PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2019, **21**, 5499

Investigating coordination flexibility of glycerophosphodiesterase (GpdQ) through interactions with mono-, di-, and triphosphoester (NPP, BNPP, GPE, and paraoxon) substrates†

Gaurav Sharma,[a] Qiaoyu Hu,[a] Vindi M. Jayasinghe-Arachchige,[a] Thomas J. Paul,[a] Gerhard Schenk[b] and Rajeev Prabhakar*[a]

Received 13th November 2018,
Accepted 14th February 2019

DOI: 10.1039/c8cp07031h

rsc.li/pccp

In this study, interactions of the catalytically active binuclear form of glycerophosphodiesterase (GpdQ) with four chemically diverse substrates, *i.e.* NPP (a phosphomonoester), BNPP and GPE (both phosphodiesters), and paraoxon (a phosphotriester) have been investigated using all-atom molecular dynamics (MD) simulations. The roles of metal ions and key amino acid residues, coordination flexibility, and dynamic transformations in all enzyme–substrate complexes have been elucidated. The roles of important first and second coordination shell residues in substrate binding and coordination flexibility of the enzyme suggested by simulations are supported by experimental data. The chemical nature of the substrate is found to influence the mode of binding, electrostatic surface potential, metal–metal distance, and reorganization of the active site. The experimentally proposed association between the substrate binding and coordination flexibility is analyzed using principal component analysis (PCA), movements of loops, and root-mean-square-fluctuations (RMSF) as parameters. The PCA of these substrates provides different energy basins, *i.e.* one, three, two and five for NPP, BNPP, GPE, and paraoxon, respectively. Additionally, the area of an irregular hexagon (268.3, 288.9, 350.8, and 362.5 Å^2) formed by the residues on these loops illustrates their distinct motions. The substrate binding free energies of NPP, BNPP, and GPE are quite close (22.4–24.3 kcal mol^{-1}), but paraoxon interacts with the smallest binding free energy (14.1 kcal mol^{-1}). The metal binding energies in the presence of these substrates are substantially different, *i.e.* the lowest for NPP and the highest for paraoxon. These results thus provide deeper insight into the chemical promiscuity and coordination flexibility of this important enzyme.

I. Introduction

Glycerophosphodiesterase (GpdQ) from *Enterobacter aerogenes* is a binuclear metallohydrolase that catalyzes the hydrolysis of the phosphoester bond of a wide range of critical molecules (Fig. 1). This enzyme possesses unprecedented versatility and degrades representatives from each of the three phosphate ester groups (*i.e.* mono-, di-, and triesters) such as 4-nitrophenyl phosphate (NPP), bis(4-nitrophenyl) phosphate (BNPP), glycerophosphoethanolamine (GPE) and several organophosphates, including diethyl 4-nitrophenyl phosphate (paraoxon) (Fig. 1b).[1–6] In general, phosphoester bond hydrolyzing enzymes have been implicated in several important processes such as DNA replication, bone turnover, iron transport and the generation of reactive oxygen species (ROS).[7–16] However, due to its substrate promiscuity, GpdQ has also been proposed as a promising agent for the remediation of pesticides and the deactivation of nerve agents,[17,18] similar to the phosphotriesterase from *Pseudomonas diminuta* (PTE)[19] or the organophosphate-degrading enzyme from *Agrobacterium radiobacter* (OpdA).[20–22]

The X-ray structures of the Co^{2+}- and Fe^{2+}-substituted forms of GpdQ have been resolved at 1.9 Å (PDB ID: 3D03) and 2.2 Å (PDB ID: 2ZO9), respectively.[23] They show that this enzyme is structurally similar to other α/β sandwich binuclear phosphoesterases like purple acid phosphatases (PAPs) and Mre11 nuclease.[24–26] Its active site contains two metal centers, known as α and β sites. The α metal ion coordinates to Asp8, His10, and His197, and the β metal to Asn80, His156, and His195 (Fig. 1a).[23] Both metal ions are also bridged through residue Asp50 and a hydroxyl molecule. This binuclear active site is identical to that of the cyclic phosphodiesterase Rv0805[27] from *Mycobacterium tuberculosis* and differs only in one amino acid substitution (His10 → Tyr) from PAPs.[14,28] The exact *in vivo*

[a] *Department of Chemistry, University of Miami, Coral Gables, FL 33146, USA. E-mail: rpr@miami.edu; Fax: +1 305-284-4571; Tel: +1 305-284-9372*
[b] *School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, QLD4072, Australia*

† Electronic supplementary information (ESI) available: Fig. S1–S4 and Tables S1–S3. See DOI: 10.1039/c8cp07031h

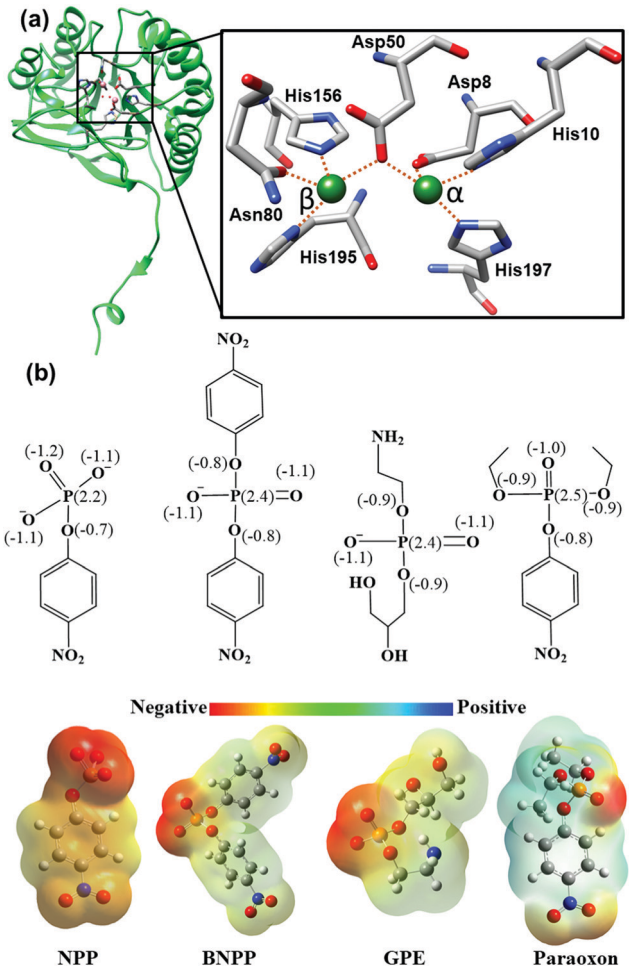

Fig. 1 (a) Structure of the binuclear active site of GpdQ, (b) chemical structures of the NPP, BNPP, GPE, and paraoxon substrates including their electrostatic surface potentials.

metal ion composition of GpdQ is not known.[25] However, several divalent metal ions like Zn(II), Cd(II), Mn(II), and Co(II) have been shown *in vitro* to regenerate its catalytic activity.[1,3,23,29–31] These metal ion compositions include both homonuclear (*e.g.* Zn–Zn, Mn–Mn, and Co–Co) and heteronuclear (*e.g.* Fe–Zn, and Fe–Co) variants. Among these metals, Fe(II) was proposed to be the preferred metal ion in the more buried α site based on anomalous scattering experiments.[25] Importantly, the catalytically active binuclear form (Fig. 1a) of this enzyme is created only upon binding of a substrate (or a substrate mimic) to the inactive monometallic resting state.[4,23,32] In a relatively fast process (rate = $\sim$40 s^{-1}), the bimetallic form is created upon the initial binding of the substrate in the vicinity of the β site. However, it takes approximately another minute for the enzyme to reach its optimal catalytic efficiency.[4] The metal ion binding site (*i.e.* the primary coordination sphere) and the substrate binding pocket (*i.e.* the outer coordination sphere) are intricately connected *via* an extensive hydrogen bonding network.[4,23,31] In this network, the first coordination shell residue Asn80 (Fig. 1a) plays a pivotal role. Upon the initial interaction with a substrate, Asn80 assists metal binding in the β site. However, this residue is hydrogen-bonded to Ser127, which, together with His81, His217, and Tyr19, lines the outer sphere substrate binding pocket in the active site of GpdQ. Upon a rearrangement of the bound substrate in this pocket (including a direct coordination to the metal in the β site), the bond between Asn80 and the β metal ion is broken. This process takes approximately one minute, thus coinciding with the time it takes the enzyme to attain optimal catalytic efficiency. Consequently, the coordination flexibility of Asn80 has been proposed as a regulatory mechanism for GpdQ activity.[4,23,31] The observed dynamics and plasticity of the active site, termed here "breathing of the active site cleft", may also enable GpdQ to utilize a vast range of substrates with distinct chemical properties and sizes.[29] This substrate-promoted coordination flexibility has also been reported for other metalloenzymes including soybean lipoxygenase[33,34] and *E. coli* DNA polymerase I enzyme.[11,35]

