Calibration of the dianionic phosphate group: Validation on the recognition site of the homodimeric enzyme phosphoglucose isomerase

Marion Devillers,[a] Jean-Philip Piquemal,[b,c] Laurent Salmon,[a] and Nohad Gresh *[b]

We calibrate and validate the parameters necessary to represent the dianionic phosphate group (DPG) in molecular mechanics. DPG is an essential fragment of signaling biological molecules and protein-binding ligands. It is a constitutive fragment of biosensors, which bind to the dimer interface of phosphoglucose isomerase (PGI), an intracellular enzyme involved in sugar metabolism, as well as an extracellular protein known as autocrine motility factor (AMF) closely related to metastasis formation. Our long-term objective is to design DPG-based biosensors with enhanced affinities for AMF/PGI cancer biomarker in blood. Molecular dynamics with polarizable potentials could be used toward this aim. This requires to first evaluate the accuracy of such potentials upon representing the interactions of DPG with its PGI ligands and tightly bound water molecules. Such evaluations are done by comparisons with high-level ab initio quantum chemistry (QC) calculations. We focus on the Sum of Interactions Between Fragments Ab initio computed (SIBFA) polarizable molecular mechanics procedure. We present first the results of the DPG calibration. This is followed by comparisons between ΔE(SIBFA) and ΔE(QC) regarding bi-molecular complexes of DPG with the main-chain and side-chain PGI residues, which bind to it in the recognition site. We then consider DPG complexes with an increasing number of PGI residues. The largest QC complexes encompass the entirety of the recognition site, with six structural water molecules totaling up to 211 atoms. A persistent and satisfactory agreement could be shown between ΔE(SIBFA) and ΔE(QC). These validations constitute an essential first step toward large-scale molecular dynamics simulations of DPG-based biosensors bound at the PGI dimer interface. © 2020 Wiley Periodicals, Inc.

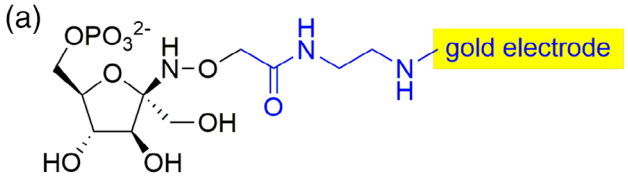
DOI: 10.1002/jcc.26134

Introduction

The dianionic phosphate group (DPG) is ubiquitous in biochemical processes. It is the terminal end of various DNA and RNA nucleotides and several coenzymes including adenosine and guanosine di- and triphosphates (ADP, ATP, GDP, GTP),[1,2] thiamine diphosphate (ThDP),[3] pyridoxal phosphate (PLP)[4] and nicotinamide adenine dinucleotide phosphate (NADPH).[2] Numerous proteins in their lifetimes become phosphorylated/dephosphorylated at Tyr, Ser, Thr, and His side-chains.[5] Although this list is by no means limitative, it is encountered in numerous phosphate sugars in cellular metabolic pathways such as glycolysis, gluconeogenesis, glycogenesis, glycogenolysis,[6,7] pentose phosphate pathway,[8] hexose monophosphate shunt,[9] and so on. One of our laboratories has a long-standing interest in the design of DPG-based sugar inhibitors of aldose-ketose isomerases, such as phosphomannose isomerase (PMI) and phosphoglucose isomerase (PGI). PMI (type I) is a zinc metalloenzyme, essential for the biosynthesis of mannosylated glycoproteins and a potential target in anti-pathogen strategies.[10–12] PGI reversibly isomerizes glucose 6-phosphate into fructose-6-phosphate (F6P) in glycolysis. Outside the cell, it is also known as the cytokine autocrine motility factor (AMF), a cancer progression biomarker.[13] Therefore, high-affinity inhibitors of PGI/AMF could be used both in anticancer therapy and for the early detection of tumor development in serum or urine.[14] We have recently designed and synthesized an F6P-based electrochemical biosensor enabling picomolar detection of AMF in human serum.[15] It embodies a fructose scaffold connected to diphosphate along with an N-oxime connector linking it to a spacer covalently bound to a gold electrode (Scheme 1A). There are high-resolution X-ray studies of the complexes of DPG-based ligands with PGI homodimers. While both monomers contribute to ligand binding, DPG binds tightly to a single monomer.[16–22]

To determine the pK_a of DPG, we have considered several PDB 3D structures of aldose-ketose isomerases complexed with

[a] M. Devillers, L. Salmon
Equipe de Chimie Bioorganique et Bioinorganique, Institut de Chimie Moléculaire et des Matériaux d'Orsay (ICMMO), Univ Paris-Saclay, Univ Paris-Sud, UMR 8182 CNRS, rue du Doyen Georges Poitou, F-91405, Orsay, France
[b] J.-P. Piquemal, N. Gresh
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, Paris, France
E-mail: gresh@lct.jussieu.fr
[c] J.-P. Piquemal
Department of Biomolecular Engineering, The University of Texas at Austin, Texas, 78712
Contract Grant sponsor: Centre Régional Informatique et d'Applications Numériques de Haute Normandie (CRIANN, Rouen, France).; Contract Grant number: 1998053; Contract Grant sponsor: Grand Equipement National de Calcul Intensif (GENCI), Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), Centre Informatique de l'Enseignement Supérieur (CINES, France); Contract Grant number: x2009-075009



Scheme 1. a, Representation of the highly sensitive F6P-based electrochemical biosensor. b, Its model fructose-6-phosphate-*N*-methyloxime (F6P-NOMe) used in this study. [Color figure can be viewed at wileyonlinelibrary.com]

sugar phosphates. It was computed with the PROPKA 3.1[23,24] and PDB2PQR 2.1[25] softwares available on the Web application "Playmolecule ProteinPrepare."[26] Thus, in the human PGI-E4P complex (PDB 1IRI),[21] the pK_a has values of 6.6 and 6.9 for the first and the second monomer, respectively, indicating DPG to be mostly dianionic.

We have previously reported several theoretical studies of the interaction of PMI in complex with its substrate mannose 6-phosphate and inhibitors such as 5-phosphate-D-arabinonohydroxamic acid.[27–30] In this study, in view of molecular dynamic (MD) studies on the complexes of PGI-dimers with DPG-containing molecules, and the interconversion between conformers, we report calibration of DPG and validation on the recognition site of the homodimeric enzyme PGI-AMF using F6P-*N*-methyloxime (F6P-NOMe, Scheme 1B) as a simple mimic of the highly sensitive F6P-based electrochemical biosensor.

Our objective is to perform long-duration polarizable molecular dynamics (PMD) simulations of the complexes of DPG-based ligands with a dimer of PGI, the ligands occupying the hinge inter-monomer position. But for these simulations to be reliable, we deem it essential to prior ensure accurate reproduction of high-level ab initio quantum chemistry interaction energies, ΔE (QC), with the residues and structural waters at the dimer interface. We will resort for such an evaluation to the SIBFA procedure,[31] which was recently coded in a massively parallel PMD code, Tinker-HP,[32] affording the first SIBFA MD simulations of liquid water and enabling numerous perspectives of its extension to ligand-macromolecule complexes. It is useful to recall in the present context that the first SIBFA computations of DPG-containing ligands bore on the complexes of triphosphate and a nonhydrolyzable ATP analog with one up to three divalent cations and with the recognition site of a binuclear enzyme addressing the issue of conformation-dependency of ΔE(SIBFA) compared to ΔE(QC).[33] This was followed by studies of the complexes of PMI with DPG-containing hydroxamate inhibitors[27,30] and DPG-bound mannose surrogates,[29] and of a trianionic diphosphate-containing ligand with the bimetallic isopentenyl diphosphate isomerase enzyme DPPI.[34] In the present study, we intend to further refine the representation of DPG in SIBFA profiting from recent assets:

- Regarding SIBFA, the possibility, for each individual ΔE contribution, to refine automatically and simultaneously the relevant ligand-specific parameters. This is done by least-squares fitting of the difference between this contribution and its QC-counterpart for a range of complexes of the ligand with either a dicationic probe, Zn(II), or a dipolar probe, water, in a diversity of in- and out-of-plane positions (*vide infra*).
- Regarding the correlated ΔE(QC) computations, the recourse to the symmetry adapted perturbation theory (SAPT).[35,36] In the present study, SAPT will be used to complement the Reduced Variational Space (RVS) energy decomposition analyses (EDA)[37] performed at the Hartree-Fock (HF) level in a preliminary stage. SAPT derives values for the first-order electrostatic Coulomb (E_C) and short-range repulsion (E_X) contributions, and for the second-order polarization (E_{pol}) and charge-transfer (E_{ct}) contributions. It also delivers a value of the dispersion contribution, E_{disp}, with a separate value of its repulsive exchange-dispersion component, $E_{disp-exch}$, to which the corresponding SIBFA component could be tailored in model cases.

The outline of this article is as follows. We first focus on the reproduction of ΔE(HF) and its contributions as an essential validation step prior to correlated computations, starting from bimolecular DPG-PGI individual residue complexes and ending with the complex of the DPG-based biosensor with the entirety of the PGI-dimer recognition site, which totals 211 atoms. We first probe DPG with Zn(II) and with water upon performing a series of radial and of in- and out-of-plane variations of their positions of approach. The corresponding variations of the SIBFA versus QC electrostatic, repulsion, polarization, and charge-transfer contributions are plot to evaluate how close an agreement the automatic calibration enables to obtain. This calibration is extended to N-methyl-formamide (NMF), the basic building block of the protein backbone. There are two successive PGI NMFs that anchor DPG in the complex. Correctly accounting for the magnitude of these interactions depends upon the prior and separate DPG and NMF calibrations. ΔE(SIBFA) and ΔE(QC) and their contributions are then compared for each of the bimolecular complexes DPG forms with the PGI side chains of the binding site and with a tightly bound structural water. We next analyze the nonadditivities of DPG binding to an entity made out of two successive, covalently bound NMF fragments making up the backbones of Gly_{211} and Thr_{212} and the side chain of Thr_{212}. It is worth mentioning that most our previous studies addressing nonadditivity[38–40] had a divalent cation in the center of a polyligated complex. Thus, nonadditivity stemmed from the shielding of a dicationic charge by electron-rich ligands. The reverse situation occurs with the polyligated complexes of DPG, now with a dianion in the central binding site, and it is now its negative polarizing field that is shielded by electron-deficient ligands, some of which are covalently bound or in close proximity. A proper control of the variation of E_{pol} and its nonadditivities should be evaluated in this case as well, the more so, as DPG serves as a probe for the polarization response of NMF and covalently bound successive NMF

dimers, as the basic building-block of oligopeptides and proteins. The comparisons are extended to two halves of the recognition site, amenable to EDA with the RVS procedure, and then to the entirety of the DPG-mimic with a completed recognition site, with and without the structural waters.

In the second part of this study, comparisons are resumed at the correlated levels, following a recalibration of E_{disp} and $E_{disp\text{-}exch}$ on the basis of SAPT energy decomposition. Perspectives of applications of large-scale MD simulations are then discussed.

Methods

QC computations

The computations were done with both cc-pVTZ(−f) and aug-cc-pVTZ(f) basis sets.[41,42] The intermolecular interaction energies without EDA analysis were done with the G09 software.[43] At the HF level, the EDAs resorted to the reduced variation space analysis[37] coded in the GAMESS software.[44] At the correlated level, they resorted to the SAPT[35,36] coded in the Psi4 software[45] and performed with the au-cc-pVTZ basis set at the SAPT2 level.[46,47] The DFT calculations were done with the B3LYP-D3 and B97-D3 functionals.[48] The basis set superposition error (BSSE) was computed using the Boys and Bernardi procedure.[49]

SIBFA computations

In the SIBFA procedure, the intermolecular interaction energy is computed as the sum of five contributions: electrostatic multipolar (E_{MTP*}), short-range repulsion (E_{rep}), polarization (E_{pol}), charge transfer (E_{ct}), and dispersion (E_{disp})

$$E_{TOT} = E_{MTP*} + E_{rep} + E_{pol} + E_{ct} + E_{disp}$$

The expressions for the five contributions were given in previous publications.[31] In the present context, the distributed multipoles of DPG and the sugar moiety of the biosensor were derived from the GDMA procedure by Stone[50] and their distributed polarizabilities computed by the Garmer and Stevens[51] procedure. The constitutive fragments making up PGI are taken from the SIBFA library of protein fragments.

Procedure followed

We have for each of the first four contributions refined the relevant parameters of DPG and NMF by an automated procedure minimizing the Least Squares Difference between its SIBFA and QC value in a diversity of their bimolecular complexes with a Zn(II) or a water probe. We resorted to a BFGS procedure, coded under name VA10 from the Harwell library.[52] We specify below the bimolecular complexes and relevant parameters.

