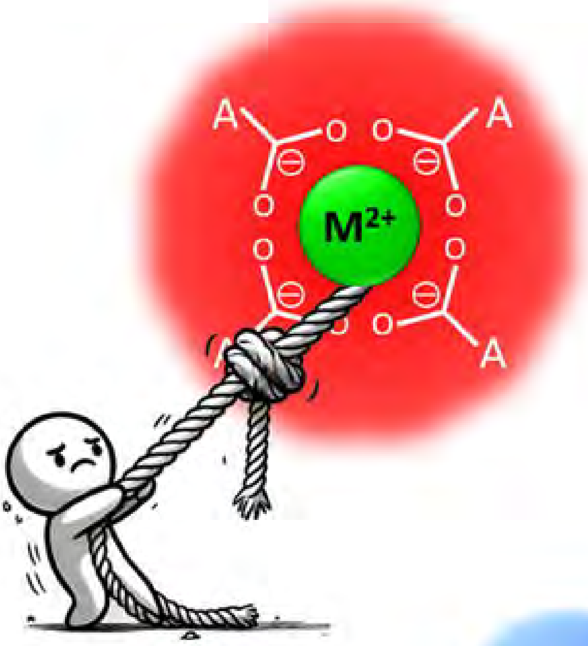
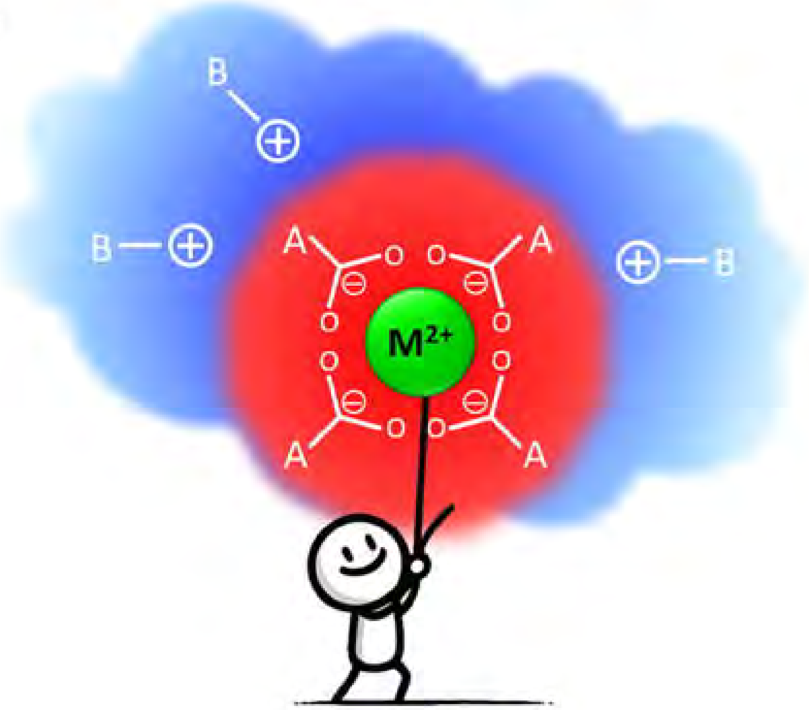
PROTEIN SCIENCE

A PUBLICATION OF THE PROTEIN SOCIETY
www.proteinscience.org Vol 33 No 12 December 2024



4A



4A/3B

WILEY

Received: 27 June 2024 | Revised: 16 October 2024 | Accepted: 17 October 2024

DOI: 10.1002/pro.5206

RESEARCH ARTICLE

Why is binding of a divalent metal cation to a structural motif containing four carboxylate residues not accompanied by a conformational change?