Despite a wealth of functional data, structures of GpdQ bound to chemically distinct substrates (*e.g.* NPP, BNPP, GPE, and paraoxon) are currently not yet available. Thus, the precise binding mode(s) of such substrates, and how their interactions in the active site trigger the observed coordination flexibility of Asn80 are also unknown. Furthermore, the exact roles of the metal ions and the remaining first and second coordination shell residues in binding of the substrates are also elusive. In the current study, all-atom molecular dynamics (MD) simulations have been employed to address these unresolved issues. The available experimental information has been fully integrated in these simulations. Our results provide a deeper understanding of the complex mechanism of not only GpdQ but of other related binuclear metallohydrolases as well, and may form the basis for the engineering of optimized enzyme variants (mutants) for applications in bioremediation and other relevant processes in biotechnology.

II. Computational details.

The structure of the monomeric form of GpdQ was obtained from the hexameric 1.9 Å resolution X-ray structure (PDB ID: 3D03).[23] 4-Nitrophenyl phosphate (NPP), bis(4-nitrophenyl) phosphate (BNPP), glycerophosphoethanolamine (GPE) and diethyl 4-nitrophenyl phosphate (paraoxon) were used as substrates. Their charges and electrostatic surface potentials (ESP) were computed at the B3LYP[36]/6-31G(d)[37] level using the Gaussian 09 program.[38] In the X-ray structure, both Co^{2+} ions were substituted with Zn^{2+}. Although these divalent metal ions are chemically similar, more accurate AMBER force field parameters for Zn^{2+} are available due to its redox stable +2 oxidation state. Furthermore, the metal ions are not directly associated with the issues addressed in this study. Antechamber,[39] an in-built tool in the AMBER program, was used to parameterize all four substrates (NPP, BNPP, GPE, and paraoxon). AutoDock Vina 1.5.6. software was used to explore the binding poses of the substrates to the GpdQ active site.[40] The following two docking protocols were used: (1) rigid docking and (2) flexible docking. In the rigid docking, the structure of the enzyme was kept fixed, while the

substrate was flexible. In the flexible docking, both enzyme and substrate were allowed to alter their conformations. Both docking procedures yielded 20 poses each with the exhaustiveness value of 20 for all four enzyme–substrate complexes. The following three separate substrate binding motifs were used for the docking: (1) singly coordinated substrate to the α metal, (2) singly coordinated substrate to the β metal, and (3) doubly coordinated (bridging) substrate to both metals. The subsequent short MD simulations on all these conformations provided a common binding motif with the bridging substrate for NPP, BNPP, GPE and the α metal binding mode for paraoxon. These motifs were used as the starting conformation to study enzyme–substrate interactions through all-atom 100 ns MD simulations. The MD simulations were performed using the GROMACS-4.5.6 [41] program utilizing the AMBER03[42] force field. This force field was reported to reproduce the key structural features of GpdQ in our previous study.[32] In these simulations, the structure of the binuclear Zn–Zn center was maintained by applying distance restraints of 1.7–2.2 Å between the Zn ion and δ-nitrogen atom of His10, His156, His195, and His197, carboxylate oxygen atom of Asp8 and Asp50, and carbonyl oxygen atom of Asn80.[43] The accuracy of the AMBER force field was further tested by performing all-atom 100 ns MD simulations on all four enzyme–substrate complexes without constraints using the AMOEBA[44,45] polarizable force field utilizing the Tinker software package.[46–48] The clustered structures of the active site including both metal ions, all first coordination shell residues (Asp8, His10, Asp50, Asn80, His195, and His197) and the substrate obtained using the AMBER and AMOEBA force fields were quite comparable for all four substrates. The computed RMSD values between these structures were quite low *i.e.* 0.54, 1.47, 1.19, and 1.27 Å for **GpdQ–NPP**, **GpdQ–BNPP**, **GpdQ–GPE**, and **GpdQ–paraoxon**, respectively (Fig. S4 in the ESI†). Furthermore, differences in all metal–ligand distances were within the 0.01–0.27 Å range (Table S1 in the ESI†). For all MD simulations, the structures were contained in a cubic box with dimensions of $10 \times 10 \times 10$ nm^3. The shortest distance from the surface of the protein to the edge of the box was 1.0 nm. Electrostatic interactions were calculated using the Particle Mesh Ewald method,[49] and a cutoff at 1.2 nm was used for both van der Waals and Coulombic interactions. The coordinates of initial structures and their force field parameters are provided in the ESI† (Tables S2 and S3 in ESI†). The box was filled with TIP3P water molecules.[50] To neutralize the system and to simulate a physiological ion concentration of 154 mM, some of the water molecules were replaced by sodium and chloride ions. Energy minimization of the starting structure was performed for 3000 steps by a steepest descent method, which resulted in the formation of the starting structure for MD simulations. These simulations were performed in two steps. In the first step, the active site of GpdQ was restrained and MD simulations was performed for 30 ns. In the second step, restraints from the active site were removed and simulations was performed for 100 ns. Furthermore, to study the effect of sampling time, we extended the simulation of the **GpdQ–NPP** complex to 200 ns and did not observe any noticeable changes in the overall structure *i.e.* RMSD of 1.28 Å for the whole protein. The MD simulations were carried out with a constant number of particles (N), pressure (P) and temperature (T) (*NPT* ensembles). To constrain the bond lengths and angles of the water molecules the SETTLE[51] algorithm was used and to constrain the remaining bond lengths the LINCS[52] algorithm was employed. The trajectories were computed for each model with a time step of 2 fs. Cluster analysis was performed to derive the most representative structure of GpdQ. The binding free energies between GpdQ and the four selected substrates were calculated using the lambda (λ) particle approach. In this methodology, a thermodynamic cycle defines the bound and unbound states. In this cycle, the relative binding energies of two ligands can be computed as the difference in free energy associated with the chemical changes in their bound and solvated states.[53] This approach has been discussed in detail previously.[53–55] In these calculations, Coulombic and van der Waals interactions between GpdQ and the substrates were turned off in a systematic way, *i.e.* first the Coulombic and then van der Waals terms, thus avoiding interactions of oppositely-charged atoms at very close distances that would have provided unfavorable configurations and inaccurate energies. Metal binding energies in the presence of each substrate were calculated by utilizing the Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) method.[56] Porcupine plots were utilized to explore modes of protein motion.[57] Maestro software was used to create the 2D interaction diagram between the enzyme and the substrate.[58]

Principal component analysis of the Cα atoms was also performed to study conformational dynamics of all four enzyme–substrate complexes.[59] The distribution of eigenvalues in this plot represents the conformational changes in these structures. A covariance matrix was generated from each MD trajectory to describe the correlation between all pairs of backbone Cα atoms of the protein. In the next step, diagonalization of the covariance matrix of the backbone Cα atoms was performed to obtain the associated eigenvectors and eigenvalues. The first two eigenvectors were found to account for more than 70% of the collective motion of the protein. They can be used to determine modes of fluctuations that contribute significantly to the overall motion of the protein. The eigenvector with the highest eigenvalue was considered as the first principal component (PC1, x-axis), whereas the eigenvector with the second highest eigenvalue as the second principal component (PC2, y-axis). The direction of motion is represented by the eigenvectors and the amount of motion along with the eigenvectors is represented by eigenvalues. VMD,[60] Yasara,[61] ChemDraw and Chimera[62] programs were used for the visualization of the MD trajectories and preparation of the figures used in this study.