Dianionic phosphate group. The choice of Zn(II) as a probe, in keeping with Refs.[33,40], was done to maximize the polarization and charge-transfer responses of probed DPG. Even though some distances of approach are very short (1.6–1.7 Å), none of the four RVS contributions was found to undergo any discontinuous radial variation that would translate the onset of a covalent bond between DPG and Zn(II).

E_{MTP*} and E_{rep}. Eight-four complexes with a Zn(II) probe and 56 complexes with a water probe were considered. Each of the three anionic oxygens was probed in three stages: (1) upon stepwise (0.10 Å) varying its Zn or H_w distance in the range of 1.50–2.10 Å, the PO-Zn/H_w angle being 180°; (2) at a fixed distance of 1.8 Å, performing stepwise (15°) variations of the θ angle PO—Zn/H_w, from 105 till 225°; (3) for four preselected values of θ (225°, 255°, 270°, and 285°), performing out-of-plane variations of the angle ϕ from 0 till 150°. The range of distances is the one encountered in complexes of the phosphate anion with divalent metal cations or water. The 1.8 Å distance selected for the angular variations is close to, or actually equal to, the equilibrium distance in such complexes. For E_{MTP*}, the only parameters are the effective radii of phosphorus and of the anionic phosphate oxygens used in the expression of the penetration component of E_{MTP*}.[53] For E_{rep}, the relevant parameters are the effective radius of the anionic oxygens (O^-); the internal coordinates of the lone pairs with respect to their O bearer, and the increments/decrements of the effective radius of their O bearer along the direction of the lone pair. In this representation of O^-'s, there are six lone pairs, each with an effective population of 1 with a 60° spacing of their successive dihedral angles. This confers better isotropy for out-of-plane variations than a representation with three sp^3 lone pairs with an effective population of 2 on each and 120° dihedral angle spacing. The lone-pair variables are coupled together, that is, the variation of each of these five parameters is constrained to be the same on all lone-pairs.

E_{pol}. We resorted to the sole Zn(II) probe, on account of two points: -Zn(II) exerts a much stronger polarizing field than water; this enables to prevent the induced dipoles of water, if it were used as a probe, to impact DPG polarization. We deemed it preferable to account for these additional effects only in a forthcoming treatment, after the water-relevant E_{pol} parameters are themselves refined. The DPG-relevant parameters are the multiplicative factor and the exponent of the Gaussian function, which screens the polarizing field undergone by DPG, and the effective radii of O^- and of P which enter in the expression of this function.[54] In addition, we resorted to a simplified expression of the three lone pair dipole polarizability tensors, α, on each O^-. The off-diagonal elements, which are on average small, were set to zero. The diagonal elements α_{XX}, α_{YY}, and α_{ZZ} were averaged together. Thus one and the same value was allotted to all 27 elements. We proceeded in similar fashion for the two lone pair polarizabilities of the ester oxygen. A global multiplicative value for the lone pair was then used as an additional fitting parameter. This should, on the one hand, improve the match of E_{pol}(SIBFA) to E_{pol}(RVS) at long as well as at short cation-ligand distances and thus partly remedy some previously noted short-coming of the Garmer-Stevens polarizabilities, with which E_{pol}(SIBFA) was in several instances seen to have a faster decay with distance than E_{pol}(RVS). On the other hand, it should

allow for a more efficient implementation into the Tinker-HP code for MD simulations.

E_{ct}. With both Zn(II) and the water probe in the 140 complexes, there were only two parameters to be fit: the effective O^- radius and its increment along the direction of each lone-pairs, which had been previously optimized upon E_{rep} fitting.

NMF

E_{MTP*} and E_{rep}. Both Zn(II) and water were used to probe O along the CO bond, the Zn/H_w-O distances varying in the range of 1.6–2.2 Å found for NMF complexes with metal cations or water. NMF was also probed along verticals to its plane passing through its O, C, and N atoms, the corresponding H_w/Zn distances being in the range of 2.0–2.4 Å. The amide hydrogen atom was probed by a water molecule approaching along its external bisector the NH bond, the N, H, and O atoms being collinear. The range of distances was 1.5–2.1 Å. This enabled hand fitting of the amide H effective radius.

E_{MTP*}. The sole variables were the effective radii of the amide H and N atoms, those of carbonyl oxygen having been previously calibrated.[40]

E_{rep}. The relevant parameters are the effective O, C, and N radii. Both N and C have two π lone pair orbitals, with starting total occupation numbers of one and of one-half, respectively. O has two in-plane sp^2 lone-pairs and two "smeared" lone-pairs on both sides of the NMF plane projecting out of the C⊠O bond.[40] The total starting occupation number is four. It also has two lone pairs, with starting occupation numbers of one-half each. The variables for the fit are the effective radii of C, N, and O, and five lone-pair variables, namely the three internal coordinates, the electronic population, and the increment of effective-radii along the lone-pair direction. The total population number is constrained to be equal to eight. The variables of each of the C, N, and O π lone-pairs and of the smeared sp^2 O lone-pairs are coupled to retain symmetry above and below the NMF place. Those of the two sp^2 lone-pairs are similarly coupled to retain symmetry on both sides of the CO bond.

E_{pol}. Only Zn(II) was used as a probe, for the same reason as for DPG. The parameters are the effective radii of the N and C amide atoms, that of the carbonyl O having been previously calibrated;[40] the prefactor, and the exponent of the Gaussian screening of the polarizing field; and two multiplicative factors, for the lone pair and for the bond polarizabilities.

E_{ct}. The parameters are the effective radii of N, C, and O atoms, and the increment of effective radii along the directions of the sp^2 and of the sp lone pairs.

Calibration of the hydroxyl groups of Ser and Thr side chains. We have also performed an RVS analysis of the binding of water molecule approaching the hydroxyl proton of methanol, its external bisector being aligned with the direction of the OH bond. This led us to use slightly modified effective radii for methanol-like molecules, as found in Ser and Thr side chains, than for water hydrogens.

We have considered the complex of the DPG biosensor with a PGI dimer, in which the first monomer was made out of residues Met_{84} till Leu_{529}, and the second monomer of residues 335 till 479. The second monomer is the one involved in the binding of the sugar moiety. All these residues cover a zone of 20 Å around the biosensor mimic. We started from the X-ray crystal structure of human PGI/AMF in complex with erythrose-4-phosphate (PDB code 1IRI).[21] Energy minimization was done with the Merlin package.[55] We relaxed the following side-chain torsional angles: *monomer 1:* residues Ile_{157}, Ser_{160}, Asp_{161}, Leu_{162}, Ser_{185}, Asn_{186}, Ile_{187}, Asp_{188}, Ser_{210}, Lys_{211}, Thr_{212}, Thr_{214}, Thr_{215}, Gln_{216}, Glu_{217}, Thr_{218}, Ile_{219}, Asn_{221}, Arg_{273}, Tyr_{274}, Tyr_{351}, Phe_{352}, Gln_{353}, Gln_{354}, Asp_{356}, Met_{357}, Glu_{358}, Ser_{359}, Asp_{511}, Gln_{512}, Trp_{513}, Val_{515}, Glu_{516}, Lys_{519}; and *monomer 2:* residues Tyr_{341}, Glu_{382}, Thr_{385}, Asn_{386}, Gln_{388}, His_{389}, Phe_{391}, Tyr_{392}, Gln_{393}, Leu_{394}, Ile_{395}, His_{396}, Ile_{424}, Leu_{425}, Asn_{428}, Gln_{432}. The backbones of both PGI monomers were not relaxed. The DPG biosensor was relaxed in terms of its six intermolecular variables and the six intermolecular variables defining its binding position in the cavity. Six structural water molecules identified from the X-ray structure were included and relaxed. Our present purpose is to identify the predominant DPG biosensor mimic-PGI interactions, and how well they are accounted for in SIBFA as compared to QC. This is a prerequisite to reliable MD simulations with a full relaxation of the energy hypersurface.

Results and Discussion

Figure 1 represents the structure of the complex of the DPG biosensor mimic with the closest residues of the dimer recognition site. An extended view is given as Supporting Information S1. DPG is bound by the backbone of Lys_{211} and Thr_{212}, and three hydroxyl groups belonging to the side chains of Ser_{210}, Th_{r212}, Thr_{215}, as well as with structural water W_1, which also acts as a proton acceptor from Ser_{210}. At the outcome of energy minimization, additional interactions were found to take place: between one water molecule which acts on the one hand, as a proton acceptor from the hydroxyl connected to C_3, and on the other hand, as a proton donor to the hydroxyl oxygen connected to C_4; and between the hydroxyl oxygen of the extracyclic CH_2OH entity, which accepts a proton from the main-chain amide group of Gly_{159}. There is an intramolecular interaction between one anionic DPG oxygen and the hydroxyl group connected to C_4. Additional interactions involve two water molecules. W_2 bridges Thr_{215} and Lys_{519} side chains, acting as a proton acceptor from the latter and a proton donor to the former. W_3 bridges the side chains of Glu_{217} and Lys_{519}, acting as proton donor to the former and as a proton acceptor to the latter. The His_{389} side chain belonging to monomer B acts as the third proton acceptor from Lys_{519} side chain. This is a consequence of the minimization procedure, which resulted into a near −156° flip of the χ_2 His_{389} torsion angle (86° after

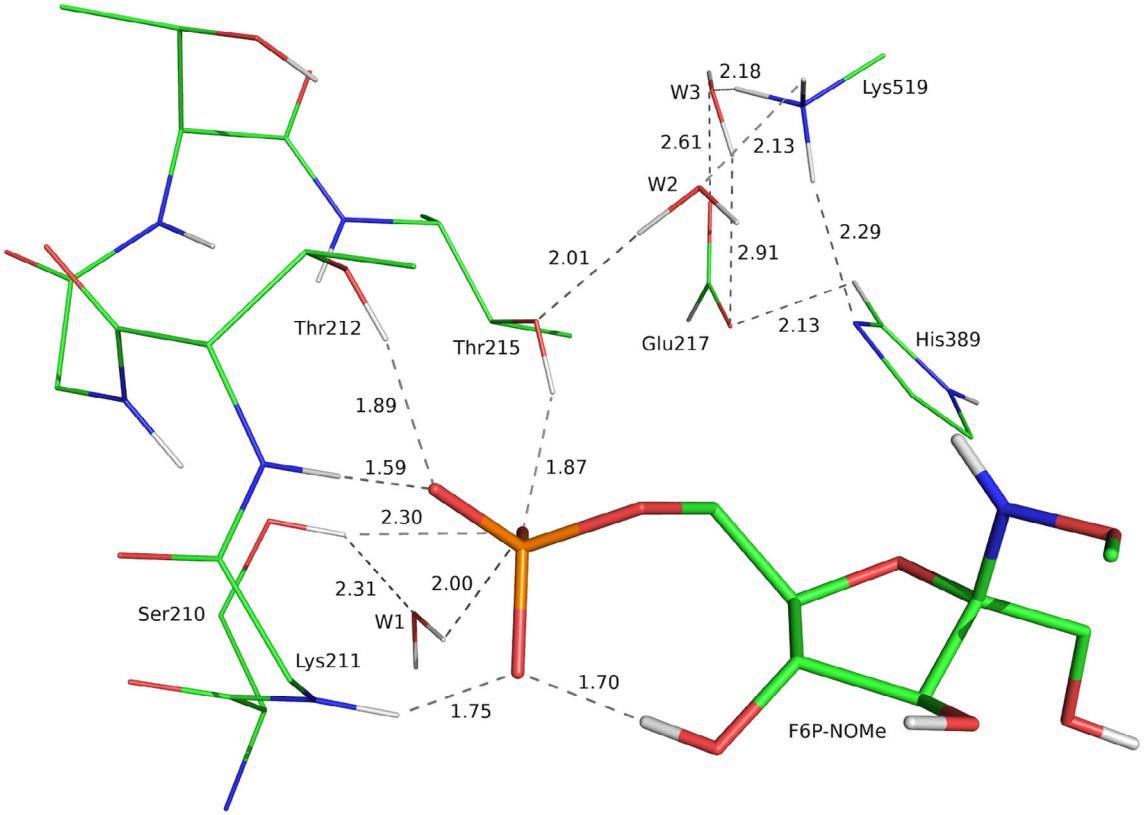


Figure 1. Representation of the complex of the dianionic inhibitor with the PGI recognition site. [Color figure can be viewed at wileyonlinelibrary.com]

minimization and 242° in the starting structure). This also enabled the partly acidic C_ε-bound proton to come into close contact to one anionic Glu_{217} oxygen.