Sofya Lushchekina[1] | **Lev Weiner**[2,3] | **Yacov Ashani**[1] | **Reeki Emrizal**[4] | **Mohd Firdaus-Raih**[4,5] | **Israel Silman**[2] | **Joel L. Sussman**[6,7]

[1]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

[2]Department of Brain Sciences, Weizmann Institute of Science, Rehovot, Israel

[3]Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel

[4]Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

[5]Institute of Systems Biology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

[6]Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot, Israel

[7]Structural Proteomics Unit, Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel

Correspondence

Sofya Lushchekina, Department of Brain Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel.
Email: sofya.lushchekina@weizmann.ac.il

Joel L. Sussman, Department of Chemical and Structural Biology, The Weizmann Institute of Science, Rehovot 7610001, Israel.
Email: joel.sussman@weizmann.ac.il

Funding information

Instruct-ERIC; Helen and Milton A. Kimmelman Center for Biomolecular Structure & Assembly; Weizmann Institute of Science; Horizon 2020 Framework Programme, Grant/Award Number: 824087; Ministry of Higher Education, Malaysia, Grant/Award Number: TRGS/1/2022/UKM/01/9/1

Review Editor: Lynn Kamerlin

Abstract

We earlier showed that *Torpedo californica* acetylcholinesterase (AChE) contains a cluster of four conserved aspartates that can strongly bind divalent cations, which we named the 4D motif. Binding of the divalent metal cations greatly increases its thermal stability. Here we systematically examined all available crystallographic structures of *T. californica* AChE. Two additional metal-binding sites were identified, both composed of acidic and histidine residues. Relative binding to the 4D and additional sites was studied using metadynamics simulations. It was observed that in crystal structures devoid of metal ions in the 4D site, the conformation of *T. californica* AChE is almost identical to that in structures in which it is occupied by a divalent metal ion. Closer examination of the 4D motif reveals that three of the four *acidic* residues form ion pairs with conserved *basic* residues surrounding them. We named this new motif the 4A/3B motif. Molecular dynamics with quantum potential simulations was used to quantify the 4D motif's binding strength compared with that of the metal-binding site in the protein fXIIIa, which consists of four aspartates, but is devoid of adjacent cationic residues. Whereas fXIIIa's 4D site, in the absence of a metal cation, expanded significantly in the simulation, that of *Torpedo* AChE displayed only minor periodic changes in size. Furthermore, the energy of metal ion unbinding from the two sites differs by ca. 10 kcal/mol. We identified several other proteins in the PDB that contain the 4A/3B motif, whose conformations are identical in the presence or absence of a metal ion.

https://doi.org/10.1002/pro.5206

An animated Interactive 3D Complement (I3DC) is available in Proteopedia at https://proteopedia.org/w/Journal:Protein_Science:4.

KEYWORDS

acetylcholinesterase, divalent metal cation, metadynamics, metal-binding site, molecular dynamics, QM/MM, structural motif

1 | INTRODUCTION

Acetylcholinesterase (AChE) is a powerful enzyme that hydrolyses the neurotransmitter acetylcholine at a rate approaching diffusion control (Bazelyansky et al., 1986; Rosenberry, 1975), in accordance with its principal biological role in terminating transmission at cholinergic synapses (Silman, 2021). Upon thermal denaturation, *Torpedo californica* AChE (*Tc*AChE) unfolds irreversibly from its native state to a molten globule in a two-state transition (Kreimer, Dolginova, et al., 1994; Kreimer, Szosenfogel, et al., 1994). This thermal transition is retarded in the presence of the divalent cations, Ca^{2+}, Mg^{2+}, and Mn^{2+} (Millard et al., 2003). It was revealed by the use of electron paramagnetic resonance that the effects observed could be ascribed to a specific binding site for the divalent metal ions and to several weak binding sites (Silman et al., 2021). Indeed, a novel motif of four aspartate residues, which we called the 4D motif, was subsequently identified in *Tc*AChE (Silman et al., 2021). It has been found to be present in numerous sequences of both AChE and butyrylcholinesterase (BChE) (Silman et al., 2021).

The 4D motif consists of residues D326, D389, D392, and D393 (*Tc*AChE numbering). Thus, it differs from well-characterized metal-cation binding structural motifs, such as DxDxDG (Rigden et al., 2011). Furthermore, it differs from both the EF-hand and calcium blade domains (Permyakov, 2021; Valdez et al., 2014), from active sites of metalloenzymes that include His residues in addition to Asp and/or Glu residues (Valdez et al., 2014), and from other metal-binding structural elements (Permyakov, 2021).