III. Results and discussions

In this study, interactions of the catalytically active binuclear form of GpdQ with four chemically diverse substrates, *i.e.* NPP (phosphomonoester), BNPP and GPE (both phosphodiesters) and paraoxon (phosphotriester) have been investigated using all-atom MD simulations. The chemical structures, atomic

charges and electrostatic potentials of these substrates are substantially different from each other (Fig. 1b). The phosphate oxygen atoms of NPP contain the highest electronegative charge, followed by GPE, BNPP and paraoxon. The analysis of the root-mean-square-deviations (RMSDs) confirmed the equilibration of all four complexes (**GpdQ–NPP**, **GpdQ–BNPP**, **GpdQ–GPE**, and **GpdQ–paraoxon**) within the simulation time (Fig. S1 in ESI†). The preferential binding of these substrates and coordination flexibility of the enzyme have been discussed by comparing their ESPs, secondary structures, non-covalent interactions, radius of gyration (R_g), root-mean-square-fluctuations (RMSF), principal components (PC), substrate binding free energies, and metal binding energies (see below).

IIIa. GpdQ–NPP interactions

NPP is a phosphomonoester with one *p*-nitrophenyl group attached to the phosphorus atom (Fig. 1b). In the **GpdQ–NPP** complex, the protein surrounding the active site can be divided into six loops (I–VI); *i.e.* loop I (E16-D23), loop II (V52-R56), loop III (P78-D83), loop IV (S122-G129), loop V (G163-A171) and loop VI (Y221-P228). The ESP map of this complex shows that loops I, II, III and VI are negatively charged, loop V is slightly positively charged, and loop IV is neutral (Fig. 2a). A 2D graph of the residues located within 5 Å of the binuclear metal center also shows several negatively charged residues adjacent to the substrate, and the *p*-nitrophenyl group of NPP is exposed to the solvent water (Fig. 2b). The binding of this negatively charged substrate in such a microenvironment is likely to be driven by its interaction with the positively charged α and ß metal ions. The most representative structure of the active site derived from the **GpdQ–NPP** simulation is shown in Fig. 2c. In this structure, Asp8 (1.82 Å), His10 (1.92 Å), and His197 (2.02 Å) are coordinated to the α site, whereas Asp50 (1.81 Å), Asn80 (2.08 Å), His156 (2.11 Å), and His195 (2.46 Å) to the β site (Table 1). The hydroxyl group is located on the opposite side of the substrate and bridges both metals. The NPP substrate thus binds to the binuclear active site in a μ-1,3 bidentate fashion. In an asymmetric manner, one phosphate oxygen atom of NPP binds to the α metal (M_α–O = 2.06 Å; Table 1) and the second oxygen atom coordinates more strongly to the β metal (M_β–O = 1.88 Å). The coordination numbers of the α and β metals in this enzyme–substrate complex are five and six, respectively, contrasting the six- and five-coordinate α and β metals observed in the crystal structure of the phosphate-bound complex.[23] A superposition of the equilibrated and X-ray structures shows the displacement of key active site residues upon the binding of NPP (Fig. 2d). The use of the actual NPP substrate in MD simulations (*vs.* a PO_4 analogue as in the X-ray structure) substantially alters the active site of the enzyme. The radius of gyration (R_g) of 18.25 Å for the **GpdQ–NPP** complex is also smaller than the corresponding value (18.77 Å) computed for the X-ray structure (Fig. S2 in ESI†).

Rapid kinetics measurements suggest that after substrate binding the active site of GpdQ undergoes a swift structural rearrangement, facilitated by the hydrogen bonding network involving several first and second coordination shell residues.[4] Similarly, in the MD simulations multiple second coordination

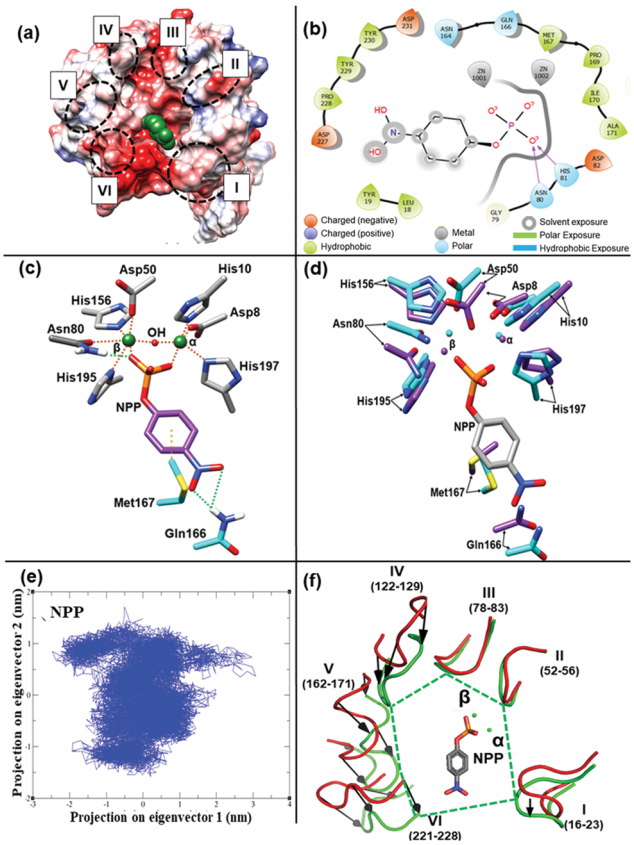

Fig. 2 **GpdQ–NPP** complex: (a) electrostatic surface potential; (b) 2D interaction graph of the **GpdQ–NPP** complex; (c) equilibrated structure of the active site; (d) superposition of the equilibrated (purple) and X-ray (cyan carbons) structures; (e) principle component analysis; and (f) superposition of the six loops from the equilibrated (green) and X-ray (red) structures (the direction and length of black arrows describe the direction and magnitude of motion).

shell residues are reoriented upon substrate binding, thus stabilizing the **GpdQ–NPP** complex either through direct or water-mediated non-covalent interactions (Table 2). For instance, in the X-ray structure Asn80 forms a hydrogen bond with Ser127 (loop IV), a second sphere residue lining the substrate binding pocket. The disruption of this hydrogen bond through mutagenesis (Ser127 was replaced by an alanine) led to a mutant with catalytic properties similar to those of the wild-type enzyme but with an enhanced affinity for the metal in the β site.[18] In the equilibrated enzyme–NPP complex Asn80 forms a direct hydrogen bond (2.26 Å) with the non-esterified phosphate oxygen of the substrate, indicating that the addition of a substrate indeed leads to a rearrangement of the hydrogen bonding interactions of Asn80 (Table 2). This rearrangement was proposed to promote a change in the coordination environment of the β metal, leading to a gradual increase in catalytic activity and a concomitant disruption of the coordination bond between Asn80 and the β metal (see also below).[4,23,31] The β metal site of the **GpdQ–NPP** complex exhibited lower flexibility than the GpdQ–phosphate complex possibly because the larger NPP substrate restricts the structural flexibility required to alter

Table 1 Metal coordination number and metal–metal, metal–substrate, metal–OH, and metal–ligand distances (Å) for the first coordination shell residues. Metal coordination sites (α and β) are shown in the parentheses

	GpdQ–NPP	GpdQ–BNPP	GpdQ–GPE	GpdQ–paraoxon
Coordination number	5 (α), 6 (β)	5 (α), 6 (β)	5 (α), 6 (β)	5 (α), 6 (β)
M_α–M_β	3.15	3.14	3.11	3.17
Substrate	2.06 (α), 1.88 (β)	2.06 (α), 1.87 (β)	1.71 (α), 1.83 (β)	2.16 (α)
μ-OH	1.66 (α), 1.66 (β)	1.69 (α), 1.68 (β)	1.70 (α), 1.74 (β)	1.67 (α), 1.70 (β)
Asp8	1.82 (α)	1.83 (α)	1.71 (α)	1.81 (α), 1.90 (β)
His10	1.92 (α)	1.94 (α)	1.99 (α)	2.01 (α)
His197	2.02 (α)	2.02 (α)	2.58 (α)	2.07 (α)
Asp50	1.81 (β)	1.89 (β)	1.77 (β)	1.72 (β)
Asn80	2.08 (β)	2.01 (β)	2.18 (β)	2.00 (β)
His156	2.11 (β)	2.23 (β)	2.14 (β)	2.14 (β)
His195	2.46 (β)	2.57 (β)	2.48 (β)	3.01 (β)