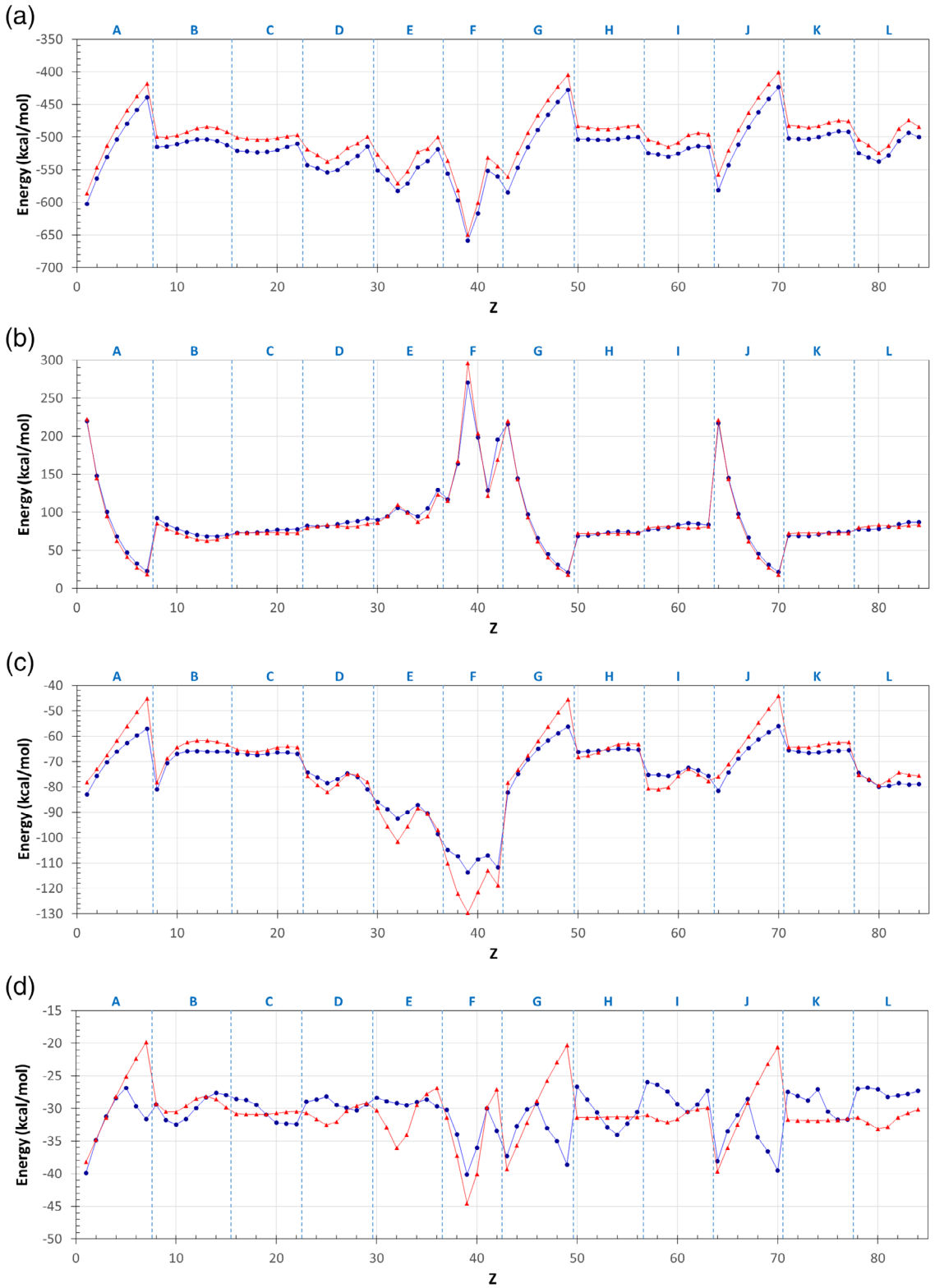
Such a flip occurred at the early stages of energy minimization. It is a consequence of the choice of the His_{389} tautomer having a protonated N_δ and a deprotonated N_ε, since it is the most commonly occurring tautomer in proteins. However, using instead the tautomer with a deprotonated N_δ and a protonated N_ε enables, in the starting X-ray structure, N_ε to interact favorably with the Glu_{217} side-chain. Such an interaction could compete with the complex formed between Lys_{519} and the deprotonated N_ε of the "major" tautomer. This competing possibility implies that future MD simulations of DPG-PGI complexes should consider both His tautomers, involving different interplays of intra- and inter-monomer and water–PGI interactions. The extent to which both forms could coexist and/or compete could be evaluated in the framework of newly developed constant pH simulation methods.[56–60]

Toward this perspective, we also provide, following the results from the fitting process, an evaluation of SIBFA as compared to QC for single-point computations on the starting X-ray structure, now involving the deprotonated N_δ His_{389} tautomer. It will bear on a model complex involving DPG, the His_{389} side chain and the closest PGI side chains to them: Thr_{215}, Glu_{217}, and Lys_{519}. It should enable to evaluate how well ΔE(SIBFA) and its contributions compare to their QC counterparts with different interplay of interactions than those resulting from the χ_2 flip.

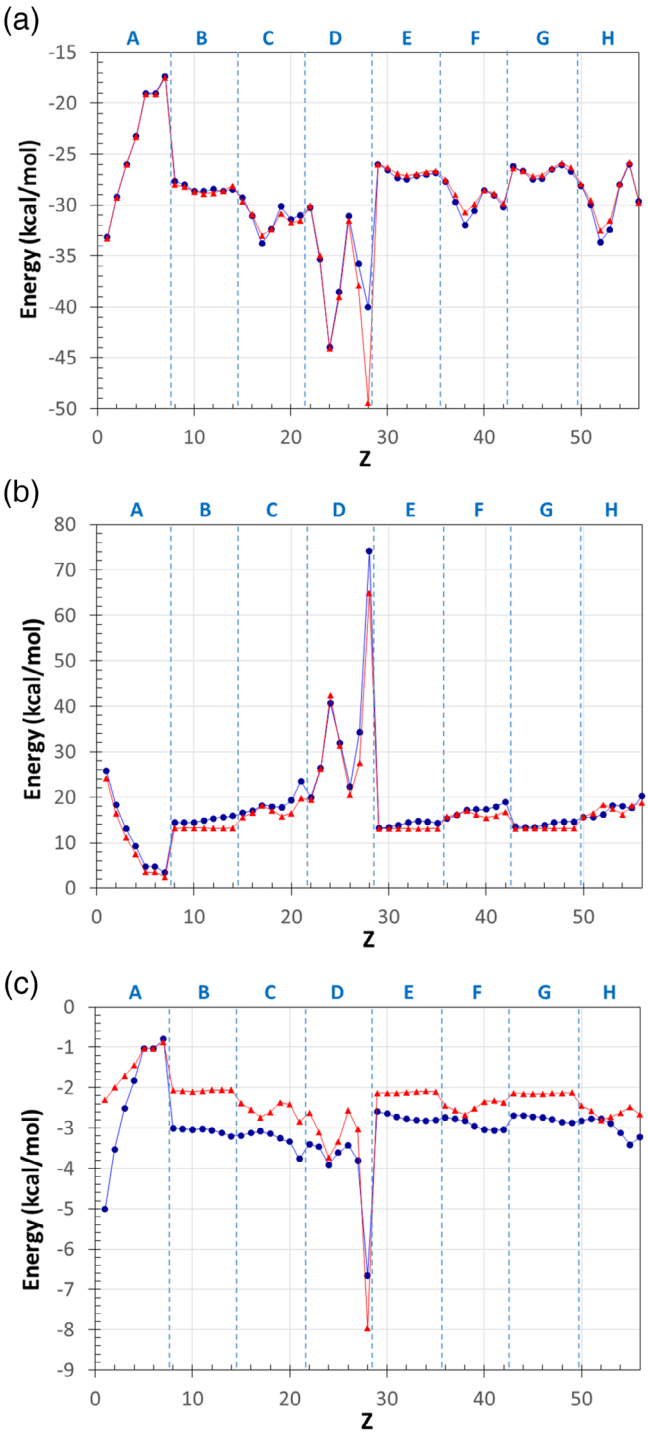
Result of the fits

DPG. For the DPG-Zn(II) complex, the compared radial and in- and out-of-plane variations of E_{MTP}, E_{rep} E_{pol}, and E_{ct} and of their QC counterparts are plotted in Figures 2a–2d. For the DPG–water complex, the corresponding variations of E_{MTP*}, E_{exch}, E_{rep}, and E_{ct} and of their QC counterparts are plotted in Figures 3a–3c. The SIBFA and QC curves are displayed in red and blue, respectively.

E_{MTP*} reproduces closely the trends of E_C throughout. It underestimates E_C in all Zn(II) complexes. Although a much closer match is possible by increasing the values of the anionic O effective radius of the anionic oxygen, this would result into a concomitantly overestimated E_{MTP}* value in the complexes with water. A lack of simultaneous accurate reproduction of both Zn(II) and water complexes could translate some lack of flexibility of the present calibration of E_{pen}, which underestimates the contribution of charge-quadrupole to the penetration component of E_{MTP}*, E_{pen}.[61] This is prevented in ongoing recalibration of E_{pen} with correlated multipoles on the basis of SAPT EDAs aiming to construct a new library of SIBFA fragments (see discussion below). Overall, the present tests make aware of the need to have an equally improved reproduction of both Zn(II) and water complexes with a diversity of targeted ligands. E_{rep} reproduces closely E_X in the water as well as in the Zn(II) complexes. This implies that formulation of E_{rep} with six lone pairs with occupation numbers of one each does lend



Figures 2. a–d) Complexes of dianionic phosphate with a dicationic probe, Zn(II). Compared radial and in- and out-of-plane variations of a) E_C and E_{MTP}; b) E_{exch} and E_{rep}; c) E_{pol}; d) E_{ct}. The QC and SIBFA curves are drawn in blue and red, respectively. We denote by O_1, O_2, and O_3 the three anionic O atoms of DPG. The points are defined as follows. Points 1–7 (interval A): 0.1 Å variations of the Zn-O_1 distance from 1.5 to 2.1 Å. Points 8–15 (B): 15° variations of the P-O-Zn angle, from 105 to 225°. Points 16–22 (C): 30° variations of the O-P-O-Zn angle, from 0 to 180°, at fixed θ angle of 225°. Points 23–29 (D): 30° variations of the O-P-O-Zn angle, from 0 to 180°, at fixed θ angle of 255°. Points 30–36 (E): 30° variations of the O—P—O—Zn angle, from 0 to 180°, at fixed θ angle of 270°. Points 37–42 (F): 30° variations of the O-P-O-Zn angle, from 0 to 180°, at fixed θ angle of 270°. Points 43–49 (G): 0.1 Å variations of the Zn-O_2 distance from 1.50 to 2.1 Å. Points 50–56 (H): 30° variations of the O—P—O—Zn angle, from 0 to 180°, at fixed θ angle of 225°. Points 57–63 (I): 30° variations of the O-P-O-Zn angle, from 0 to 180°, at fixed θ angle of 255°. Points 64–70 (J): 0.1 Å variations of the Zn-O_3 distance from 1.50 to 2.1 Å. Points 71–77 (K): 30° variations of the O-P-O-Zn angle, from 0 to 180°, at fixed θ angle of 225°. Points 78–84 (L): 30° variations of the O—P—O—Zn angle, from 0 to 180°, at fixed θ angle of 255°. [Color figure can be viewed at wileyonlinelibrary.com]



Figures 3. a–c) Complexes of dianionic phosphate with a dipolar probe, water. Compared radial and in- and out-of-plane variations of a) E_C and E_{MTP}; b) E_{exch} and E_{rep}; c) E_{ct}. The QC and SIBFA curves are drawn in blue and red, respectively. We denote by O_1, O_2, and O_3 the three anionic O atoms of DPG. The points are defined as follows. Points 1–7 (interval A): 0.1 Å variations of one H-O_1 distance from 1.6 to 2.2 Å. Points 8–14 (B): 15° variations of the O-P-O_1-H angle, from 105 to 225°. Points 15–21 (C): 15° variations of the O-P-O_1-H angle, from 105 to 225°, at fixed θ angle of 255°. Points 22–28 (D): 15° variations of the O-P-O_1-H angle, from 105 to 225°, at fixed θ angle of 270°. Points 29–35 (E): 30° variations of the O-P-O_2-H angle, from 0 to 180°, at fixed θ angle of 225°. Points 36–42 (F): 30° variations of the O—P—O_2—H angle, from 0 to 180°, at fixed θ angle of 255°. Points 43–49 (G): 30° variations of the O—P—O_3—H angle, from 0 to 180°, at fixed θ angle of 225°. Points 50–56 (H): 30° variations of the O—P—O_3—H angle, from 0 to 180°, at fixed θ angle of 255°. [Color figure can be viewed at wileyonlinelibrary.com]

itself to a close reproduction of E_X. We report in Supporting Information S2 the relevant parameters defining the lone pairs in terms of internal coordinates, and increments of effective radii of the bearer O atoms.

