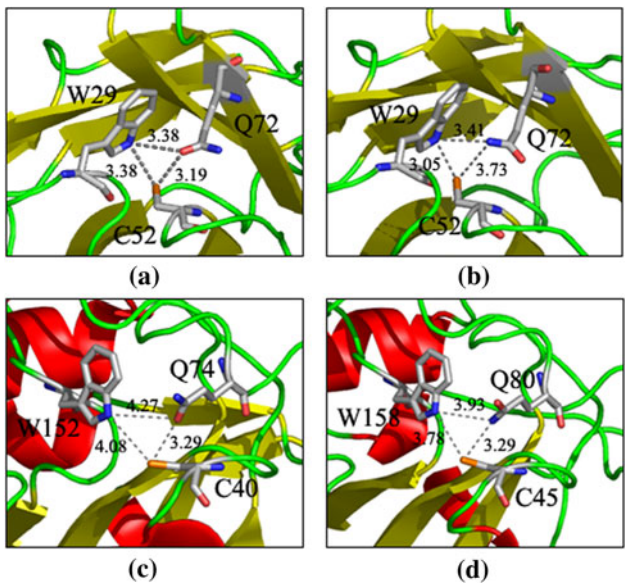


Fig. 4 The conformation of the catalytic triad in Se-scFv2F3 and native GPx where Gln interacts with Cys using amide *O* or amide *N*. **a** Se-scFv2F3 (Gln^{72} with the dihedral angle chi3 of about 280°). **b** Se-scFv2F3 (Gln^{72} with the dihedral angle chi3 of about 75°). **c** Human GPx (PDB ID: 2HE3). **d** Bovine erythrocyte GPx (PDB ID: 1GP1, U45C mutant). The distances are marked in *black dotted lines* (unit: Å)

that the interactions between the thiolate ion of Cys^{52} and its binding partners make the ionized state more favorable. To evaluate the pK_a shift of Cys, the double free energy difference of *Model I*, *Model II*, *Model III*, and *Model IV* relative to the model compound was calculated, giving $\Delta\Delta G = -0.24463$, -0.39400, -0.40068, -0.59028 kcal/mol, respectively. From the computational results, the free energy changes were found to be sensitive to the environment of the ionizable Cys^{52}. *Model I*, with the smallest value of $\Delta\Delta G$, has a relatively unfavorable environment for the ionized state of Cys^{52}, due to the non-specific interaction of the cysteine thiolate anion with the side chains of Phe^{29} and Val^{72}. However, *Model IV* has the most negative value of $\Delta\Delta G$ and a favorable environment for the ionized state of Cys^{52}, indicating that hydrogen-bonding between the ionized thiol and the side chain micro-dipoles of Trp^{29} and Gln^{72} renders ionisation energetically favorable. As for *Model II* and *Model III*, there is no significant difference between the dipole effect of Trp^{29} and Gln^{72} on thiol ionisation because the calculated $\Delta\Delta G$ of *Model III* is only 0.00668 larger than that of *Model II*. This is in accordance with the experimental data showing that the GPx activity of *Mutant III* is only by 3 % greater than that of *Mutant II*.

Before calculating pK_a values for Cys^{52}, we hypothesized that Cys in the model compound has a pK_a value of 8.30, equal to the pK_a value of Cys (Huber and Criddle 1967). As shown in Table 2, the pK_a shifts of Cys in these mutant models were found to decrease linearly with $\Delta\Delta G$. The pK_a of *Model I* was estimated to be 8.12, 0.18 U lower than that of free Cys, indicating that Cys^{52} is in the thiolate form at physiological pH. In addition, mutation of either Phe^{29} to Trp^{29} or Val^{72} to Gln^{72} results in a decrease in pK_a, showing downward shifts by 0.11 or 0.10 U relative to *Model I*, respectively. This suggests that the role of Trp^{29} and Gln^{72} in catalysis is to stabilize the thiolate form of Cys^{52}. As for *Model IV*, pK_a value was calculated to be 7.87, and the pK_a downshift relative to *Model I* is 0.25 U. This value is larger than the sum of the pK_a downshift (0.21 U) of *Model II* and *Model III*, suggesting a cooperative effect of Phe^{29} and Val^{72} on thiol ionization. The smaller the value of pK_a, the larger the extent of dissociation. Thus, the catalytic efficiency of the abzyme can be adjusted by the designed catalytic triad when the active ionized form Cys^{52} acts as nucleophile. Based on the pK_a calculations, *Model IV* should be the abzyme with the best