Table 2 Non-covalent interaction (in Å) between the substrate and 1st and 2nd coordination shell (*) residues. The nature of interaction is provided in the parentheses

	GpdQ–NPP	GpdQ–BNPP	GpdQ–GPE	GpdQ–paraoxon
His10	—	—	—	3.36 (π–π)
Asn80	2.26 (H-bond)	2.53 (H-bond)	2.28 (H-bond)	—
His195	—	—	2.75, 3.18 (H-bond)	—
Arg12*	—	—	—	3.30, 3.35 (H-bond)
His81*	—	2.87 (CH–π)	—	—
Gln166*	1.93, 2.69 (H-bond)	2.18, 2.24 (H-bond)	—	—
Met167*	2.70 (CH–π)	2.71 (CH–π)	—	—
Ile170*	—	3.20 (CH–π)	—	—
Asn196*	—	—	2.26, 2.75 (H-bond)	—
His217*	—	—	—	3.37 (CH–π)
Thr226*	—	—	—	3.34 (CH–π)

the coordination environment of that site. The interactions of Gln166 (loop V) *via* two hydrogen bonds (1.93 and 2.69 Å) and Met167 (loop V) through a CH–π interaction (2.70 Å) with the substrate may contribute to a restricted structural flexibility of the β site when compared to the GpdQ–phosphate complex (Fig. 2c and Table 2). In the latter, these residues are positioned quite far from the nearest oxygen atom of phosphate (*i.e.* 9.0 Å and 5.0 Å for Gln166 and Met167, respectively) and thus have no effect on the binding of this substrate mimic. Similar to Ser127, the second shell residues His81 (loop III) and His217 (loop VI) have also been shown to influence the affinity of the metal in the β site without directly affecting the catalytic parameters of the enzyme.[4] In the **GpdQ–NPP** complex, these residues indirectly interact with the substrate through water-mediated interactions, *i.e.* His81 *via* two waters and His217 through three waters. Additionally, two other second coordination shell residues, Tyr19 and Asn53, associate with NPP using one and three waters, respectively.

Although the MD model with bound NPP does not reflect the coordination flexibility proposed for the β site of GpdQ upon binding of phosphate, the model does nonetheless illustrate changes in the coordination environment triggered upon binding of a substrate. Instead of the β site the α site displays flexibility, mediated *via* residue Asp50 (Fig. 2c). In particular, the observed five- and six-coordinate environments of the two metals are also in agreement with spectroscopic data that demonstrate that upon substrate binding the coordination numbers do not change but different residues may display flexibility.[31] The inherent "breathing of the active site cleft" was further investigated using the following three parameters: principal component analysis (PCA), movements of loops and RMSF. Both principal components (PC1 and PC2 in Fig. 2e) of the Cα atoms were observed to contribute equally on the basis of PCA. In the **GpdQ–NPP** complex, the ensemble of conformations is restricted to a single energy basin. Kinetic measurements have revealed that after the initial formation of the catalytically active binuclear form of GpdQ it takes approximately another minute before optimal catalytic efficiency is reached.[4] This lag period may be associated with the reorganization of the active site involving residues such as Ser127, His81, and His217 (see above) and movement of the six loops containing these residues (shown using arrows in Fig. 2f) to promote the binding of the substrate in the active site. Among those loops I, V, and VI undergo major fluctuations, while loops II, III, and IV are more rigid. The RMSF indicates dominant fluctuations in loops I, V, and VI (Fig. 3). The relative movement of these loops can be analyzed quantitatively by computing the area of an irregular hexagon formed by the six residues located at their tips (*i.e.* Tyr19, Cys54, His81, Thr126, Ile170, and Asp227; Fig. 2f). The binding of NPP expands the active site and increases the area (268.3 $Å^2$; Table 3) of the hexagon by 36.5 $Å^2$ in comparison to the X-ray structure (231.8 $Å^2$).

The free energy of NPP binding to GpdQ calculated using λ dynamics is −24.3 kcal mol^{-1} (Table 3). This energy is dominated by interactions of the phosphate oxygens of NPP with the two divalent metals (α and β). The metal binding energies in the presence of this substrate computed using the MM-PBSA

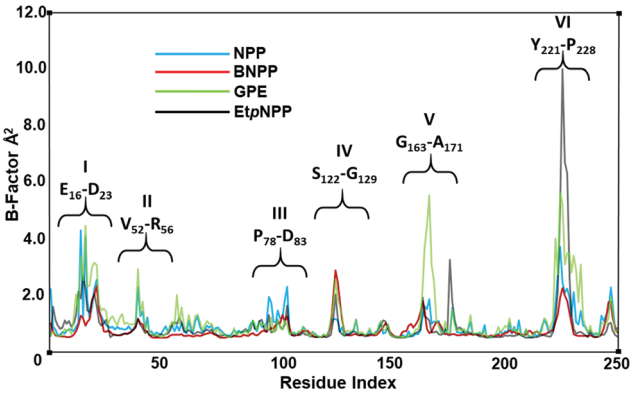

Fig. 3 Root-mean-square-fluctuations (RMSF) of the **GpdQ–NPP**, **GpdQ–BNPP**, **GpdQ–GPE**, and **GpdQ–paraoxon** complexes.

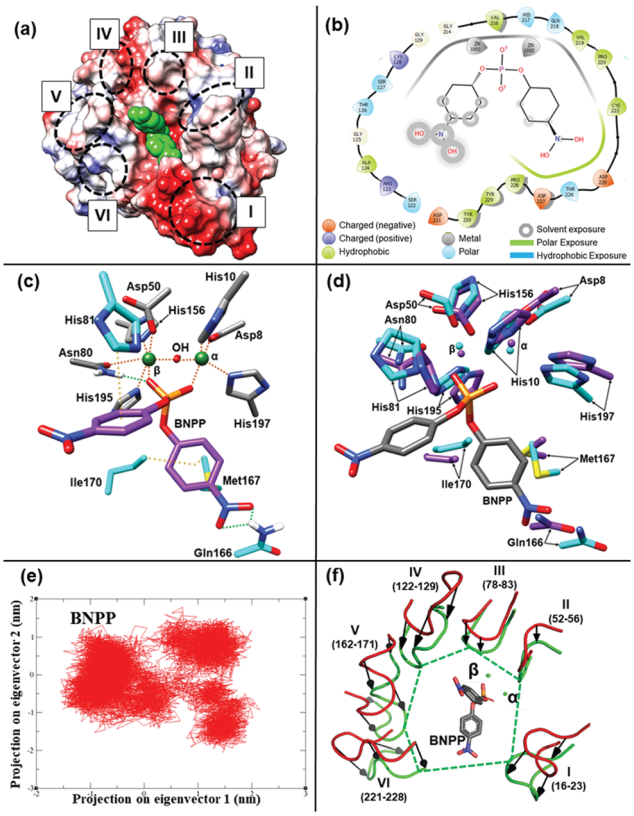

Fig. 4 **GpdQ–BNPP** complex: (a) electrostatic surface potential; (b) 2D interaction graph of the **GpdQ–BNPP** complex; (c) equilibrated structure of the active site; (d) superposition of the equilibrated (purple) and X-ray (cyan carbons) structures; (e) principle component analysis; and (f) superposition of the six loops from the equilibrated (green) and X-ray (red) structures (the direction and length of black arrows describe the direction and magnitude of motion).

approach[63] show that the binding energy (−209.2 kcal mol^{-1}) of the β site is 7.6 kcal mol^{-1} lower than the corresponding energy (−201.6 kcal mol^{-1}) of the α site (Table 3). The lower energy for the former is supported by the shorter bond distance (1.88 Å *vs.* 2.06 Å) and higher coordination number (6 *vs.* 5) for this site (Table 1).