Although the first resolved crystal structure of *Tc*AChE contained a uranyl cation in the 4D site, which was used as a heavy atom to solve the phase problem (PDB ID 7B2W), it was only recently added to the PDB (Silman et al., 2021; Sussman et al., 1991). In the high-resolution *Tc*AChE apo-structure, 1EA5, no metal ions were observed (Dvir et al., 2002). However, the crystal structures of a number of *Tc*AChE complexes reveal the presence of metal cations both within the 4D motif and at other sites. These include the crystal structures of the complexes obtained by soaking $(UO_2)^{2+}$, Mg^{2+}, and Ca^{2+} into native crystals (Silman et al., 2021), PDB IDs 7B2W, 7B38, and 7B8E, respectively, solved to 2.65, 1.85, and 2.23 Å resolution.

Comparison of apoenzyme crystal structures of *Tc*AChE with those of structures containing divalent metal ions revealed that their 4D motifs overlap very closely. This was a rather surprising observation since one might expect that, due to the electrostatic repulsion between the four negatively charged Asps, the crystal structure of the apoenzyme would differ significantly from those of the divalent metal ion complexes. This prompted us to perform a closer inspection of the 4D motif to identify the molecular basis for this anomalous behavior.

Molecular dynamics with quantum potentials (QM/MM-MD) was the method of choice for assessing the global protein and local 4D site dynamics, and the energetics of metal binding/unbinding from the site, taking into account polarization effects and possible water dissociation events.

Analyzing available crystallographic data for *Tc*AChE, we observed additional metal binding sites which, in most cases, were located at a crystal packing interface. This raised two questions: are they crystallographic artifacts or do they also exist in solution? And, if so, how do they compare with the 4D site, the primary site of our attention? We explored these questions using the metadynamics (MTD) technique.

2 | METHODS

2.1 | Structural analysis

Analysis of the 3D structures and electron density maps deposited in the PDB (https://www.rcsb.org/) (Berman et al., 2000) was performed with the PDB-REDO webserver (https://pdb-redo.eu/) (Joosten et al., 2014; van Beusekom et al., 2018), the CheckMyMetal web server (Gucwa et al., 2023; Zheng et al., 2017), Phenix 1.21 (Liebschner et al., 2019), and Coot 0.8.9 (Emsley et al., 2010).

The FindGeo web-tool (https://metalweb.cerm.unifi.it/tools/findgeo/) (Andreini et al., 2012) was used to confirm the mode of coordination of protein-bound metal ions.

The web-based tools and programs, ASSAM (http://mfrlab.org/grafss/assam/) (Nadzirin et al., 2012), IMAAAGINE (http://mfrlab.org/grafss/imaaagine) (Nadzirin et al., 2013), pyScoMotif (https://github.com/3BioCompBio/pyScoMotif) (Cia et al., 2023), and the PDB structural search (https://www.rcsb.org/search/advanced) (Guzenko et al., 2020) were used to search for structural motifs.

For searches using ASSAM, pyScoMotif, and the PDB structural search two templates were used: the *Tc*AChE 4D motif, viz., the side chains of D326, D389, D392, D393, and the 4D motif together with the side chains of the neighboring cationic residues, K325, R388, R517 from the 7B8E PDB structure (Silman et al., 2021). For the template-free searches using IMAAAGINE, the query contained 4 acidic and 3 basic residues, excluding histidine. A cutoff of 5 Å separating the acidic and basic residues was imposed, since ionic pairs rather than salt bridges were sought (Kumar & Nussinov, 2002).

Analysis of retrieved structures was performed with PyMol (The PyMOL Molecular Graphics System, Version 2.5.8 Schrödinger, Inc.) and VMD (Hsin et al., 2008).

Sequences of AChEs from various species were retrieved from UniProt (UniProt, 2023) and aligned with Clustal Omega (Sievers et al., 2011).

2.2 | Molecular modeling

2.2.1 | System preparation

For the *Tc*AChE protein model, the 7B8E crystal structure (Silman et al., 2021) was used, and for the human fXIIIa protein model, the 4KTY crystal structure (Stieler et al., 2013). Both structures were treated with the Charmm-GUI (Brooks et al., 2009; Jo et al., 2008) PDB manipulator (Park et al., 2023) to restore missing residues and side chains. The crystallographic water molecules and ions were retained, and protonation of the ionizable residues was assigned according to the PROPKA3 (Olsson et al., 2011) predictions via the VEGA (Pedretti et al., 2021) online service (https://www.ddl.unimi.it/vegaol/propka.htm).