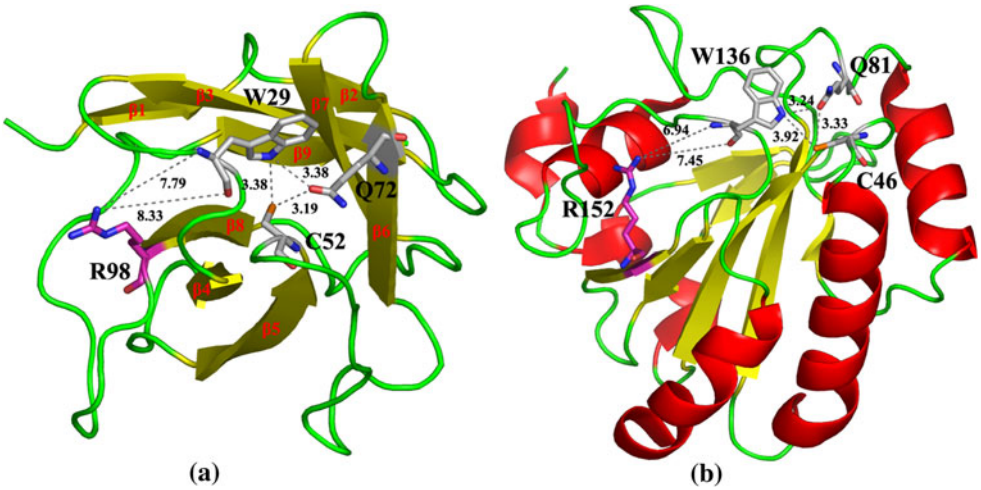


Fig. 5 Structural comparison of Se-scFv2F3 with native GPx. A similar structural motif consisting of Cys^{52}, Trp^{29}, Gln^{72}, and Arg^{98} was observed in the active site. **a** Se-scFv2F3; **b** human GPx4 (PDB ID: 2OBI). The distances are marked in *black dotted lines* (unit: Å)

Table 2 The MDFE results of the model compound and each mutant model

	Small model compound	*Model I*	*Model II*	*Model III*	*Model IV*
$\partial G/\partial\lambda$ ($\lambda = 0.0$)	−5.13751	−4.81513	−4.76394	−4.89037	−5.20626
$\partial G/\partial\lambda$ ($\lambda = 0.11270$)	−5.17571	−4.92972	−5.18335	−5.37815	−5.61459
$\partial G/\partial\lambda$ ($\lambda = 0.5$)	−5.24217	−5.58824	−5.77082	−5.62567	−5.95416
$\partial G/\partial\lambda$ ($\lambda = 0.88729$)	−5.48052	−6.05348	−6.04545	−6.10695	−6.02750
$\partial G/\partial\lambda$ ($\lambda = 1.0$)	−5.56249	−6.11459	−6.21772	−6.56226	−6.22460
ΔG	−5.28981	−5.53444	−5.68381	−5.69049	−5.88009
$\Delta\Delta G$		−0.24463	−0.39400	−0.40068	−0.59028
$\Delta pK_a = \Delta\Delta G/2.303kT$		−0.17820	−0.28700	−0.29187	−0.42998
pK_a value of Cys	8.30	8.12	8.02	8.01	7.87

All energies in kcal/mol. For the free energies, it is calculated by Gaussian integration

ΔG the derivative of the free energy, $\Delta\Delta G$ change in the derivative of the free energy, ΔpK_a change in dissociation constant, pK_a dissociation constant

GPx activity, with *Model III* and *Model II* also showing enhanced GPx activity. This is in accordance with the rank order of activity test: *Mutant IV* > *Mutant III* > *Mutant II* > *Mutant I*. Nevertheless, the theoretical results are not comprehensive in describing the system. The GPx activity of *Mutant IV* is ~80 % greater than that of *Mutant I* and approaches native GPx activity (5,780 U/μmol). Compared with the pK_a shift of *Model I*, the calculated pK_a shift for *Model IV* seems to be too small to cause such a drastic change in activity. Therefore, the activity of Se-scFv2F3 cannot simply be attributed to the dissociation of the nucleophile.