IIIa1. GpdQ–BNPP interactions. BNPP is a phosphodiester that contains two *p*-nitrophenyl groups in comparison to only one in NPP (Fig. 1b). This introduces significant changes to its structure, charges, and mode of binding. The ESP map of the **GpdQ–BNPP** complex is different from the corresponding map of the NPP complex. Here, loop I is dominated by negative charge, loop VI contains a slightly positive charge, while loops II, III, IV, and V are largely neutral (Fig. 4a). The overall 2D graph of the **GpdQ–BNPP** complex shows multiple negatively charged residues around the substrate (Fig. 4b). However, there are some noticeable differences in these maps generated for the NPP and BNPP substrates. More positively charged residues are located near BNPP. Additionally, while the nitro group of NPP is exposed to solvent, one of the two nitro groups of BNPP is surrounded by hydrophobic residues (Fig. 4b). While one half of BNPP is exposed to solvent, the other half interacts with the side chains of His81, Met167, and Ile170 through CH–π interactions. The total number of hydrogen bonds formed by NPP and BNPP are the same (3), but the number of CH–π interactions are different (one for NPP and three for BNPP). The most representative structure of the active site of the **GpdQ–BNPP** complex is shown in Fig. 4c. In this complex, the direct ligand environment of both metals [α (Asp8, His10, and His197) and β (Asp50, Asn80, His156, and His195)] and their coordination number (five and six for the α and β metal, respectively) are the same as in the **GpdQ–NPP** complex. The most notable difference is observed in the metal–ligand distances in the β site (Table 1). While the metal–Asp80 distance shrinks by 0.08 Å in the BNPP complex, the remaining three coordination bonds with Asp50, His156, and His195 are longer by 0.08, 0.12, and 0.11 Å, respectively (Fig. 4c and Table 1). The phosphate group of BNPP also forms a μ-1,3 bidentate bridge and the metal–metal distance is virtually identical to that of the NPP complex. The superposition of the BNPP complex with the X-ray structure of GpdQ again reveals some significant deviations in the location of several key second coordination shell residues such as His81,

Table 3 The area of hexagon (Å^2), substrate binding free energy (in kcal mol^{-1}) and metal (α and β) binding energies (in kcal mol^{-1}) in the presence of each substrate

Complexes	Area of hexagon (Å^2)	Substrate binding free energy (kcal mol^{-1})	Metal binding energy (kcal mol^{-1})	
			α metal	β metal
GpdQ–NPP	268.3	−24.3	−201.6	−209.2
GpdQ–BNPP	288.9	−22.4	−182.0	−192.4
GpdQ–GPE	350.8	−23.4	−193.2	−200.3
GpdQ–paraoxon	362.5	−14.1	−163.8	−168.0

Gln166, Ile170, and His197 (Fig. 4d). The RMSD of 1.66 Å between the **GpdQ–BNPP** and **GpdQ–NPP** models indicates some noticeable structural difference between these two complexes, largely due to different conformations in the substrate binding pocket. The R_g value (18.5 Å) of the **GpdQ–BNPP** complex is slightly (0.25 Å) higher than the corresponding value for its NPP counterpart due to the binding of a bulkier BNPP substrate (Fig. S2 in ESI†).

Similar to **GpdQ–NPP**, Asn80 (ligand of the β metal) forms a hydrogen bond (2.53 Å) with an oxygen atom of the phosphate moiety, while Gln166 and Met167 stabilize one of the two nitro groups *via* a bifurcated hydrogen bond and a CH–π interaction, respectively (Fig. 4c and Table 2). Furthermore, His81 and Ile70 move and rotate towards the aromatic phenyl ring of BNPP to associate with it through a hydrogen bond and a CH–π interaction, respectively. Other second coordination shell residues (Tyr19 (loop I), Asn53 (loop II), and His217 (loop VI)) also move 3–8 Å from their corresponding locations in the X-ray structure towards the substrate.

With respect to the flexibility of the first coordination sphere of GpdQ, PCA shows a much broader ensemble (three energy basins) of conformations for the **GpdQ–BNPP** complex than in the NPP complex (one basin; Fig. 2e and 4e). Due to the larger size of BNPP, major fluctuations occur in all six loops of the enzyme, while only three of them (I, V, and VI) are significantly affected in the NPP complex (Fig. 3). Dominant fluctuations are observed in loops III (*i.e.* Asn80 and His81), IV (no direct interactions with BNPP) and V (*i.e.* Gln166, Met167, and Ile170). In comparison to the **GpdQ–NPP** complex, the binding of BNPP increases the area of the irregular hexagon by 20.6 $Å^2$ to 288.9 $Å^2$ for the **GpdQ–BNPP** complex (Table 3). This increase is caused by the expansion of the active site to accommodate the additional *para*-nitrophenyl group of BNPP. The RMSF in all six loops for the binding of BNPP is the smallest when compared to all the other substrates (Fig. 4f). The binding free energy (−22.4 kcal mol^{-1}) for BNPP is computed to be only slightly (1.9 kcal mol^{-1}) lower than the corresponding energy computed for NPP (Table 3). The smaller negative charge of the phosphate group of the BNPP substrate would be expected to lead to a weaker interaction with the metals in the active site. Indeed, the metal binding energies (using MM-PBSA technique) in the presence of BNPP are weaker by nearly 20 kcal mol^{-1} when compared to the NPP complex (Table 3). Thus, since difference in binding affinity between BNPP and NPP is relatively small, the stronger contribution of the metal-phosphate in the NPP complex is somewhat compensated by stronger interactions between the nitro groups and the second coordination sphere in the BNPP complex.

IIIa2. GpdQ–GPE interactions. Glycerophosphoethanolamine (GPE), a natural substrate of GpdQ, is also a phosphodiester with a glycerolate and aminoethanolate group (Fig. 1b). Due to the presence of these chemically distinct groups and absence of nitrophenyl group in GPE, the bonding interactions of GPE are anticipated to differ significantly from those observed for NPP and BNPP. For instance, GPE interacts predominantly through hydrogen bonding with GpdQ; no π–π or CH–π interactions are

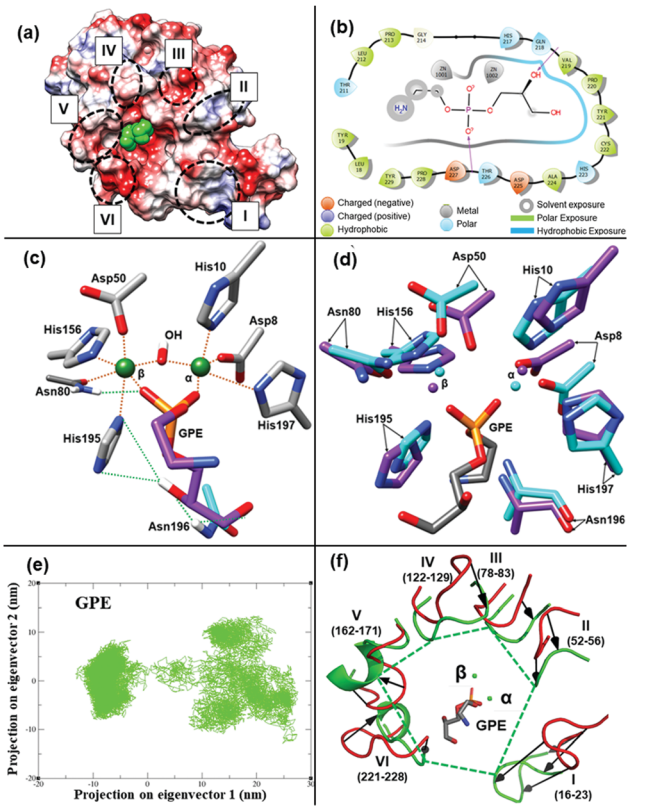

Fig. 5 **GpdQ–GPE** complex: (a) electrostatic surface potential; (b) 2D interaction graph of the **GpdQ–GPE** complex; (c) equilibrated structure of the active site; (d) superposition of the equilibrated (purple) and X-ray (cyan carbons) structures; (e) principle component analysis; and (f) superposition of the six loops from the equilibrated (green) and X-ray (red) structures (the direction and length of black arrows describe the direction and magnitude of motion).