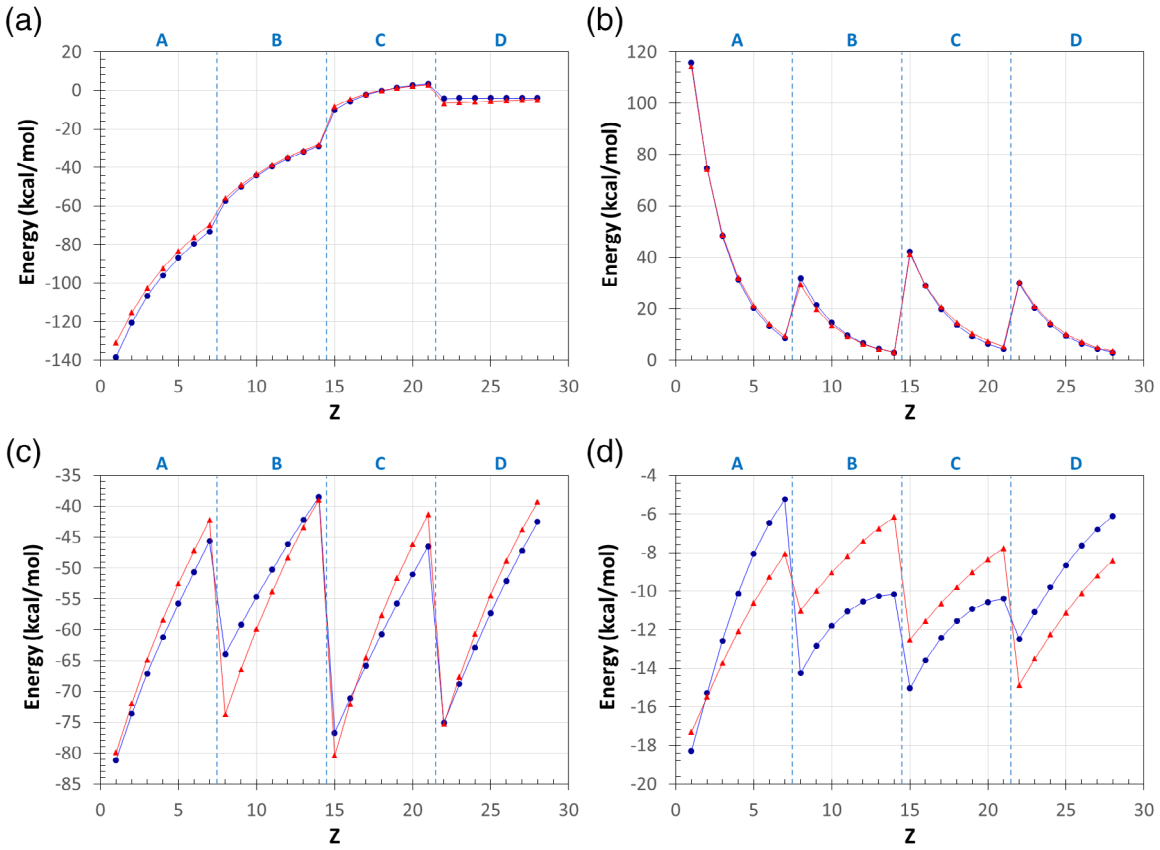
E_{pol} and E_{ct} are also correctly reproduced except in four high-lying energy zones. This is the case: (1) for position 39, where E_{rep} and Ex are maximal, with values over 250 kcal/mol, so that this position is unlikely to be populated; and (2) for positions 7, 49, and 70, corresponding to the longest probed Zn-O distance of 2.10 Å, at which E_{MTP}* and E_c have their smallest magnitudes. Regarding E_{pol}, this could translate some shortcoming of the present representation of E_{pol}(SIBFA), which in divalent cation complexes decreases faster than E_{pol}(QC) upon increasing the intermolecular distance. It should be attenuated in polyligated complexes of divalent cations, due to the screening effects of the dication charge exerted on each ligand by the fields of the other ligands. The fact that E_{ct}(QC) undergoes an actual increase, rather than a decrease, in magnitude upon increasing the Zn(II)-phosphate distance as occurs in regions 45–49 and 67–70, translates the onset of asymptotic states in which an electron is transferred to the dication.[62] This was already noted in the earliest SIBFA calibration of divalent cations and their complexes with anionic ligands,[63] and it should not be considered as a limitation of SIBFA. Such states are not anyway encountered in polyligated complexes of divalent cations.

NMF. All four SIBFA contributions reproduce closely their QC counterparts, except, as with DPG, for a limited number of high-lying energy points. In Supporting Information S3, we report the relevant parameters for the NMF lone pairs (Figs. 4a–4d).

Bimolecular complexes

Table 1 reports the comparisons between ΔE(SIBFA) and ΔE (RVS) and their contributions for the complexes of DPG with the NMF entities of Lys_{211}, Thr_{212}, and the side chains of Ser_{210}, Thr_{212}, and with structural water W_1.

ΔE(SIBFA) has a close agreement with ΔE(RVS) for the DPG-NMF_{211} complex, regarding both E_1 and E_2. There are, however, some compensation of errors within E_1 between E_{MTP*} and E_C, the former being overestimated with respect to E_C, while E_{rep} is overestimated with respect to E_{exch}. The DPG-NMF_{212} complex gives rise to the least agreement in the series of the five complexes of Table 1. In this case, the overestimation of E_{rep} with respect to E_{exch} exceeds that of E_{MTP*} with respect to E_C, so that E_1(RVS) is larger than E_1(SIBFA) by 2.2 kcal/mol, while the difference was 0.1 kcal/mol with NMF_{211}. This could to some extent reflect the shorter phosphate-amide H-bond distances (1.59 Å versus 1.75 Å). It implies that in the next stage of refinements, planned to be at the correlated level, the approach to the NMF entity should encompass a larger number of points at short distances than presently done, and that the interactions should not be limited to linear approaches along the NH bond but involve as well out-of-plane approaches as performed in the case of DPG. For the NMF_{212} complex, another source of error stems from E_{pol}, which is by 3 kcal/mol underestimated with respect to

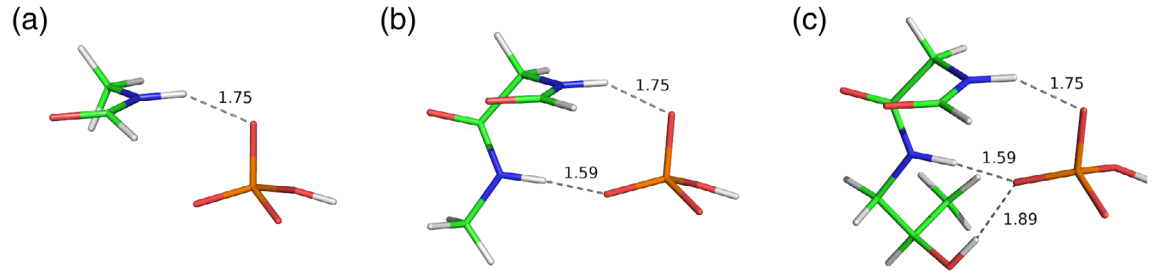


Figures 4. a–d) Complexes of NMF with a dicationic probe, Zn(II). Compared radial and in- and out-of-plane variations of a) E_C and E_{MTP}; b) E_{exch} and E_{rep}; c) E_{pol}; d) E_{ct}. The QC and SIBFA curves are drawn in blue and red, respectively. The points are defined as follows. Points 1–7 (interval A): 0.1 Å variations of the Zn—O distance, from 1.6 to 2.2 Å; the C—O—Zn angle is 180°. Points 8–14 (B): 0.1 Å variations of the Zn—O distance from 2.0 to 2.6 Å; the Zn—O bond is perpendicular to the formamide plane. Points 15–21 (C): 0.1 Å variations of the Zn—N distance from 2.0 to 2.6 Å; the Zn—N bond is perpendicular to the formamide plane. Points 22–28 (D): 0.1 Å variations of the Zn—C distance from 2.0 to 2.6 Å; the Zn—C bond is perpendicular to the formamide plane. [Color figure can be viewed at wileyonlinelibrary.com]

E_{pol}(QC). E_{pol} is clearly dominated by the NMF polarization by the doubly anionic phosphate. The present result implies that upon passing to correlated levels, NMF refinements should probe in more detail than done presently, not only electron-rich sites of NMF by cationic probes such as Zn(II) but also electron-deficient sites, such as the NH bond, by anionic probes, and, similar to E_{rep}, explore out-of-plane binding and not be limited to in-plane approaches. The three DPG complexes with Ser_{210}, Thr_{212} side chains have, similar to the NMF ones, overestimated E_{MTP} values. This is to a large extent a consequence of the overestimation of E_{MTP}, which is inherent to its calibration as it was not possible to have simultaneously close reproductions of E_c for DPG complexes with Zn(II) and with water probes (see discussion above). Except for the DPG-Thr_{212} complex, this gives rise to overestimations of ΔE(RVS) by 1.6 out of 12 kcal/mol for DPG-Ser_{210} and by 2.2 kcal/mol out of 8 for DPG-W_1.

Table 1. QC and SIBFA intermolecular interaction energies and their contributions (kcal/mol) of diphosphate with: (a) the main-chain NMF entities of Lys_{211} and Thr_{212}; (b) the side chains of Ser_{210}, Thr_{212}, and structural water W_2.

	(a)				(b)					
	NMF_{211}		NMF_{212}		Ser_{210}		Thr_{212}		W2	
Complexes A	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA
E_C/E_{MTP}	−37.6	−40.6	−43.3	−46.0	−10.1	−11.8	−22.9	−24.0	−11.6	−13.0
E_{exch}/E_{rep}	27.9	30.7	29.9	34.7	7.0	8.6	12.9	12.3	13.8	13.8
E_1	−9.7	−9.8	−13.5	−11.3	−3.0	−3.2	−10.0	−11.7	2.1	0.9
E_{pol}	−19.7	−19.4	−20.9	−17.7	−7.9	−8.0	−12.4	−9.8	−6.8	−7.3
E_{ct}	−3.7	−4.8	−5.3	−4.4	−0.7	−2.0	−2.0	−2.8	−1.1	−1.6
E_2	−23.4	−24.2	−26.2	−22.0	−8.6	−10.0	−14.4	−12.6	−7.9	−8.8
ΔE	−33.2	−34.0	−39.6	−33.3	−11.6	−13.2	−24.3	−24.3	−5.8	−8.0



Figures 5. a–c) Representation of the complex of dianionic phosphate with: a) one single main-chain NMF fragment; b) two successive NMF fragments; c) the two connected NMF_{211}, NMF_{212}, and Thr_{212} side-chain. [Color figure can be viewed at wileyonlinelibrary.com]

Table 2. ΔE(QC) and ΔE(SIBFA) values (kcal/mol) or dianionic phosphate binding to: (a) the two connected NMF_{211} and NMF_{212} main-chain fragments and (b) the connected two NMF fragments and Thr_{212} side chain.

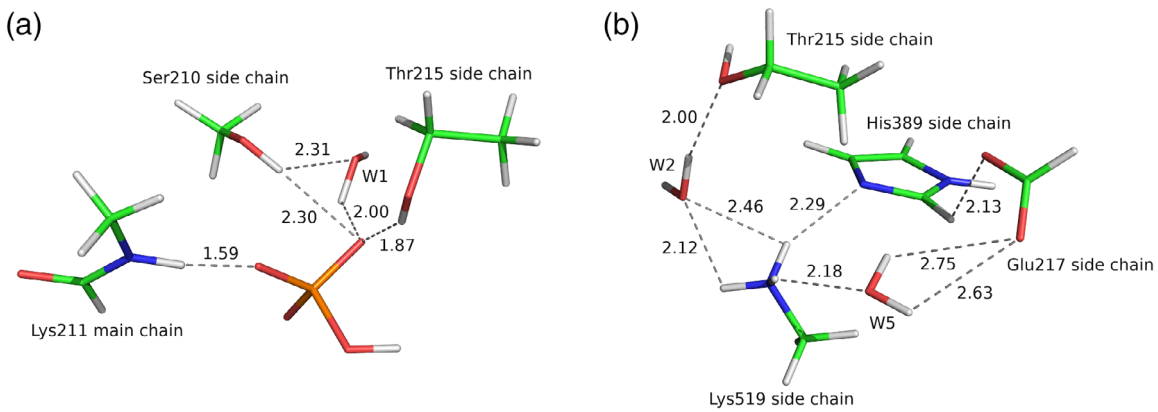
	(a)		(b)	
	QC	SIBFA	QC	SIBFA
E_C/E_{MTP}	−79.5	−82.5	−99.1	−107.2
δ_{nadd}	1.4	4.1	4.7	3.4
E_{exch}/E_{rep}	58.8	62.5	72.5	75.6
δ_{naddp}	1.0	−2.9	1.8	−2.1
E_1	−20.7	−20.0	−26.6	−31.6
δE_{1nadd}	2.5	1.1	6.6	1.2
E_{pol}	−34.7	−30.5	−44.1	−36.7
δE_{pol}	**5.9**	**6.6**	**8.9**	**10.2**
E_{ct}	−8.4	−6.4	−9.8	−10.3
δE_{ct}	**0.6**	**2.8**	**1.2**	**1.7**
δE_2	**6.5**	**9.4**	**10.1**	**11.9**
ΔE	**−63.6**	**−57.0**	**−80.8**	**−78.7**
δE	**9.3**	**10.3**	**16.3**	**12.9**

DPG binding to two or three covalently bound fragments

Because PGI is assembled from its constitutive fragments, we deemed it instructive to compare ΔE(SIBFA) and ΔE(RVS) for DPG binding to: (1) the two successive, covalently bound, NMF entities of Lys_{211} and Thr_{212} backbones, denoted as NMF_{211}-NMF_{212}; (2) this dipeptide covalently bound to the side chain of Thr_{212}, denoted as NMF_{211}-NMF_{212}-Thr_{212}. The QC calculations bear on those actual entities, as represented in Figure 5.