Molecular docking and MM-PBSA analysis

The refined GSH model was docked into the binding cleft of these mutant models (Fig. 6a, b). As shown in Fig. 6c and d, the replacement of Phe^{29} and Val^{72} with bulkier residues (Trp^{29} and Gln^{72}) does not cause steric hindrance between GSH and the residues of the binding site. GSH can interact with these mutant models in a similar manner. In these enzyme–substrate complexes, one side of the substrate GSH makes close contact with Ser^{74}, with a hydrogen bond between the carboxyl group of GSH and the hydroxy group of Ser^{74}. The carboxyl group of GSH on the other side of the substrate is able to form a hydrogen bond with the amide group of Asn^{55} and the backbone NH of Gly^{56}. The binding contact is further reinforced by the interactions with the side chain of Ile^{51}, The^{71}, and Ala^{57}. Moreover, the planar peptide bond between the glutamine and cysteine of GSH is stabilized by the hydrogen bond with the carbonyl oxygen of Tyr^{54} to maintain a suitable orientation of the sulfhydryl group of GSH for initiating the nucleophilic attack. Compared to GSH in *Model I* and *Model II*, GSH in *Model III* and *Model IV* adopts a tighter conformation, suggesting that substrate binding is strongly influenced by replacement of Val^{72} with Gln^{72}. The superimposition of the mutant models reveals that Tyr^{54}, Asn^{55}, and Gly^{56} on the CDR2 loop were translated about 1 Å toward the CDR3 loop (Fig. 6b). This shift is probably due to the highly flexibility of the CDR3 loop responding to the conformational change of substrate, which gives greater weight to the previous hypothesis that the CDR3 loop is responsible for the diversity of antigen-binding sites.

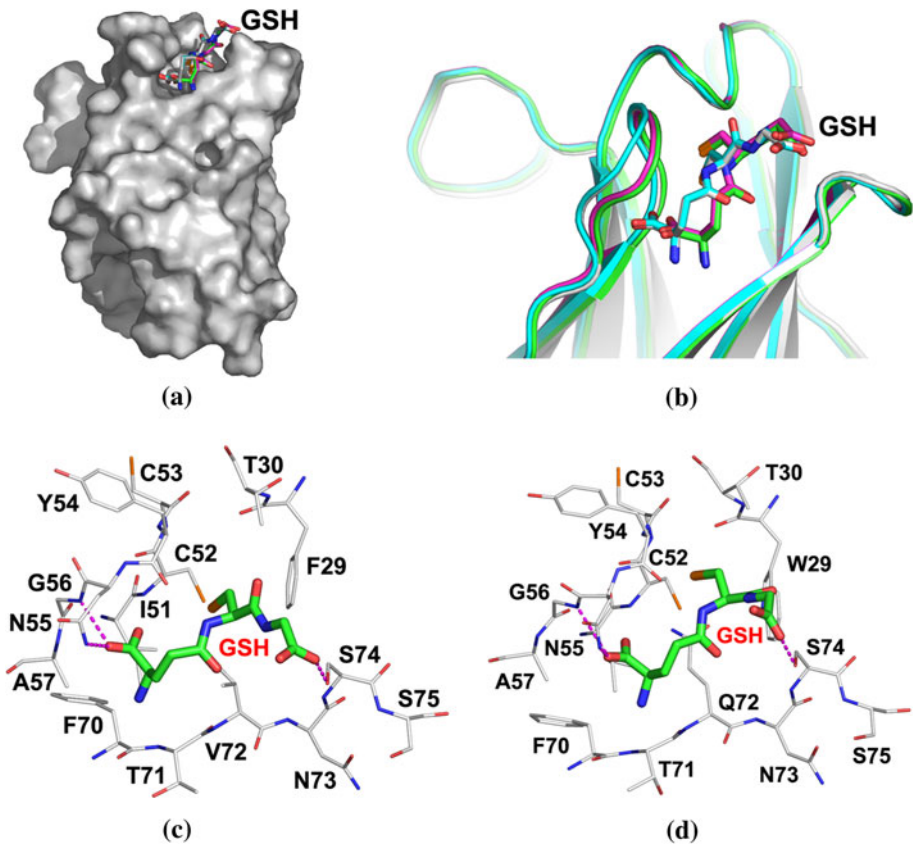


Fig. 6 The conformation of substrate GSH in the binding cleft of the mutant models (**a**). A comparison of the binding mode of GSH with different mutant models by superimposing the coordinates of backbone Cα atoms. *Model I*: *cyan*, *Model II*: *white*, *Model III*: *magenta*, *Model IV*: *green* (**b**). The detailed binding models of abzyme–substrate complexes: *Model I*-GSH (**c**) and *Model IV*-GSH (**d**) (color figure online)