observed (see also Fig. S3 in ESI†). Consequently, substantial differences in the ESP maps of the BNPP and GPE complexes are observed (Fig. 5a), and the R_g value for the **GpdQ–GPE** complex is also higher by 0.25 Å than the one for the BNPP complex (Fig. S2 in ESI†). In contrast to **GpdQ–BNPP**, loops III, V, and VI acquire predominantly negative charge, loop II becomes positive, and loops I and IV are neutral in the **GpdQ–GPE** complex (Fig. 5a). These variations are also reflected in the 2D graphs for the two complexes (Fig. 4b and 5b). The glycerolate moiety of GPE is surrounded by polar residues, while the remaining part of the substrate is exposed to solvent water. This binding mode of GPE has also been suggested from an analysis of the crystal structure of GpdQ.[29] The most representative structure of the **GpdQ–GPE** complex is shown in Fig. 5c. Among the four substrates compared in this study, GPE forms the most hydrogen bonds (5) with the enzyme (Table 2). Among the first coordination shell residues Asn80 again performs an important dual role in metal and substrate binding. Another ligand to the β metal, His195, forms two hydrogen bonds with the substrate. The remaining two hydrogen bonding interactions originate from the second coordination sphere residue Asn196 (Table 2). The other second coordination shell residues (Tyr19, His81, Asn53, and His217) also assist the

binding of GPE through water-mediated interactions. While the overall arrangement of the first coordination shell residues in the **GpdQ–GPE** complex is similar to that observed in the BNPP complex, the metal–ligand distances differ significantly (Fig. 5c and Table 1). In the former, the bond distances of Asp8, Asp50, His156, and His195 with the corresponding metal are shorter by 0.09–0.12 Å, while those of His10, Asn80 and His197 are longer by 0.05–0.56 Å (Table 1). These significant changes in coordination bond lengths are also reflected in the superposition of the equilibrated **GpdQ–GPE** complex onto the X-ray structure of GpdQ that indicates large deviations in the positions of the two aspartate residues in the first coordination sphere, *i.e.* Asp8 and Asp50 (Fig. 5d). Despite these variations in bond lengths, the phosphate group of GPE adopts the same μ-1,3 bidentate binding mode as observed in the NPP and BNPP complexes, and also maintains the coordination numbers for the α and β metal (5 and 6, respectively) and a similar metal–metal distance (Table 1). However, the phosphate group of GPE coordinates significantly tighter to the two metal ions than its counterpart in NPP and BNPP, in particular in the α site (bond contraction by over 0.3 Å).

The conformational variations induced upon GPE binding when compared to NPP or BNPP binding are also evident in a comparison of their respective PCAs. The PCA of the **GpdQ–GPE** complex shows a scattered conformation with a large surface area (Fig. 5e). The ensemble of conformations with two energy basins of this complex is broader than in the BNPP complex, suggesting that the presence of GPE expands the population of conformational ensembles. Major fluctuations occur in each of the six loops in this complex, with magnitudes that are higher than those in the BNPP complex (Fig. 3 and 5f). Loops I, II, and VI display the largest fluctuations. The surface area of the hexagon in the GPE complex is 61.9 $Å^2$ greater than that of the BNPP complex (Table 3). The binding free energy for GPE (-23.4 kcal mol^{-1}) lies in the middle between the corresponding values of the NPP and BNPP complexes, and the same is observed for the binding energies to both metals (Table 3). Thus, despite considerably shorter bonds between the metals and the phosphate group of GPE, no obvious trend that links metal coordination to substrate binding energy is observed. While this indicates that the binding energies of different substrates to GpdQ are based on a series of complex interactions in both the primary and secondary coordination spheres, it also illustrates how flexible this enzyme is with respect to accommodating vastly different reactants.

IIIb. GpdQ–paraoxon interactions

The paraoxon substrate is a phosphotriester with one nitrophenol group (like NPP) and two ethyl groups and has the highest positive charge on the phosphorous atom (Fig. 1b). The ESP map of **GpdQ–paraoxon**, in particular in the ethyl group binding pocket, is significantly different from the corresponding maps of the other three substrates (Fig. 6a). In the paraoxon complex, loops I, III, V, and VI are negatively charged, and loops II and IV are positive (Fig. 6a). These differences are also reflected in (i) the 2D graph of the **GpdQ–paraoxon** complex when compared to the 2D graphs of the other substrates (Fig. 6b), and (ii) the relatively large RMSD deviations obtained from a pairwise structural comparison of **GpdQ–paraoxon** with the NPP (3.12 Å), BNPP (3.23 Å) and GPE (3.00 Å) complexes, respectively. For instance, only paraoxon interacts with residues His10, Arg12, His217, and Thr226 in the active site. Specifically, this substrate forms a π–π interaction with His10 and CH–π interactions with both His217 and Thr226, while Arg12 forms two weak hydrogen bonds with the nitro oxygens of paraoxon (Fig. 6c and Table 2). However, unlike all other substrates, paraoxon does not interact with Asn80 and only binds to the α metal. Similar to the complexes with the other substrates Tyr19, Asn53, His81, and His217 interact with paraoxon through water-mediated interactions.

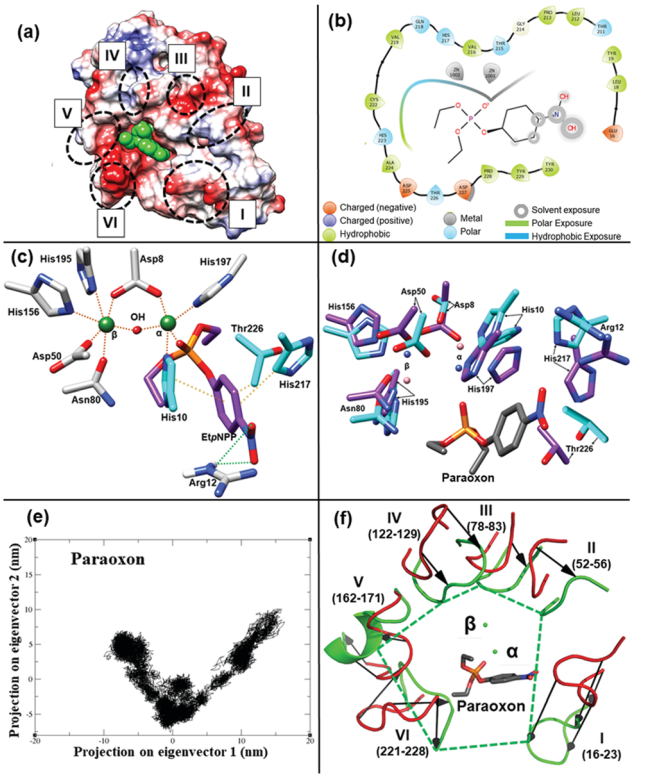

Fig. 6 **GpdQ–paraoxon** complex: (a) electrostatic surface potential; (b) 2D interaction graph of the **GpdQ–paraoxon** complex; (c) equilibrated structure of the active site; (d) superposition of the equilibrated (purple) and X-ray (cyan carbons) structures; (e) principle component analysis; and (f) superposition of the six loops from the equilibrated (green) and X-ray (red) structures (the direction and length of black arrows describe the direction and magnitude of motion).