Following a procedure detailed in previous papers,[64] the SIBFA NMF_{211}-NMF_{212} dipeptide is assembled a as juxtaposition of two successive NMF entities, the fictitious hydrogens of the C_α-H(NMF_{211}) and HC(NMF_{212}) junction bonds donating their multipoles to the midpoint of the created $C_\alpha C$ bond. The polarization energy is computed with a set of multipoles without fictitious H redistributions, the two H atoms being instead retracted along the connecting bond on their bearer C atom. This enables E_{pol} to be computed between covalently connected fragments retaining their net original charge of zero, while preventing too close overlaps between the CH bonds. A similar procedure is followed to handle the connection between theThr_{212} backbone and side-chain fragments, namely NMF_{212} and methanol. The results are reported in Table 2 which lists the values of ΔE(SIBFA), ΔE(RVS) and their contributions, and their nonadditivities, δE_{cov}. For each contribution, δE_{cov} denotes the difference between its value in the covalently bound complex and that of the sum of its values in the DPG complexes with the individual fragments.

DPG-NMF_{211}-NMF_{212}. ΔE(SIBFA) differs from ΔE(RVS) by a quantity, 6.6 kcal/mol, slightly larger than the 5.5 kcal/mol sum of the corresponding differences in the DPG complexes with NMF_{211} and NMF_{212}. E_1 has small δE_{cov} values in both RVS and QC calculations. These are larger and of opposite signs for E_{MTP} and E_{rep}(SIBFA). This could be due to the fact that with the individual fragments DPG "sees" two separate CH bonds rather than a single CC bond in the connected dipeptide backbone, whence their smaller magnitudes in the latter. The summed values of



Figures 6. a and b) Representation of subcomplexes making up the recognition site. a) local interactions with DPG; b) local interactions between PGI side chains and two structural waters. [Color figure can be viewed at wileyonlinelibrary.com]

Table 3. Values (kcal/mol) of ΔE(QC) and ΔE(SIBFA) and their contributions in two complexes extracted from the recognition site: (a) dianionic phosphate-NMF_{212}-Ser_{210}-Thr_{215}-W_2 and (b) Thr_{215}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5.

	(a)			(b)		
	aug-cc-pvtz	cc-pVTZ	SIBFA	aug-cc-pvtz	cc-pVTZ	SIBFA
E_C/E_{MTP}	−88.6	−84.3	−99.2	−135.3	−131.6	−138.5
E_{exch}/E_{rep}	74.0	62.3	81.4	43.8	35.8	43.8
E_1	−14.6	−22.0	−17.9	−91.5	−95.8	−94.6
E_{pol}	−39.0	−37.5	−36.1	−23.1	−22.5	−25.2
E_{ct}	−8.5	−9.3	−7.7	−3.6	−3.9	−5.0
E_2	−47.5	−46.8	−43.8	−26.7	−26.4	−30.2
ΔE	−62.3	−66.8	−61.7	−118.3	−122.1	−124.8

δE_{cov} for E_{pol} and E_{ct} are larger, 6.5 and 9.4 kcal/mol in QC and SIBFA, respectively, but the final δE_{cov} have closer values, 9.3 and 10.3 kcal/mol, respectively. It is to be noted that δE_{cov} behaves differently from the energy difference denoted δE_{nadd},[40] which denotes the nonadditivity in multimolecular, noncovalent intermolecular complexes. In these, the nonadditivity of E_C/E_{MTP} is null, and the redistribution of multipoles upon complex formation is taken into account by the dipolar polarization.[64,65]

DPG-NMF_{211}-NMF_{212}-Thr_{212}. ΔE(SIBFA) is much closer to ΔE (RVS) than in the former case, the difference being reduced to 2 kcal/mol out of 80. The summed QC and SIBFA δE_{cov} values are themselves close for E_2 (10.1 and 11.9 kcal/mol, respectively) but the corresponding E_1 values differ significantly (6.6 and 1.2 kcal/mol). While for E_1 δE_{cov} is small in SIBFA (1.2 kcal/mol), it is large in QC (6.6 kcal/mol). This translates the further increases of the differences in QC electronic distribution in the covalently bound entity as compared to the one in the isolated fragments, which is not the purpose of E_{pol} to account for, as it is destined for noncovalent interactions.

Multimolecular complexes. As an intermediate step prior to considering the complete binding site, we have considered two noncovalent intermolecular complexes contributing to stabilize the complexes in the dimer cavity. The first takes place between DPG, NMF_{211}, the side chains of Ser_{210} and Thr_{215}, and structural water W_1. The second does not involve DPG, but PGI residues involved in stabilizing the dimer interface, and two other structural waters, denoted as W_2 and W_5 (Figs. 6a and 6b). The results are given in Table 3. The RVS analyses were done with both aug-cc-pVTZ(−f) and cc-pVTZ(−f) basis sets. The latter was the only one tractable for the large complex of the DPG biosensor mimic with the PGI dimer recognition site, and it was essential to compare the magnitudes of cc-pVTZ(−f) ΔE(RVS) and their contributions to those of the larger basis set.

For complex a, ΔE(SIBFA) amounts to −61.7 kcal/mol and is much closer to the −62.3 kcal/mol aug-cc-pVTZ(−f) value than to the −66.8 kcal/mol cc-pVTZ(−f) one. However, there are some error compensations. Thus compared to the aug-cc-pVTZ (−f), E_{MTP} is larger than E_C by 10.6 kcal/mol, while E_{rep} is larger than E_{exch} by 7.4 kcal/mol, so that E_1(SIBFA) differs from E_1(RVS) by 3.3 kcal/mol. This is a reflection of the trends observed in the bimolecular complexes of DPG with NMF_{211}, Ser_{210} side chain, and W_1. It is noted that the cc-pVTZ(−f) basis set gives rise to an even larger magnitude of E_1 (−22 kcal/mol) than SIBFA. On the other hand, E_2(SIBFA) is smaller than E_2(RVS) with either basis set. For completeness, we provide a new table in Supporting Informaiton S4 comparing ΔE(SIBFA) and ΔE (QC) and their contributions in the most relevant binary complexes extracted from the penta-molecular complex b of Table 3.

Table 4. Values (kcal/mol) of ΔE(QC) and ΔE(SIBFA) in two complexes extracted from the recognition site. a

	aug-cc-pVTZ	cc-pVTZ	SIBFA
a	−79.2	−87.5 (−75.9)	−75.2
b	−181.2	−193.1 (−179.6)	−188.7
c	−158.1	−167.4 (−158.1)	−160.2
d	−111.7 (−111.0)	−117.2 (−112.4)	−114.8
e	−87.3 (−86.8)	−89.3 (−87.3)	−85.6
f		**−227.8 (−209.1)**	**−213.8**
g		**−273.0 (−242.6)**	**−253.2**

a) dianionic phosphate-NMF_{212}-Thr_{212}-Ser_{210}-Thr_{215}-W_1; b) dianionic phosphate-NMF_{211}-Thr_{215}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; c) dianionic phosphate-NMF_{211}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; d) Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; e) Glu_{217}-His_{389}-Lys_{519}; f) Complete binding site without structural waters; g) Complete binding site with structural waters. The QC values under parentheses are the BSSE-corrected ones. The SIBFA values under parentheses are those with a vdW radius for E_{rep} of 1.79 Å on all three imidazole carbons.

Intermolecular interactions in the complete binding site and in subsites. The SIBFA and QC intermolecular interaction energies of the DPG-based biosensor mimic with the entirety of the PGI dimer recognition site, both without and with the six structural waters are given in Table 4. The complex totals 211 atoms. Prior to considering this complex, we list the results found in some of it subsets. As a follow-up to the previous section, this should ensure that the overall agreement on the complete site does not stem from compensating imbalances in some of these sites. We have considered the complex of DPG with: (a) NMF_{212}-Thr_{212}-Ser_{210}-Thr_{215}-W_1; (b) NMF_{211}-Thr_{215}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; and (c) NMF_{211}-Glu_{271}-His_{389}-Lys_{519}-W_1-W_5. We have also considered complexes (d) Glu_{217}-His_{389}-Lys_{519}-W_1-W_5 and (e) Glu_{217}-His_{389}-Lys_{519} both without DPG. For complex a, ΔE (SIBFA) has close agreements with ΔE(QC) with both aug-cc-pVTZ and cc-pVTZ basis sets after removal of BSSE. However, for complex b, the agreement is better with the smaller basis set, and prior to the BSSE correction, while for complex e it is close throughout, the BSSE correction being small in any case.

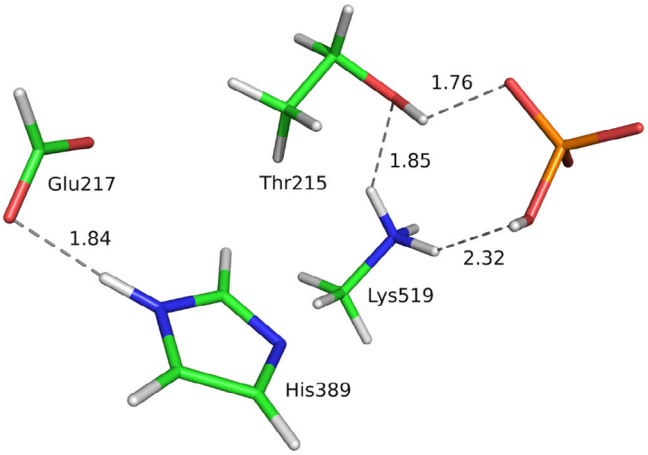


Figure 7. Representation of a five-fragment complex extracted from the X-ray structure and involving DPG and the side-chains His_{389} (protonated N_ε tautomer), Thr_{215}, Glu_{217}, and Lys_{519}. [Color figure can be viewed at wileyonlinelibrary.com]

For complexes c, d, f and g, the values of ΔE(SIBFA) are intermediate between the ΔE(QC) BSSE-uncorrected and corrected cc-pVTZ(−f) values.

At this stage preceding the correlated ΔE calculations, and given the size of the complexes (193 and 211 atoms for f and g, respectively), the fact that ΔE(SIBFA) retains values close to those of ΔE(QC), constitutes an essential validation in the prospect of future large-scale molecular dynamics of the PGI dimer-DPG biosensor mimic complexes.

Complex with the deprotonated N_δ His_{389} tautomer

The flip of the χ_2 torsion angle of His_{389} enabling its N_ε to interact with Lys_{519} could be a consequence of our present choice of the protonated N_δ His_{389} tautomer. Thus, we need to also consider in a preliminary evaluation stage the alternative, minor His_{389} tautomer with a protonated N_ε instead. In the starting X-ray structure, N_ε can donate its proton to the side chain of Glu_{271}. There could thus be two competing forms for the PGI-DPG complex, each corresponding to one defined tautomer. Could such two forms coexist in solution, or would one outcompete the other and how well could MD simulations with polarizable potentials rank them energetically? This could be addressed by future constant-pH simulations.[57–60] But it is necessary to perform a preliminary evaluation of the performance of SIBFA at least in a simplified model complex extracted from the starting X-ray structure. We have thus considered, without optimization, one five-fragment complex. It involves DPG and the side chains of the two closest residues: Thr_{215}, and Lys_{519}, together with His_{389} and Glu_{271}. This complex is represented in Figure 7. The values of ΔE(SIBFA) and ΔE (QC) and their contributions are given in Table 5, which also gives the results regarding four "binary" complexes, denoted as Glu_{217}-His_{389}, Thr_{215}-DPG, Lys_{519}-DPG, and Thr_{215}-Lys_{519}.