To characterize the binding affinities of docked complexes, the MM-PBSA method was used to calculate the binding free energy components. Table 3 summarizes the binding free energies along with the contribution of individual energy terms. The estimated total binding free energy is negative for all of the complexes, indicating favorable binding contacts. Both the electrostatic energy (ΔE_{ELE}) and the van der Waals energy (ΔE_{VDW}) are important for binding. The non-polar solvation energy ($\Delta E_{\mathrm{PBSUR}}$) provides slightly favorable contributions, but the electrostatic contribution to solvation ($\Delta E_{\mathrm{PBCAL}}$) leads to a significant loss of the total binding free energy. The resulting balance of the ΔE_{ELE} and $\Delta E_{\mathrm{PBCAL}}$ contribution, $\Delta E_{\mathrm{PBELE}}$, is unfavorable to the total binding free energy for all systems except *Model IV*. In addition, unfavorable entropic contributions ($T\Delta S$) are observed for all complexes.

The order of the binding free energy is *Model IV* < *Model III* < *Model II* < *Model I*. *Model IV* has the most favorable binding free energy, which is −12.91 kcal/mol, about 4.5 kcal/mol more negative than *Model III*. The binding free energies of *Model II* and *Model I*, with values of −0.46 and −0.17 kcal/mol, are significantly higher than that of *Model IV*. This result indicates that both the single and double mutations cause a small rise in the substrate-binding capacity, but a cooperative effect between the catalytic triad amino acids Trp, Cys, Gln at position 29, 52, and 72 greatly enhances the affinity of the abzyme. The calculated electrostatic contributions to the binding free

Table 3 Results of the free energy calculation by MM-PBSA and NMODE

	Model I	*Model II*	*Model III*	*Model IV*
ΔE_{ELE}	−17.92	−15.82	−19.33	−201.37
ΔE_{VDW}	−23.39	−22.85	−27.44	−19.53
ΔE_{INT}	0.00	0.00	0.00	0.00
ΔE_{GAS}	−41.31	−38.67	−46.77	−220.90
$\Delta E_{\mathrm{PBSUR}}$	−3.67	−3.55	−3.73	−3.62
$\Delta E_{\mathrm{PBCAL}}$	26.18	22.42	22.61	190.38
$\Delta E_{\mathrm{PBSOL}}$	22.51	18.87	18.88	186.76
$\Delta E_{\mathrm{PBELE}}$	8.25	6.60	3.28	−10.99
$\Delta E_{\mathrm{PBTOT}}$	−18.80	−19.80	−27.89	−34.14
TΔS	−18.63	−19.34	−19.40	−21.23
ΔG	−0.17	−0.46	−8.49	−12.91

All energies are in kcal/mol

ΔE_{ELE} electrostatic energy, ΔE_{VDW} van der Waals energy, ΔE_{INT} internal energy, ΔE_{GAS} total gas phase energy ($\Delta E_{\mathrm{ELE}} + \Delta E_{\mathrm{VDW}} + \Delta E_{\mathrm{INT}}$), ΔE_{PBSUR} non-polar solvation energy, ΔE_{PBCAL} polar solvation energy, ΔE_{PBSOL} total solvation energy ($\Delta E_{\mathrm{PBSUR}} + \Delta E_{\mathrm{PBCAL}}$), ΔE_{PBELE} total electrostatic energy ($\Delta E_{\mathrm{ELE}} + \Delta E_{\mathrm{PBCAL}}$), ΔE_{PBTOT} binding free energy calculated from the terms above ($\Delta E_{\mathrm{GAS}} + \Delta E_{\mathrm{PBSOL}}$), $T\Delta S$ entropy contribution to binding, ΔG total change of binding free energy ($\Delta E_{\mathrm{PBTOT}} - T\Delta S$)