In the substrate complexes with NPP, BNPP, and GPE the primary coordination spheres did not differ significantly (apart from some variations in bond lengths). The phosphate groups coordinate in a μ-1,3 mode to the two metal ions in the active site, and the only other metal-bridging ligand is a OH^-/H_2O. Residue Asp50, shown to be located in a metal-bridging position in the crystal structure of GpdQ (Fig. 1a), is moved towards a terminal position in the coordination sphere of the β metal in these complexes (Fig. 2c, 4c and 5c). In contrast, in the **GpdQ–paraoxon** complex, residue Asp8 bridges both metals in a μ-1,3 mode while the substrate only interacts

directly with the α metal with a bond length that is longer than in the other substrate complexes (Fig. 6c and Table 1). Studies with phosphotriesterases, enzymes that prefer triesters such as paraoxon as substrates, support a similar binding mode for this substrate in their active sites.[20,64,65] Additionally, QM/MM calculations with these enzymes also support the same binding mode.[66] Although the rearrangement of Asp8 may lead to a significant elongation of the metal-His195 bond in the β site (Table 1), the coordination numbers of the α and β metals are also 5 and 6, respectively, as in the other GpdQ–substrate complexes (Fig. 6c). The R_g value of 18.75 Å of the **GpdQ–paraoxon** complex is similar to GPE complex, but higher than the corresponding values for the NPP and BNPP complexes (Fig. S2 in ESI†).

The PCA of the **GpdQ–paraoxon** complex provides conformations that are distinct from all other substrate complexes (Fig. 6e). In this V-shaped graph, there are many well spread and narrow basins, which correspond to multiple conformational structures. Fluctuations occur in all loops upon the binding of paraoxon (Fig. 6f). In particular, loops I and VI move downwards, whereas loops II and III shift upwards. Common to all loop motions is that they move away from the substrate to accommodate this bulky molecule in the active site. Consequently, the area of the irregular hexagon is the greatest (362.5 $Å^2$) when paraoxon is bound (Table 3). Similarly, the RMSF of the loops, especially loops V and VI, are greatest in this complex (Fig. 3 and 6d). The binding free energy for paraoxon (−14.1 kcal mol^{-1}) is smaller than that for the other substrates, as are the metal binding energies (163.8 and 168.0 kcal mol^{-1} for the α and β metal, respectively), Table 3. Thus, despite the presence of a μ-1,3 metal-bridging aspartate ligand (Asp8) the loss of the bidentate substrate coordination leads to a significant loss of affinity for both the substrate and the metals.

IV. Summary and conclusions

In this MD study, interactions of the catalytically active binuclear form of GpdQ with four chemically distinct substrates, *i.e.* NPP (a phosphomonoester), BNPP and GPE (both phosphodiesters), and paraoxon (a phosphotriester) have been investigated. In particular, the roles of metal ions and the first and second coordination shell residues, coordination flexibility and dynamic transformations in preferential binding of these substrates have been elucidated.

The chemical nature of the substrate influences the overall charge of all six loops (I–VI) that surround the active site of GpdQ, and in turn also modifies their electrostatic surface potential (ESP) maps. Substrate binding in the mostly negatively charged active site is driven by interactions with the positively charged metals in the α and β sites. All substrates bind in a μ-1,3 bidentate fashion with the exception of paraoxon that coordinates in a monodentate manner only to the α metal and consequently has the lowest affinity. The metal–metal (M_α–M_β) distances appear to be little affected by the type of substrate (ranging from 3.11 Å to 3.17 Å) but are significantly shorter than in the crystal structure (3.68 Å) with a PO_4 analogue, indicating that substrate binding has a significant effect on the geometry of the active site (Table 1). Several second coordination shell residues stabilize the enzyme–substrate complex through distinct non-covalent (direct or water-mediated) interactions (Table 2). These conformational changes are in agreement with experimental data that demonstrate that modifications (mutations) of Asn80 in the first coordination sphere and various residues in the second sphere have a significant effect on both metal and substrate binding.[4,23,31] Experimental data also suggest a link between substrate binding and coordination flexibility in GpdQ, similar to the non-heme iron lipoxygenase from soybean.[33,34] Specifically, for GpdQ, it has been proposed that substrate binding promotes a gradual transformation in the active site where the bond between Asn80 and the β metal is broken as the enzyme attains a conformation optimal for catalysis.[4,23] Although our MD simulations support a role for Asn80 in both metal and substrate binding, they do not demonstrate coordination flexibility of this residue. However, the simulations do indeed support structural flexibility in the active site, with Asp50 moving from a metal-bridging conformation into a monodentate location on the β metal. The only water (–OH) molecule present in the active site of the simulated models is bridging the two metals. Thus, the resulting coordination numbers of the metals in the resting state are five in each case, providing vacant binding positions for the incoming substrate. Interestingly, in agreement with spectroscopic studies,[4,23,31] the enzyme–substrate complex has a five- and a six-coordinate metal independent of the nature of the substrate. The mono- and diester substrates bind in a μ-1,3 bidentate mode, whereas the triester substrate coordinates only to the α metal (however the metals are still connected *via* a μ-1,3 bridge due to the rearrangement of Asp8; Fig. 6).

The PCA suggests distinct conformational dynamics and coordination flexibility for all four enzyme–substrate complexes, *i.e.* one, three, two, and five energy basins for NPP, BNPP, GPE, and paraoxon, respectively. Additionally, based on the shape and size of the substrate, different loops exhibit dominant fluctuations, *i.e.* NPP (I, V and VI), BNPP (I–VI), GPE (I, II and, VI), and paraoxon (I–VI). The movements of these loops are also supported by the computed RMSF values (Fig. 3). Furthermore, substantial differences in the area of irregular hexagon (268.3, 288.9, 350.8, and 362.5 $Å^2$ for NPP, BNPP, GPE, and paraoxon, respectively) formed by the residues on these loops explicitly indicate flexibility of the active site (Table 3). The substrate binding free energies of most of the substrates are quite similar (22.4–24.3 kcal mol^{-1}) except for paraoxon (14.1 kcal mol^{-1}). The metal binding energy is the highest (201.6 and 209 kcal mol^{-1} for the α and β metal, respectively) in the presence of NPP and the lowest (163.8 and 168.0 kcal mol^{-1} for the α and β metal, respectively) in the paraoxon case (Table 3).

Overall, the simulations demonstrate how GpdQ can accommodate largely diverse substrates through its inherent flexibility around the active site. The simulations also support a mechanistic model whereby substrate binding triggers changes in the primary coordination environment. Instead of Asn80 (as proposed based on experimental data) Asp50 was singled

out as the metal ligand that undergoes the most significant change upon substrate binding, changing from a μ-1,3 mode in the free form to a monodentate, terminal coordination in the substrate-bound state. This discrepancy may be connected to the fact that experimental investigations of the enzyme–substrate interactions employed the substrate mimic and product phosphate. It is likely that this small molecule triggers conformational changes different from those of the significantly larger substrate molecules. Nonetheless, the computational studies described here underline the extensive conformational changes that take place upon substrate binding to GpdQ, thus providing an appropriate framework to study its reaction mechanism and to design variants (mutants) with properties optimal for this enzyme's application in bioremediation.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

The authors declare no competing financial interests. This material is based upon work supported by the grant from the National Science Foundation (Grant Number CHE-1664926) to R. P. Computational resources from the Center for Computational Science (CCS) at the University of Miami to R. P. are also greatly appreciated.