In the five-fragment complex, ΔE(SIBFA) has a close numerical agreement with ΔE(QC), −170.0 as compared to −170.9 kcal/mol, respectively. There is also a close agreement regarding the "pivotal" Glu_{217}-His_{389} complex, ΔE(QC) and ΔE(SIBFA) amounting to −24.1 and −25.1 kcal/mol, respectively, and this carries out to all four energy contributions. There is a close agreement as well regarding the Lys_{519}-DPG complex, ΔE(QC) and ΔE(SIBFA) amounting to −161.8 and −158.0 kcal/mol, respectively, also found again with the individual contributions. The Thr_{215}-Lys_{519} complex has a satisfactory agreement of ΔE(SIBFA) with ΔE(QC) (−9.3 as compared to −8.1 kcal/mol). This complex has a value of E_1 that is either slightly repulsive (1.4 kcal/mol for QC) or slightly attractive (−1.2 kcal/mol for SIBFA), due to the fact that the complex was not minimized at this stage. The least agreement is for the Thr_{215}-DPG complex (−5.4 and −1.5 kcal/mol for ΔE(QC) and ΔE(SIBFA), respectively). The unusually repulsive values of E_1 (14.0 and 15.3 kcal/mol in QC and SIBFA respectively) are a clear indication of the need for energy optimization. We did not carry these again here. Indeed, our previous results on binary complexes of DPG with Ser and Thr side chains in energy-minimized structures reported in Table 1 did prove the reliability of SIBFA to handle these.

Correlated calculations

The complete intermolecular interaction energy embodies correlation/dispersion. In the standard SIBFA calibration, this is accounted for by a dispersion contribution which embodies $1/R^6$, $1/R^8$, $1/R^{10}$ and an exchange-dispersion component.[66] A more involved procedure resorts to distributed multipoles and polarizabilities derived at the correlated level, and an accordingly rescaled E_{disp} to match the results from SAPT. This approach is being generalized to all building blocks of proteins, nucleic acids, and to ligands. The present study remains in the context of the standard calibration, and it should be instructive

Table 5. QC and SIBFA intermolecular interaction energies and their contributions in a starting X-ray complex encompassing His_{389} side-chain, DPG, and their closest PGI side chains. A penta-molecular Thr_{215}-Glu_{217}-Lys_{519}-His_{389}-DPG complex is considered together with four bimolecular complexes: Glu_{217}-His_{389}, Thr_{215}-DPG, Lys_{519}-DPG, and Thr_{215}-Lys_{519} (Energies in kcal/mol).

	Penta-molecular		Glu_{217}-His_{389}		Thr_{215}-DPG		Lys_{519}-DPG		Thr_{215}-Lys_{519}	
Complexes	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA
E_C/E_{MTP}	−194.4	−194.1	−27.5	−27.6	−19.4	−17.4	−154.9	−150.6	−14.1	−16.4
E_{exch}/E_{rep}	69.7	66.4	14.8	14.1	33.4	32.7	3.2	2.1	15.6	15.2
E_1	−124.7	−127.7	−12.7	−13.5	14.0	15.3	−151.7	−148.5	1.4	−1.2
E_{pol}	−36.5	−34.7	−8.9	−9.4	−15.7	−14.6	−9.4	−8.9	−7.4	−7.3
E_{ct}	−9.7	−7.7	−2.5	−2.2	−3.7	−2.3	−0.7	−0.6	−2.1	−0.8
E_2	−46.2	−42.4	−11.4	−11.6	−19.4	−16.9	−10.1	−9.5	−9.5	−8.1
ΔE(RVS)/ΔE(SIBFA)	−170.9	−170.0	−24.1	−25.1	−5.4	−1.5	−161.8	−158.0	−8.1	−9.3

Table 6. Values (kcal/mol) of the SAPT and SIBFA exchange-dispersion and total dispersion contributions at equilibrium distances for the water complexes of (a) dianionic phosphate; (b) *N*-methylformamide; (c) methanol.

	Phosphate		NMF		Methanol	
	(d_{O-H} = 1.60 Å)		(d_{O-H} = 2.00 Å)		(d_{OH} = 1.90 Å)	
Complexes	QC	SIBFA	QC	SIBFA	QC	SIBFA
$E_{exch-disp}$	1.2	1.1	0.3	0.4	0.3	0.4
E_{disp}	−4.5	−5.1	−2.1	−2.0	−2.3	−2.1

to evaluate its present accuracy, pending ongoing refinements. We thus present below first the calibration for E_{disp} of the effective radii of O atoms of DPG, and those of the hydrogen of the amide bond on the one hand, and of the hydroxyl group on the other. We then compare the values of ΔE_{tot}(SIBFA) to correlated ΔE(QC) values in the binary complexes of DPG with NMF_{211}, NMF_{212}, Ser_{210}, Thr_{212}, and W_1. These comparisons are resumed for the complexes of the DPG-based biosensor with the same sites as above and with the complete recognition site.

Calibration. DPG was probed by a water molecule donating a proton to one anionic oxygen, the PH and OH bonds being collinear, and the O—H distance varying in the range of 1.60–2.00 Å. NMF and methanol had their amide and hydroxyl H atoms probed by a water molecule approaching the NH and OH bond, respectively, along its external bisector, the H—O distance varying in the range of 1.70–2.20 Å. Table 6 lists at equilibrium distances the SAPT values of E_{disp} and $E_{exch-disp}$ and the SIBFA values resulting for this calibration, indicating a satisfactory match. Lesser correspondences between SIBFA and SAPT E_1 and E_2 values would have to be expected, since the present SIBFA computations resort to uncorrelated multipoles and polarizabilities. A new library of fragments with correlated multipoles and polarizabilities is presently being constructed (for preliminary results, see Refs. [67–69]. The results of the following sections will serve as additional validations when this is achieved. Nevertheless, the comparisons reported below on the smaller subsystems should be instructive prior to considering

Table 7. Correlated QC and SIBFA intermolecular interaction energies and their contributions of diphosphate with: (a) the main-chain NMF entities of Lys_{211} and Thr_{212}; (b) the side chains of Ser_{210}, Thr_{212}, and structural water W_1.

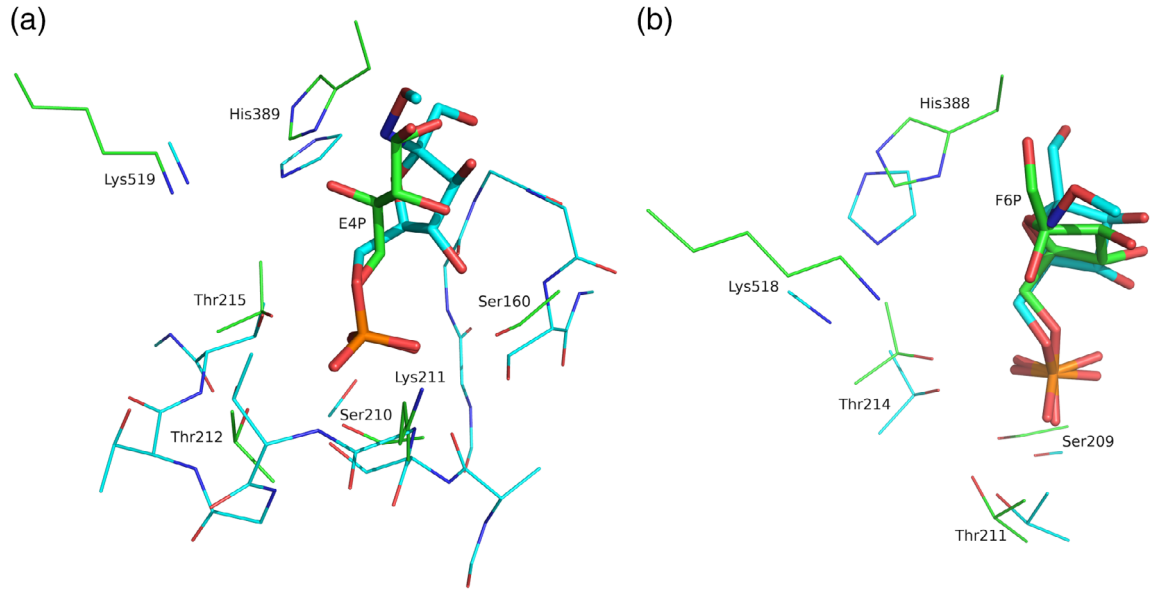
	a								b	
	NMF_{211}		NMF_{212}		Ser_{210}		Thr_{212}		W_1	
Complexes	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA	QC	SIBFA
E_C/E_{MTP}	−37.0	−40.6	−40.8	−46.0	−10.1	−11.8	−23.3	−24.0	−12.6	−13.0
E_{exch}/E_{rep}	31.9	30.7	31.9	30.7	6.6	8.6	15.7	12.3	16.0	13.8
E_1	−5.1	−9.8	−8.9	−11.3	−3.5	−3.2	−7.6	−11.7	3.4	0.9
E_2	−24.0	−24.2	−25.4	−22.0	−8.6	−10.0	−14.8	−12.6	−8.1	−8.8
$E_{exch-disp}$	1.9	1.4	1.6	1.7	0.5	0.4	0.8	0.6	0.7	0.5
$E_{disptot}$	−8.0	−9.1	−6.5	−9.1	−3.5	−4.7	−4.3	−4.6	−3.5	−3.3
ΔE(SAPT)	−37.1	−43.2	−40.8	−42.3	−15.6	−17.9	−26.7	−28.9	−8.2	−11.3
ΔE(DFT/B3LYP-D3)	−39.0		−43.2		−17.1		−29.7		−9.4	
	(−33.5)		(−35.9)		(−12.4)		(−22.3)		(−9.1)	
ΔE(DFT/B97-D3)	−36.4		−43.2				−26.7		−11.6	
	(−26.0)		(−30.5)				(−13.2)		(−6.0)	
ΔE(MP2)	−39.6		−43.6		−16.2		−29.0		−9.9	
	(−36.5)		(−40.1)		(−14.3)		(−26.6)		(−8.5)	

Values in parentheses are the BSSE-corrected ones.

Table 8. Values of correlated ΔE(QC) and ΔE(SIBFA) in two complexes extracted from the recognition site.

	B97-D3			B3LYP-D3		
	aug-cc-pVTZ	cc-pVTZ	SIBFA	aug-cc-pVTZ	cc-pVTZ	SIBFA
a	−94.6	−114.9 (−92.8)	−106.7	−100.6	−116.8 (−97.8)	−106.7
b	−201.7	−227.9	−218.3	−210.5	−231.7 (−201.9)	−218.3
c	−180.9	−197.5	−180.6	−177.6	−199.0	−180.6
d	−119.6	−129.5 (−121.7)	−126.8	−123.9	−133.3 (−125.8)	−126.8
e	−91.4	−94.7 (−92.0)	−91.0	−92.1	−95.2 (−92.5)	−91.0
f		**−280.5**	**−288.2**		**−303.4 (−275.5)**	**−288.2**
g		**−363.0**	**−359.0**		**−396.6 (−348.9)**	**−359.0**

a) dianionic phosphate-NMF_{212}-Thr_{212}-Ser_{210}-Thr_{215}-W_1; b) dianionic phosphate-NMF_{211}-Thr_{215}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; c) dianionic phosphate-NMF_{211}-Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; d) Glu_{217}-His_{389}-Lys_{519}-W_1-W_5; e) Glu_{217}-His_{389}-Lys_{519}; f) complete binding site without structural waters; g) Complete binding site with structural waters. The QC values under parentheses are the BSSE-corrected ones. Some BSSE-corrected QC computations were not converged.



Figures 8. a and b) Superimposition of the energy-minimized complex depicted in blue with: a) the complex of linear inhibitor erythrose-4-phosphate (E4P) with human PGI/AMF dimer (PDB 1IRI)[21] depicted in green; b) the complex of cyclic substrate fructose-6-phosphate (F6P) with rabbit muscle PGI dimer (PDB 1HOX)[19] depicted in green. [Color figure can be viewed at wileyonlinelibrary.com]

the entirety of the recognition site and could be considered as an upper bound of the error with respect to correlated ΔE(QC).

Binary complexes. Table 7 bears on the five binary complexes of DPG with the ligands reported in Table 4 at the uncorrelated level. It compares the values of ΔE_{tot}(SIBFA) to the correlated QC ones, namely ΔE(SAPT), ΔE(B3LYP-D3), ΔE(B97-D3), and ΔE (MP2). For all but the Ser_{210} complex, E_1(SAPT) has smaller magnitudes than E_1(SIBFA) but also than E_1(RVS). The latter feature stems more from a larger magnitude of E_X(SAPT) than E_X(RVS) value than from a correspondingly larger magnitude of E_C(SAPT), which was not anticipated. On the other hand, E_2(SAPT) differs by only small amounts from E_2(RVS). ΔE_{tot}(SIBFA) has a satisfactory agreement with ΔE(SAPT) regarding the NMF_{212}, Ser_{210} and Thr_{212}. It overestimates ΔE(SAPT) for the NMF_{211} complex by 6.1 kcal/mol out of 40, in marked contrast with the 0.8 kcal/mol difference at the RVS level. The overestimation in the W_1 complex amounts to 3.1 kcal/mol, larger than at the RVS level (2.2 kcal/mol).