energies for the four complexes are quite different, with $\Delta E_{ELE} = -17.92$, -15.82, -19.33, -201.37 kcal/mol for *Model I*-GSH, *Model II*-GSH, *Model III*-GSH, and *Model IV*-GSH, respectively. Correspondingly, the electrostatic contribution to solvation calculated for the four complexes is also very different, ranging from 22.42 to 190.38 kcal/mol. Although the total electrostatic energy (ΔE_{PBELE}) is offset by the increase of the electrostatic component of solvation, the ΔE_{PBELE} of *Model IV*, with a value of -10.99 kcal/mol, is still much lower than that of other models, suggesting that the binding ability of the electrostatic components of *Model IV* provides a thermodynamic benefit to GSH. The increased and positive values of ΔE_{PBELE} for other mutants are most likely due to the lack of cooperative effects in the active site to offset desolvation penalties. The calculated van der Waals contributions to the binding free energies (ΔE_{VDW}) for the four complexes are relatively close to each other, with values ranging from -27.44 to -19.53 kcal/mol. This is due to the similar binding modes and minor structural differences of these docked complexes that cause a small change in van der Waals energy. Moreover, the calculated entropy contributions for all of the complexes are also very close to each other, ranging from -21.23 to -18.63 kcal/mol. This is reasonable because the entropic contribution is mainly determined by the sizes and shapes of the abzyme and substrate involved in the binding. Overall, the calculated binding free energies provide a semi-quantitative estimate of substrate binding capacity, and the trend of binding affinity is *Model IV* > *Model III* > *Model II* > *Model I*, which agree well with the GPx activity of scFv2F3 mutants.

Conclusions

Based on a theoretical model of scFv2F3, the catalytic triad (Trp^{29}–Sec^{52}–Gln^{72}) has been precisely incorporated into the antibody to yield a GPx-like abzyme (Se-scFv2F3) with high catalytic efficiency. A series of scFv2F3 mutants with replacements at position 29, 52, and 72 were constructed to study the GPx activity of Se-scFv2F3 affected by the designed catalytic triad. MD simulations and molecular docking clearly demonstrated that the conformation of Trp^{29}–Sec^{52}–Gln^{72} is not interfered by the dynamic behavior of the CDR3 loop. Furthermore, the structural change of this loop triggered by Tyr^{101} is only associated with the diversity of antigen-binding sites. MDFE simulation and MM-PBSA analysis indicated that the designed catalytic triad functions synergistically and is directly involved in changing the residual activity and substrate recognition of this abzyme to increase its reaction rate. These results presented an alternative strategy to design novel artificial enzymes with efficient and quite new activities by modifying pre-existing substrate-binding sites on protein scaffolds.

Acknowledgments This work was financially supported by the Natural Science Foundation of China (No: 21234004, 91027023, 20921003, 21004028) and 111 project (B06009). We gratefully acknowledge professor David A. Case et al. for giving us the AMBER 11 software as a freeware.

References

Bashford D, Case DA (2000) Generalized born models of macromolecular solvation effects. Annu Rev Phys Chem 51:129–152

Battin EE, Brumaghim JL (2009) Antioxidant activity of sulfur and selenium: a review of reactive oxygen species scavenging, glutathione peroxidase, and metal-binding antioxidant mechanisms. Cell Biochem Biophys 55:1–23

Bhabak KP, Mugesh G (2010) Functional mimics of glutathione peroxidase: bioinspired synthetic antioxidants. Acc Chem Res 43:1408–1419

Braman J, Papworth C, Greener A (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol Biol 57:31–44

Devasagayam TP, Tilak JC, Boloor KK, Sane KS, Ghaskadbi SS, Lele RD (2004) Free radicals and antioxidants in human health: current status and future prospects. J Assoc Physicians India 52:794–804

Esworthy RS, Swiderek KM, Ho YS, Chu FF (1998) Selenium-dependent glutathione peroxidase-GI is a major glutathione peroxidase activity in the mucosal epithelium of rodent intestine. Biochim Biophys Acta 1381:213–226

Frisch MJ, Trucks GW, Schlegel HB (2003) Gaussian 03 (Revision A.1) Gaussian Pittsburgh

Fujiwara S, Amisaki T (2008) Identification of high affinity fatty acid binding sites on human serum albumin by MM-PBSA method. Biophys J 94:95–103

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723

Huang X, Liu XM, Luo Q, Liu JQ, Shen JC (2011) Artificial selenoenzymes: designed and redesigned. Chem Soc Rev 40:1171–1184

Huber RE, Criddle RS (1967) Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. Arch Biochem Biophys 122:164–173

Hummer G, Szabo A (1996) Calculation of free-energy differences from computer simulations of initial and final states. J Chem Phys 105:2004–2010

Kollman P (1993) Free energy calculations: applications to chemical and biochemical phenomena. Chem Reviews 93:2395–2417