References

1 L. J. Daumann, B. Y. McCarthy, K. S. Hadler, T. P. Murray, L. R. Gahan, J. A. Larrabee, D. L. Ollis and G. Schenk, *Biochim. Biophys. Acta, Proteins Proteomics*, 2013, **1834**, 425–432.
2 J. A. Gerlt and F. H. Westheimer, *J. Am. Chem. Soc.*, 1973, **95**, 8166–8168.
3 J. A. Gerlt and G. J. Whitman, *J. Biol. Chem.*, 1975, **250**, 5053–5058.
4 K. S. Hadler, N. Mitić, F. Ely, G. R. Hanson, L. R. Gahan, J. A. Larrabee, D. L. Ollis and G. Schenk, *J. Am. Chem. Soc.*, 2009, **131**, 11900–11908.
5 G. Schenk, I. Mateen, T.-K. Ng, M. M. Pedroso, N. Mitić, M. Jafelicci, R. F. C. Marques, L. R. Gahan and D. L. Ollis, *Coord. Chem. Rev.*, 2016, **317**, 122–131.
6 E. Ghanem, Y. Li, C. Xu and F. M. Raushel, *Biochemistry*, 2007, **46**, 9032–9040.
7 H. Kaija, S. L. Alatalo, J. M. Halleen, Y. Lindqvist, G. Schneider, H. K. Väänänen and P. Vihko, *Biochem. Biophys. Res. Commun.*, 2002, **292**, 128–132.
8 S. R. Räisänen, S. L. Alatalo, H. Ylipahkala, J. M. Halleen, A. I. Cassady, D. A. Hume and H. K. Väänänen, *Biochem. Biophys. Res. Commun.*, 2005, **331**, 120–126.
9 G. W. Oddie, G. Schenk, N. Z. Angel, N. Walsh, L. W. Guddat, J. de Jersey, A. I. Cassady, S. E. Hamilton and D. A. Hume, *Bone*, 2000, **27**, 575–584.
10 P. Nuttleman and M. Roberts, *J. Biol. Chem.*, 1990, **265**, 12192–12199.
11 N. Mitić, S. J. Smith, A. Neves, L. W. Guddat, L. R. Gahan and G. Schenk, *Chem. Rev.*, 2006, **106**, 3338–3363.
12 M. D. Jackson and J. M. Denu, *Chem. Rev.*, 2001, **101**, 2313–2340.
13 D. E. Wilcox, *Chem. Rev.*, 1996, **96**, 2435–2458.
14 M. W. Crowder, J. Spencer and A. J. Vila, *Acc. Chem. Res.*, 2006, **39**, 721–728.
15 W. T. Lowther and B. W. Matthews, *Chem. Rev.*, 2002, **102**, 4581–4608.
16 A. J. Ribeiro, M. E. Alberto, M. J. Ramos, P. A. Fernandes and N. Russo, *Chem. – Eur. J.*, 2013, **19**, 14081–14089.
17 E. Ghanem and F. M. Raushel, *Toxicol. Appl. Pharmacol.*, 2005, **207**, 459–470.
18 L. J. Daumann, J. A. Larrabee, D. Ollis, G. Schenk and L. R. Gahan, *J. Inorg. Biochem.*, 2014, **131**, 1–7.
19 D. P. Dumas, H. D. Durst, W. G. Landis, F. M. Raushel and J. R. Wild, *Arch. Biochem. Biophys.*, 1990, **277**, 155–159.
20 F. Ely, K. S. Hadler, L. R. Gahan, L. W. Guddat, D. L. Ollis and G. Schenk, *Biochem. J.*, 2010, **432**, 565–573.
21 M. E. Alberto, G. Pinto, N. Russo and M. Toscano, *Chem. – Eur. J.*, 2015, **21**, 3736–3745.
22 G. Pinto, G. Mazzone, N. Russo and M. Toscano, *Chem. – Eur. J.*, 2017, **23**, 13742–13753.
23 K. S. Hadler, E. A. Tanifum, S. H.-C. Yip, N. Mitić, L. W. Guddat, C. J. Jackson, L. R. Gahan, K. Nguyen, P. D. Carr, D. L. Ollis, A. C. Hengge, J. A. Larrabee and G. Schenk, *J. Am. Chem. Soc.*, 2008, **130**, 14129–14138.
24 K.-P. Hopfner, A. Karcher, L. Craig, T. T. Woo, J. P. Carney and J. A. Tainer, *Cell*, 2001, **105**, 473–485.
25 C. J. Jackson, K. S. Hadler, P. D. Carr, A. J. Oakley, S. Yip, G. Schenk and D. L. Ollis, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2008, **64**, 681–685.
26 G. Schenk, N. Mitić, G. R. Hanson and P. Comba, *Coord. Chem. Rev.*, 2013, **257**, 473–482.
27 M. M. Pedroso, J. A. Larrabee, F. Ely, S. E. Gwee, N. Mitić, D. L. Ollis, L. R. Gahan and G. Schenk, *Chem. – Eur. J.*, 2016, **22**, 999–1009.
28 N. Mitić, M. Miraula, C. Selleck, K. S. Hadler, E. Uribe, M. M. Pedroso and G. Schenk, in *Advances in Protein Chemistry and Structural Biology*, ed. C. Z. Christov, Academic Press, 2014, vol. 97, pp. 49–81.
29 C. J. Jackson, P. D. Carr, J.-W. Liu, S. J. Watt, J. L. Beck and D. L. Ollis, *J. Mol. Biol.*, 2007, **367**, 1047–1062.
30 R. E. Mirams, S. J. Smith, K. S. Hadler, D. L. Ollis, G. Schenk and L. R. Gahan, *J. Biol. Inorg. Chem.*, 2008, **13**, 1065–1072.
31 K. S. Hadler, N. Mitić, S. H.-C. Yip, L. R. Gahan, D. L. Ollis, G. Schenk and J. A. Larrabee, *Inorg. Chem.*, 2010, **49**, 2727–2734.
32 T. J. Paul, G. Schenk and R. Prabhakar, *J. Phys. Chem. B*, 2018, **122**, 5797–5808.
33 G. Schenk, M. L. Neidig, J. Zhou, T. R. Holman and E. I. Solomon, *Biochemistry*, 2003, **42**, 7294–7302.
34 M. L. Neidig, A. T. Wecksler, G. Schenk, T. R. Holman and E. I. Solomon, *J. Am. Chem. Soc.*, 2007, **129**, 7531–7537.

35 G. P. Mullen, E. H. Serpersu, L. J. Ferrin, L. A. Loeb and A. S. Mildvan, *J. Biol. Chem.*, 1990, **265**, 14327–14334.
36 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
37 M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654–3665.
38 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb and J. R. Cheeseman, *Gaussian 09, revision D.01*, Gaussian Inc., Wallingford CT, 2013.
39 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Am. Chem. Soc.*, 2001, **222**, U403.
40 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
41 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
42 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
43 R. Carballo, A. Castineiras, B. Covelo, E. Garcıa-Martınez, J. Niclós and E. Vázquez-López, *Polyhedron*, 2004, **23**, 1505–1518.
44 P. Ren, C. Wu and J. W. Ponder, *J. Chem. Theory Comput.*, 2011, **7**, 3143–3161.
45 Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046–4063.
46 J. W. Ponder, C. J. Wu, P. Y. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.
47 J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht and R. A. DiStasio Jr, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.
48 A. Albaugh, H. A. Boateng, R. T. Bradshaw, O. N. Demerdash, J. Dziedzic, Y. Mao, D. T. Margul, J. Swails, Q. Zeng and D. A. Case, *J. Phys. Chem. B*, 2016, **120**, 9811–9832.
49 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
50 D. J. Price and C. L. Brooks III, *J. Chem. Phys.*, 2004, **121**, 10096–10103.
51 S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952–962.
52 B. Hess, H. Bekker, H. J. Berendsen and J. G. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
53 J. L. Knight and C. L. Brooks, *J. Comput. Chem.*, 2009, **30**, 1692–1700.
54 X. Kong and C. L. Brooks III, *J. Chem. Phys.*, 1996, **105**, 2414–2423.
55 Z. Guo, C. L. Brooks and X. Kong, *J. Phys. Chem. B*, 1998, **102**, 2032–2036.
56 R. Kumari, R. Kumar and A. Lynn, *J. Chem. Inf. Model.*, 2014, **54**, 1951–1962.
57 C. P. Barrett, B. A. Hall and M. E. Noble, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2280–2287.
58 Maestro, Schrödinger, LLC, New York, NY, 2018.
59 T. Ichiye and M. Karplus, *Proteins*, 1991, **11**, 205–217.
60 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
61 E. Krieger and G. Vriend, *Bioinformatics*, 2014, **30**, 2981–2982.
62 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
63 S. Genheden and U. Ryde, *Expert Opin. Drug Discovery*, 2015, **10**, 449–461.
64 J. L. Vanhooke, M. M. Benning, F. M. Raushel and H. M. Holden, *Biochemistry*, 1996, **35**, 6020–6025.
65 C. J. Jackson, J.-L. Foo, H.-K. Kim, P. D. Carr, J.-W. Liu, G. Salem and D. L. Ollis, *J. Mol. Biol.*, 2008, **375**, 1189–1196.
66 K.-Y. Wong and J. Gao, *Biochemistry*, 2007, **46**, 13352–13369.