ΔE(MP2) has values intermediate between ΔE(SAPT) prior to, and after, removal of the BSSE correction. In the absence of the BSSE correction, both B3LYP-D3 and B97-D3 ΔE values remain close to, and invariably larger than, the SAPT ones. Surprisingly, however, the BSSE correction severely diminishes their magnitudes, which renders the BSSE-corrected ΔE(DFT-D3) values unreliable. Such large magnitudes could translate the need for the phosphate dianion for even more diffuse basis sets than afforded at the aug-cc-pVTZ(−f) level, particularly accented with solely one interacting partner as reported here. In multimolecular complexes, BSSE effects are attenuated as shown below.

Correlated intermolecular interactions in the complete binding site and in subsites. Table 8 compares the results of ΔE_{tot}(SIBFA) to those from B97-D3 and B3LYP-D3 calculations. The QC calculations were done at both aug-cc-pVTZ(−f) and cc-pVTZ(−f) levels, except for the two larger complexes f and g, for which the larger basis set calculations were not converged. For complexes a–e, and with respect to the B3LYP-D3 functional, ΔE_{tot}(SIBFA) has values consistently intermediate between the aug-cc-pVTZ(−f) and cc-pVTZ(−f) basis sets. Its values are also intermediate between the BSSE-uncorrected and corrected B3LYP-D3 cc-pVTZ(−f) ones. With respect to the B97-D3 basis set, they are on the other hand closer to the smaller than to the larger basis set results. For complex f, the complete recognition site without the structural waters, ΔE_{tot}(SIBFA) has values nearly exactly intermediate between the B3LYP-D3 and the B97-D3 ones. It is much closer, however, to the B97-D3 results in the presence of the six waters. More involved comparisons between ΔE_{tot}(SIBFA) and DFT-D3 results should await completion of the calibration of the SIBFA library of fragments with correlated multipoles and polarizabilities, which is presently in progress. It is expected to enable further improved agreements with such calculations, similar to those obtained at the uncorrelated level. It could be instructive for polarizable force-field developers and users to compare for the present DPG complexes, the outcome from other polarizable potentials, whether in their present status or undergoing developments. For this purpose, we provide the coordinates of the complete binding site as Supporting Information S5.

Conclusions and Perspectives

We have in the context of the SIBFA procedure calibrated the DPG, a ubiquitous signaling group of biomolecules. DPG is also an essential component of biosensors destined to detect submicromole blood concentrations of autocrine motility factor-phosphoglucose isomerase (AMF-PGI), an extracellular protein closely related to metastasis formation. For each individual contribution, the DPG-relevant parameters were simultaneously

optimized by a procedure that least-squares minimizes its difference with respect to its QC counterpart. We resorted to a training set of DPG complexes with two probes, Zn(II) and water. Up to 140 complexes were considered, upon generating distance and in- and out-of-plane angular variations of the probe approaching each of the three O atoms in turn. Close reproductions of the four QC(RVS) contributions were enabled, with only few exceptions, which corresponded to high-lying energy complexes. Along these lines, we have also calibrated NMF, the building block of the protein backbone, on account of its close involvement in DPG binding. This was done with Zn(II) and water probes considering in-plane approaches to the CO group and approaches perpendicular to the NMF plane over the C, N, and O atoms. Here also close reproductions of the QC contributions were enabled except in some high-lying energy zones. The H atom of the NMF NH bond and that of the OH bond of methanol as a representative of Ser and Thr side chains were probed by a water molecule approaching these two bonds through its external bisector, the O—HN and O—HO bonds being collinear.

Energy minimization on the complex of the DPG-biosensor with a PGI dimer led to a highly compact structure, with the phosphate group exclusively bound to neutral H-bond donors: two successive NMFs; three hydroxyl groups, namely those of Ser_{210}, Thr_{212} and Thr_{215}, and one water molecule. The network was further extended by Thr_{215}-mediated interactions with water W_2, connecting to Lys_{519}, Glu_{217}, and His_{389}. Figures 8a and 8b give a superimposition of this energy-minimized structure with the X-ray complexes of the human PGI dimer with the linear inhibitor erythrose-4-phosphate (AMF/hPGI-E4P) and of the rabbit muscle PGI dimer with the cyclic substrate fructose-6-phosphate (RmPGI-F6P), denoted 1IRI[21] and 1HOX,[19] respectively, in the PDB. Thus, Figure 8a shows that the phosphate groups can be superimposed together, with the ring oxygen of F6P-NOMe being at 2.0 Å from the nearest hydroxyl group to the phosphate of the linear inhibitor E4P, both being turned toward the His_{389} and Lys_{519} side chains. The other E4P hydroxyl is on the opposite side and this is also the case for the three hydroxyl groups of the DPG-based biosensor.

Figure 8b shows that both F6P and F6P-NOMe can be closely superimposed with both ring oxygens being at 0.7 Å. Respective active site residues of the energy-minimized and crystal structures are also quite close to each other: Nε atoms of Lys_{518} (RmPGI-F6P) and Lys_{519} (AMF/hPGI-F6P-NOMe) are at 2.5 Å, and Cγ of His_{388} (RmPGI-F6P) and corresponding Cγ of His_{389} (AMF/hPGI-F6P-NOMe) are at 1.7 Å. This relatively good structural alignment of both structural and energy-minimized structures shows that AMF-hPGI inhibitor F6P-NOMe, designed as a mimic of the DPG biosensor, is a good mimic of the AMF-hPGI substrate F6P, and it gives the first view of the interaction of the DPG biosensor with the cancer biomarker AMF.

In the energy-minimized structure, with respect to the 1IRI and 1HOX crystal structures,[19,21] a rotation by about 180° of the χ_2 ring of His_{389} took place, enabling its deprotonated N atom, N_ε, to bind to the Lys_{519} ammonium group. However, such an arrangement could correspond solely to a local minimum. Its occurrence indicates that in the course of long-duration MD on the complexes of DPG-based sensors with the PGI dimer, the search for possible concerted and transient His_{389} χ_2 rotations/Lys_{519} side-chain elongations should be considered. In the present study, we have considered only one monoprotonated form of His_{389}, namely with the N_δ nitrogen protonated. An intriguing possibility, suggested by the present structure, could be that of a proton transfer from the end methylammonium Lys_{519} side chain to N_ε, followed by a second proton transfer, taking place from the imidazolium to the hemiketal O atom of the furanose ring of a substrate bound in the same position as the DPG-based sensor, and leading to its cleavage. This suggests that it would be worth considering transient or rare events identified from MD or even as local minima from preliminary EM explorations as starting points for QM/MM studies of PGI-catalyzed cleavage.

Propka 3.0[24] calculations indicate that His_{389} in the DPG-PGI complex (PDB 1IRI)[21] is in its neutral form, with one proton on either N_δ or N_ε. We have thus also considered the other His_{389} tautomer having N_ε instead of N_δ protonated. The X-ray structure of the complex shows that the stabilization of this tautomer is enabled by a complex with Glu_{217}, in which N_ε acts as a proton donor to it, while in the deprotonated N_ε His_{389} tautomer on which we focused in the present study, N_ε acts as an electron acceptor from Lys_{519}. This dual character of His N atoms is noteworthy. Thus, the present results could imply that a tautomeric change in the imidazole ring, even though localized on the sole ring, could trigger large-amplitude motions of some side chains, such as those of Lys, Asp/Glu residues, enabling them to act either as proton donors to this ring, or as proton acceptors from it. Such conformational changes could be further assisted by main chain rearrangements as well. To our knowledge, no evidence has been put forth so far regarding His tautomerization-induced large amplitude conformational changes on proteins. It would be highly instructive if future long-duration polarizable constant-pH MD simulations could demonstrate an at least transient coexistence of two major conformational states, each corresponding to one tautomer.

We have extracted DPG and the His_{389} side chain from the X-ray structure and retained the closest side chains to them, namely those of Thr_{215} and Lys_{519} for DPG, and of Glu_{217} for His_{389}. ΔE(SIBFA) and its contributions gave a close agreement to their RVS counterparts in this five-fragment complex. Even though such validations are at this stage not as extensive as those reported with the N_δ-protonated tautomer, they imply that SIBFA should be able to handle in a balanced fashion both candidate tautomers in prospective MD simulations.

Long-duration MD simulations of the complexes of the PGI dimer with DPG-based ligands could unravel competing and possibly transient-binding modes. For these to be reliably sampled, it is essential for ΔE to enable a reliable reproduction of ΔE(QC) in diverse multimolecular complexes between DPG and its neighboring PGI residues and structural waters. A balanced reproduction of first- and second-order contributions is necessary for this purpose, as shown by previous studies on the complexes of inhibitors with kinases,[70] Zn-metalloenzymes[30] and cation-stacked guanine tetramer complexes[69] While

generally satisfactory, comparison of ΔE(SIBFA) to ΔE(QC) in bimolecular complexes of DPG with individual PGI residues indicated some shortcomings. They concerned principally E_{rep} for the DPG complexes with both NMF's and E_{pol} in its complexes with NMF_{212} and the Thr_{212} side chain. Thus an important inference from the present work is that in our next stage of development, namely constructing and calibrating a new library of protein building blocks with correlated multipoles and polarizabilities, the calibration of H-bond donating groups of NMF but also of hydroxyl groups, should be carried out not only with neutral electron-donating probes such as water but also with at least one representative anionic probe enhancing the short-range, electrostatic and polarization responses of the probed fragment. This should occur concomitant with the probing of electron-rich sites by a cationic probe and by an H-bond donating one, and, similar to these, not be limited to in-plane binding, but encompass out-of-plane binding as well. The present and prospective additional explorations do not appear to have been reported in the construction of any other polarizable force field. They appear all the more important, as the secondary structures of proteins involve numerous short and long-range NMF out-of-plane, out of equilibrium, interactions, in addition to their, mostly in-plane, H-bonding interactions.

An essential result from this work is related to the close reproduction by ΔE(SIBFA) of ΔE(QC) in the extended binding site, encompassing 193 and 211 atoms without and with the six structural waters, its values being intermediate between the BSSE-uncorrected and corrected ΔE(QC) ones (Table 4). Nevertheless, examination of ΔE values in subsites detected in some cases some shortcomings. Remedying these would require recalibration of other than DPG, NMF, and hydroxyl fragments. This was clearly not the initial purpose of this work. However, such complexes will be reconsidered and used as future, additional validation tests at the correlated ΔE (QC) level, following completion of the new correlated SIBFA library of fragments.

In the course of this study, we have also recalibrated the effective radii for E_{disp} concerning the anionic oxygen of DPG and the H atoms of the NMF NH bond and of the methanol OH bond, in order to match the dispersion contribution and its exchange-dispersion component from SAPT analyses of the complexes of DPG, NMF, and methanol with water. As such this calibration is provisory and will be superseded by the one undertaken with correlated multipoles impacting all E_{disp}-relevant parameters (Naseem-Khan et al., to be published). At this stage, the present calibration has enabled to reconsider all PGI and PGI subsite interactions to now compare ΔE_{tot}(SIBFA) to ΔE(QC) at the correlated MP2, B3LYP-D3 and B97-D3 levels. For both bimolecular and multimolecular complexes, ΔE (QC) could be overall satisfactorily reproduced, with one obvious exception, DPG-NMF_{211} where ΔE_{tot}(SIBFA) was larger than ΔE(SAPT) by 6 kcal/mol out of 40. With the larger multimolecular complexes, ΔE_{tot}(SIBFA) had generally intermediate values between the B3LYP-D3 and the B97-D3 ones, and also, with either functional, intermediate values between the aug-cc-pVTZ(−f) and cc-pVTZ(−f) ones. We consider such results as upper bounds to the errors expectable after we complete the construction and calibration of the correlated library of fragments. Thus, the present results should enable us in a short future to evaluate whether these remaining errors could have been effectively narrowed down. This could enable reliable long-duration polarizable MD simulations on the complexes of DPG-based sensors with the PGI dimer, the impact of complexation on local transient and/or large-scale PGI conformation changes, and global hinge-bending motions of the dimer. Such simulations could also assist in the search for novel DPG derivatives with enhanced PGI-binding affinities and biosensors with enhanced sensitivities.

Acknowledgments

The authors wish to thank the Grand Equipement National de Calcul Intensif (GENCI), Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), Centre Informatique de l'Enseignement Supérieur (CINES, France, project No. x2009-075009), and the Centre Régional Informatique et d'Applications numériques de Normandie (CRIANN, Rouen, France, project 1998053).

Keywords: dianionic phosphate · phosphoglucose isomerase dimer · biosensor · ab initio quantum chemistry · polarizable molecular mechanics

How to cite this article: M. Devillers, J.-P. Piquemal, L. Salmon, N. Gresh. *J. Comput. Chem.* **2020**, 9999, 1–16. DOI: 10.1002/jcc.26134

Additional Supporting Information may be found in the online version of this article.