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE III (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Li LW, Uversky NV, Dunker K, Meroueh SO (2007) A computational investigation of allostery in the catabolite activator protein. J Am Chem Soc 129:15668–15676

Liu JQ, Luo GM, Mu Y (2012) Selenoproteins and mimics. Springer, Berlin

Lu SY, Jiang YJ, Zou JW, Wu TX (2012) Effect of double mutations K214/A-E215/Q of FRATide on GSK3β: insights from molecular dynamics simulation and normal mode analysis. Amino Acids 43:267–277

Luo Q, Han WW, Zhou YH, Yao Y, Li ZS (2008) The 3D structure of the defense-related rice protein Pir7b predicted by homology modeling and ligand binding studies. J Mol Model 14:559–569

Luo Q, Zhou YH, Yao Y, Li ZS (2010) Theoretical design of catalytic domain of abzyme Se-scFv2F3 by introducing a catalytic triad. Chem Res Chin Univ 26:118–121

Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356:83–85

Maiorino M, Aumann KD, Brigelius-Flohé R, Doria D, van den Heuvel J, McCarthy J, Roveri A, Ursini F, Flohé L (1995) Probing the presumed catalytic triad of selenium-containing peroxidases by mutational analysis of phospholipid hydroperoxide glutathione peroxidase (PHGPx). Biol Chem Hoppe-Seyler 376:651–660

Maiorino M, Aumann KD, Brigelius-Flohé R, Doria D, van den Heuvel J, McCarthy J, Roveri A, Ursini F, Flohé L (1998) Probing the presumed catalytic triad of selenium-containing peroxidases by mutational analysis. Z Ernahrungswiss 37:118–121

Matés JM, Pérez-Gómez C, Núñez de Castro I (1999) Antioxidant enzymes and human diseases. Clin Biochem 32:595–603

Pastor RW, Brooks BR, Szabo A (1988) An analysis of the accuracy of Langevin and molecular dynamics algorithms. A Mol Phys 65:1409–1419

Ponder JW, Case DA (2003) Force fields for protein simulations. Adv Prot Chem 66:27–85

Ren X, Gao S, You D, Huang H, Liu Z, Mu Y, Liu J, Zhang Y, Yan G, Luo G, Yang T, Shen J (2001) Cloning and expression of a single-chain catalytic antibody that acts as a glutathione peroxidase mimic with high catalytic efficiency. Biochem J 359:369–374

Scheerer P, Borchert A, Krauss N, Wessner H, Gerth C, Höhne W, Kuhn H (2007) Structural basis for catalytic activity and enzyme polymerization of phospholipid hydroperoxide glutathione peroxidase-4 (GPx4). Biochemistry 46:9041–9049

Simmerling C, Strockbine B, Roitberg AE (2002) All-atom structure prediction and folding simulations of a stable protein. J Am Chem Soc 124:11258–11259

Simonson T, Carlsson J, Case DA (2004) Proton binding to proteins: pKa calculations with explicit and implicit solvent models. J Am Chem Soc 126:4167–4180

Stanfield RL, Wilson IA (1994) Antigen-induced conformational changes in antibodies: a problem for structural prediction and design. Trends Biotechnol 12:275–279

Thannickal VJ, Fanburg BL (2000) Reactive oxygen species in cell signaling. Am J Physiol Lung Cell Mol Physiol 279:L1005–L1028

Thomas JP, Maiorino M, Ursing F, Girutti AW (1990) Protective action of phospholipid hydroperoxide glutathione peroxidase against membrane-damaging lipid peroxidation. In situ reduction of phospholipid and cholesterol hydroperoxides. J Biol Chem 265:454–461

Toppo S, Flohé L, Ursini F, Vanin S, Maiorino M (2009) Catalytic mechanisms and specificities of glutathione peroxidases: variations of a basic scheme. Biochim Biophys Acta 1790:1486–1500

Tuccinardi T, Manetti F, Schenone S, Martinelli A, Botta M (2007) Construction and validation of a RET TK Catalytic domain by homology modeling. J Chem Inf Model 47:644–655

Wilson SR, Zucker PA, Huang RRC, Spector A (1989) Development of synthetic compounds with glutathione peroxidase activity. J Am Chem Soc 111:5936–5939

Wu ZP, Hilvert B (1990) Selenosubtilisin as a glutathione peroxidase mimic. J Am Chem Soc 112:5647–5648