[1] H. R. Bourne, D. A. Sanders, F. McCormick, *Nature* **1990**, *348*, 125. https://doi.org/10.1038/348125a0.
[2] R. L. Veech, M. Todd King, R. Pawlosky, Y. Kashiwaya, P. C. Bradshaw, W. Curtis, *IUBMB Life* **2019**, *71*, 565. https://doi.org/doi.org/10.1002/iub.1997.
[3] F. Jordan, N. S. Nemeria, *Bioorg. Chem.* **2014**, *57*, 251. https://doi.org/10.1016/j.bioorg.2014.08.002.
[4] J. Liang, Q. Han, Y. Tan, H. Ding, J. Li, *Front. Mol. Biosci.* **2019**, *6*, 4. https://doi.org/10.3389/fmolb.2019.00004.
[5] S. J. Humphrey, D. E. James, M. Mann, *Trends Endocrinol. Metab.* **2015**, *26*, 676. https://doi.org/10.1016/j.tem.2015.09.013.
[6] S. T. Chung, S. K. Chacko, A. L. Sunehag, M. W. Haymond, *Diabetes* **2015**, *64*, 3996. https://doi.org/10.2337/db15-0640.
[7] H.-S. Han, G. Kang, J. S. Kim, B. H. Choi, S.-H. Koo, *Exp. Mol. Med.* **2016**, *48*, e218. https://doi.org/10.1038/emm.2015.122.
[8] A. Stincone, A. Prigione, T. Cramer, M. M. C. Wamelink, K. Campbell, E. Cheung, V. Olin-Sandoval, N.-M. Grüning, A. Krüger, M. Tauqeer Alam, M. A. Keller, M. Breitenbach, K. M. Brindle, J. D. Rabinowitz, M. Ralser, *Biol. Rev. Camb. Philos. Soc.* **2015**, *90*, 927. https://doi.org/10.1111/brv.12140.
[9] M. Akram, S. M. A. Shah, N. Munir, M. Daniyal, I. M. Tahir, Z. Mahmood, M. Irshad, M. Akhlaq, S. Sultana, R. Zainab, *J. Cell. Physiol.* **2019**, *234*, 14473. https://doi.org/10.1002/jcp.28228.
[10] A. E. I. Proudfoot, M. A. Payton, T. N. C. Wells, *J. Protein Chem.* **1994**, *13*, 619.
[11] A. Cleasby, A. Wonacott, T. Skarzynski, R. E. Hubbard, G. J. Davies, A. E. I. Proudfoot, A. R. Bernard, M. A. Payton, T. N. C. Wells, *Nat. Struct. Biol.* **1996**, *3*, 470.

[12] L. Ahmad, S. Plancqueel, V. Dubosclard, N. Lazar, W. Ghattas, I. Li de la Sierra-Gallay, H. van Tilbeurgh, L. Salmon, *FEBS Lett* **2018**, *592*, 1667. https://doi.org/10.1002/1873-3468.13059.
[13] H. Watanabe, K. Takehana, M. Date, T. Shinozaki, A. Raz, *Cancer Res.* **1996**, *56*, 2960.
[14] M. Baumann, A. Kappl, T. Lang, K. Brand, W. Siegfried, E. Paterok, *Cancer Invest.* **1990**, *8*, 351. https://doi.org/10.3109/07357909009012053.
[15] M. Devillers, L. Ahmad, H. Korri-Youssoufi, L. Salmon, *Biosens. Bioelectron.* **2017**, *96*, 178. https://doi.org/10.1016/j.bios.2017.04.031.
[16] C. J. Jeffery, B. J. Bahnson, W. Chien, D. Ringe, G. A. Petsko, *Biochemistry* **2000**, *39*, 955. https://doi.org/10.1021/bi991604m.
[17] C.-C. Chou, Y. J. Sun, M. Meng, C.-D. Hsiao, *J. Biol. Chem.* **2000**, *275*, 23154. https://doi.org/10.1074/jbc.M002017200.
[18] C. J. Jeffery, R. Hardré, L. Salmon, *Biochemistry* **2001**, *40*, 1560. https://doi.org/10.1021/bi0018483.
[19] J. H. Lee, K. Z. Chang, V. Patel, C. J. Jeffery, *Biochemistry* **2001**, *40*, 7799. https://doi.org/10.1021/bi002916o.
[20] D. Arsenieva, R. Hardré, L. Salmon, C. J. Jeffery, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5872. https://doi.org/10.1073/pnas.052131799.
[21] N. Tanaka, A. Haga, H. Uemura, H. Akiyama, T. Funasaka, H. Nagase, A. Raz, K. T. Nakamura, *J. Mol. Biol.* **2002**, *318*, 985. https://doi.org/10.1016/S0022-2836.
[22] C. Davies, H. Muirhead, J. Chirgwin, *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 1111. https://doi.org/10.1107/S0907444903007352.
[23] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 2284. https://doi.org/10.1021/ct200133y.
[24] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525. https://doi.org/10.1021/ct100578z.
[25] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Res.* **2004**, *32*, W665. https://doi.org/10.1093/nar/gkh381.
[26] G. Martínez-Rosell, T. Giorgino, G. De Fabritiis, *J. Chem. Inf. Model.* **2017**, *57*, 1511. https://doi.org/10.1021/acs.jcim.7b00190.
[27] C. Roux, N. Gresh, L. E. Perera, J.-P. Piquemal, L. Salmon, *J. Comput. Chem.* **2007**, *28*, 938. https://doi.org/10.1002/jcc.20586.
[28] C. Roux, F. Bhatt, J. Foret, B. de Courcy, N. Gresh, J.-P. Piquemal, C. J. Jeffery, L. Salmon, *Proteins* **2011**, *79*, 203. https://doi.org/10.1002/prot.22873.
[29] N. Gresh, B. de Courcy, J.-P. Piquemal, J. Foret, S. Courtiol-Legourd, L. Salmon, *J. Phys. Chem. B* **2011**, *115*, 8304. https://doi.org/10.1021/jp2024654.
[30] N. Gresh, D. Perahia, B. de Courcy, J. Foret, C. Roux, L. El-Khoury, J.-P. Piquemal, L. Salmon, *J. Comput. Chem.* **2016**, *37*, 2770. https://doi.org/10.1002/jcc.24503.
[31] N. Gresh, G. A. Cisneros, T. A. Darden, J.-P. Piquemal, *J. Chem. Theory Comput.* **2007**, *3*, 1960. https://doi.org/10.1021/ct700134r.
[32] L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Y. Ren, J. W. Ponder, J. P. Piquemal, *Chem. Sci.* **2018**, *9*, 956. https://doi.org/10.1039/C7SC04531J.
[33] N. Gresh, G.-B. Shi, *J. Comput. Chem.* **2004**, *25*, 160. https://doi.org/10.1002/jcc.10312.
[34] N. Gresh, N. Audiffren, J.-P. Piquemal, J. de Ruyck, M. Ledecq, J. Wouters, *J. Phys. Chem. B* **2010**, *114*, 4884. https://doi.org/10.1021/jp907629k.
[35] B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **1994**, *94*, 1887. https://doi.org/10.1021/cr00031a008.
[36] A. J. Misquitta, B. Jeziorski, K. Szalewicz, *Phys. Rev. Lett.* **2003**, *91*, 033201. https://doi.org/10.1103/PhysRevLett.91.033201.
[37] W. J. Stevens, W. H. Fink, *Chem. Phys. Lett.* **1987**, *139*, 15. https://doi.org/10.1016/0009-2614.
[38] G. Tiraboschi, B.-P. Roques, N. Gresh, *J. Comput. Chem.* **1999**, *20*, 1379.
[39] G. Tiraboschi, N. Gresh, C. Giessner-Prettre, L. G. Pedersen, D. W. Deerfield, *J. Comput. Chem.* **2000**, *21*, 1011.
[40] L. El Khoury, S. Naseem-Khan, K. Kwapien, Z. Hobaika, R. G. Maroun, J.-P. Piquemal, N. Gresh, *J. Comput. Chem.* **2017**, *38*, 1897.
[41] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007.
[42] D. Feller, *J. Comput. Chem.* **1996**, *17*, 1571.
[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., Gaussian 09, Revision A.1, Gaussian Inc., Wallingford, CT, **2009**.
[44] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347. https://doi.org/10.1002/jcc.540141112.
[45] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, et al., *J. Chem. Theory Comput.* **2017**, *13*, 3185. https://doi.org/10.1021/acs.jctc.7b00174.
[46] J. Langlet, J. Caillet, J. Bergès, P. Reinhardt, *J. Chem. Phys.* **2003**, *118*, 6157. https://doi.org/10.1063/1.1558473.
[47] K. Patkowski, K. Szalewicz, B. Jeziorski, *J. Chem. Phys.* **2006**, *125*, 154107. https://doi.org/10.1063/1.2358353.
[48] L. Goerigk, S. Grimme, *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670. https://doi.org/10.1039/c0cp02984j.
[49] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553. https://doi.org/10.1080/00268977000101561.
[50] A. J. Stone, *J. Phys. Chem. A* **2011**, *115*, 7017. https://doi.org/10.1021/jp112251z.
[51] D. R. Garmer, W. J. Stevens, *J. Phys. Chem. A* **1989**, *93*, 8263. https://doi.org/10.1021/j100362a023.
[52] The HSL. Mathematical Software Library; Science and Technology Facilities Council. http://www.hsl.rl.ac.uk/.
[53] J.-P. Piquemal, N. Gresh, C. Giessner-Prettre, *J. Phys. Chem. A* **2003**, *107*, 10353. https://doi.org/10.1021/jp035748t.
[54] N. Gresh, *J. Comput. Chem.* **1995**, *16*, 856. https://doi.org/10.1002/jcc.540160705.
[55] G. A. Evangelakis, J. P. Rizos, I. E. Lagaris, I. N. Demetropoulos, *Comput. Phys. Commun.* **1987**, *46*, 401.
[56] A. M. Baptista, P. J. Martel, S. B. Petersen, *Proteins* **1997**, *27*, 523.
[57] J. Lee, B. T. Miller, A. Damjanović, B. R. Brooks, *J Chem Theory Comput* **2014**, *10*, 2738. https://doi.org/10.1021/ct500175m.
[58] B. K. Radak, C. Chipot, D. Suh, S. Jo, W. Jiang, J. C. Phillips, K. Schulten, B. Roux, *J Chem Theory Comput* **2017**, *13*, 5933. https://doi.org/10.1021/acs.jctc.7b00875.
[59] F. L. Barroso da Silva, L. G. Dias, *Biophys. Rev* **2017**, *9*, 699. https://doi.org/10.1007/s12551-017-0311-5.
[60] F. L. Barroso da Silva, F. Sterpone, P. Derreumaux, *J. Chem. Theory Comput.* **2019**, *15*, 3875. https://doi.org/10.1021/acs.jctc.9b00202.
[61] M. Devereux, N. Gresh, J.-P. Piquemal, M. Meuwly, *J. Comput. Chem.* **2014**, *35*, 1577. https://doi.org/10.1002/jcc.23661.
[62] G. Corongiu, E. Clementi, *J. Chem. Phys.* **1978**, *69*, 4885. https://doi.org/10.1063/1.436518.
[63] N. Gresh, D. R. Garmer, *J. Comput. Chem.* **1996**, *17*, 1481.
[64] N. Gresh, S. A. Kafafi, J.-F. Truchon, D. R. Salahub, *J. Comput. Chem.* **2004**, *25*, 823. https://doi.org/10.1002/jcc.20012.
[65] A. Stone, The Theory of Intermolecular Forces, 2nd ed., Oxford University Press, Oxford, New York, **2013**.
[66] S. Creuzet, J. Langlet, N. Gresh, *J. Chim. Phys.* **1991**, *88*, 2399. https://doi.org/10.1051/jcp/1991882399.
[67] K. Kwapien, M. Damergi, S. Nader, L. El Khoury, Z. Hobaika, R. G. Maroun, J.-P. Piquemal, L. Gavara, D. Berthomieu, J.-F. Hernandez, N. Gresh, *J. Phys. Chem. B* **2017**, *121*, 6295. https://doi.org/10.1021/acs.jpcb.7b01053.
[68] N. Gresh, J. E. Sponer, M. Devereux, K. Gkionis, B. de Courcy, J.-P. Piquemal, J. Sponer, *J. Phys. Chem. B* **2015**, *119*, 9477. https://doi.org/10.1021/acs.jpcb.5b01695.
[69] N. Gresh, S. Naseem-Khan, L. Lagardère, J.-P. Piquemal, J. E. Sponer, J. Sponer, *J. Phys. Chem. B* **2017**, *121*, 3997. https://doi.org/10.1021/acs.jpcb.7b01836.
[70] B. de Courcy, J.-P. Piquemal, C. Garbay, N. Gresh, *J. Am. Chem. Soc.* **2010**, *132*, 3312. https://doi.org/10.1021/ja9059156.

Received: 23 August 2019
Revised: 28 November 2019
Accepted: 5 December 2019