The systems were solvated with TIP3P water molecules within a 120 Å × 120 Å × 120 Å cubic simulation box. The distance from the protein to the walls of the box was more than 10 Å, regardless of the protein's orientation. The systems were neutralized by adding Na^+ and Cl^- ions to a NaCl concentration of 150 mM.

To explore the effect of a metal cation on the stability of the 4D sites, two additional systems were generated, in which the crystallographic Ca^{2+} ion was manually replaced by a single water molecule, and 2 Na^+ ions were added on the periphery of the solvent box to keep the whole system electroneutral. Otherwise, for both proteins, the systems with and without Ca^{2+} were identical.

The systems were energy minimized using 3000 steps, and the solvent box was equilibrated for 1 ns, using classical molecular dynamics (MD), with protein, crystallographic metal ions, and water molecules fixed (NPT ensemble, $p = 1$ atm, $T = 300$ K). Classical MD simulations were performed using NAMD3.0b2 (Phillips et al., 2020) and the CHARMM36 force field (Best et al., 2012), with an integration time step of 1 fs. The resulting systems were used for subsequent QM/MM-MD simulations.

2.2.2 | QM/MM-MD simulations

All calculations using QM/MM-MD were performed using a combination of the TeraChem and NAMD programs (Melo et al., 2018). The TeraChem v1.96H program (Luehr et al., 2015) (PetaChem, LLC, www.petachem.com) was used to calculate forces in the quantum mechanical (QM) region, and NAMD 2.14 (Phillips et al., 2020) in the molecular mechanical (MM) part and in the MD steps. The QM subsystem was treated at the DFT level with the PBE0 functional (Adamo & Barone, 1999), and an all-electron 6-31G** basis set was used for all atoms of the subsystem. The D3 empirical dispersion correction (Grimme et al., 2010) was utilized. The cutoff distance for point charges of the MM subsystem contributing to the QM Hamiltonian was 12 Å. The MD steps were performed as described above.

All systems were initially equilibrated during 1 ps of QM/MM-MD simulation. For each system, three parallel unconstrained production runs were performed during 50 ps. Furthermore, to check the consistency of the unbiased MD simulations and the influence of the choice of the QM subsystem on the results of simulations, three QM subsystems of different sizes were used (referred to as small, big, and medium) for these runs.

For the *Tc*AChE system, the QM subsystems consisted of:

1. The Ca^{2+} ion, present in the initial crystallographic structure for all subsystems.
2. The side chains of Asp326, Asp389, Asp392, and Asp393 (the link between the MM and QM was between the C^{α} and C^{β} atoms of these residues) for all subsystems.
3. The Lys325 side chain (the link between the MM and QM regions was between its C^{ε} and C^{δ} atoms) for the

medium and big subsystems, and additionally the Arg388 and Arg517 side chain (the link between the MM and QM regions was between their C^{δ} and C^{γ} atoms) for the big subsystem.

4. Water molecules within 4 Å of the Ca^{2+} ion and within 2.3 Å of the Asp O^{δ} atoms, and the polar H atoms of the cationic residues in the corresponding subsystems.

All three QM subsystems are shown in Figure S5.

In the case of the fXIIIa 4D site, some of the main chain NH groups and the side chains of Asn347 and Gln349 interact with carboxylates of the 4D motif. For this reason, and to avoid splitting a peptide bond between the QM and MM subsystems, we included in the QM subsystem the following groups of atoms:

1. The Ca^{2+} ion present in the initial crystallographic structure for all subsystems.
2. The side chains of Asp343, Asp345, Asp351, and Asp367 (QM and MM regions link was between C^{α} and C^{β} atoms) for the small subsystem.
3. Asp345 without the main chain CO-group (QM and MM regions link between C^{α} and main chain C atoms), and the main chain CO group of Asn344 (QM and MM regions link between C^{α} and main chain C atoms) for the medium subsystem; for the big subsystem, the full Asp345 residue was included.
4. All atoms of Ala346 were included in the big subsystem, and only the main chain CO group in the medium subsystem (QM and MM regions link between C^{α} and main chain C atoms).
5. The full Asn347 residue and NH group of Leu348 (QM and MM regions link between C^{α} and main chain N atoms) were included in the medium and big subsystems.
6. The side chain of Gln349 (MM and QM regions link between C^{α} and C^{β} atoms) was included in the big subsystem.
7. Asp351 without the main chain CO group and the main chain CO group of Met350 (QM and MM regions link between C^{α} and main chain C atoms of both Met350 and Asp351) for the medium and big subsystems.
8. The side chains of Asp343 and Asp367 (QM and MM regions link between C^{α} and C^{β} atoms) for all subsystems.
9. Water molecules within 4 Å of the Ca^{2+} ion and within 2.3 Å of the Asp O^{δ} atoms for the small and medium subsystems. For the big QM subsystem, water molecules within 2.3 Å of the subsystem's protein polar atoms were included.

The full QM subsystem is shown in Figure S6.

Due to the diffusion of water molecules, their list in QM subsystems was updated every 10 ps of the QM/MD simulation. The size of the 4D site was the radius of gyration (R_g) of the group of side chains of its four Asp residues, defined as:

$$R_g = \sqrt{\sum_i m_i \frac{\left\| \vec{r}_i - \vec{r}_c \right\|^2}{\sum_i m_i}},$$

where m_i is the mass of the ith atom, $\vec{r}_i$ is the radius vector of the ith atom, and $\vec{r}_c$ is a radius vector of the centers of masses (COMs) of the group.

2.2.3 | Umbrella sampling calculations

To compare the affinities for Ca^{2+} of the 4D motifs of *Tc*AChE and fXIIIa, umbrella sampling and weighted histogram analysis (WHAM) (Dimelow et al., 2006; Kumar et al., 1992) were used to generate a potential of mean force (PMF) (Doudou et al., 2009; Roux, 1995) for the binding/unbinding process. The pre-equilibrated systems from the previous unconstrained QM/MM-MD calculations were taken as the starting structures. Using steered classical MD simulation, with all atoms of both proteins fixed, the Ca^{2+} ions were moved from their equilibrium positions in the 4D sites to a distance of 15 Å along the vector orthogonal to the protein surface.

Pathways obtained by steered MD were divided into windows separated by 0.25 Å. The reaction coordinate ξ was set as the distance between Ca^{2+} and the COMs of the carboxylate O^{δ} atoms of all four Asps of the 4D sites. As a tradeoff between the computational cost and accuracy, an individual QM subsystem was defined for each window as described for the medium-sized QM subsystems in the previous subsection. Water molecules were selected for inclusion in the QM subsystem independently, based on proximity to the polar atoms of the QM subsystem in the specific snapshot. This ensures that upon separation of the metal cation and the 4D site, each of them maintains a hydration shell. For a separation greater than 14 Å, they form two disjointed QM subsystems.

The state within each window was controlled with an effective harmonic restraining force of 100 kcal/mol·$Å^2$ (fluctuation width 0.1 Å; scaled force constant 1 kcal/mol). These parameters were adjusted in a series of test runs, and additional windows were added to provide good overlapping of the histograms. Prior to umbrella sampling, each window was equilibrated for 100 fs with the above restraints applied. Sampling was performed until convergence (the shape of the PMF did not change), 10 ps for the *Tc*AChE system for each of 40 windows, and 30 ps for the fXIIIa system, due to the more pronounced conformational changes in the latter. To construct the PMF, the WHAM method (Gumbart et al., 2013; Roux, 1995) in the Grossfield implementation (v. 2.0.11, http://membrane.urmc.rochester.edu/content/wham) was used. Standard deviations were estimated using a blocking procedure according to Di Palma et al. (2015).

2.2.4 | Classical MD MTD simulations

The classical MD MTD simulations were performed to compare the binding of metal ions to the different sites on the *Tc*AChE surface. We referred to the previous experimental studies (Silman et al., 2021), and adjusted the system setup to match the experimental conditions as closely as possible. Most importantly, we chose Mg^{2+} rather than Ca^{2+}; the latter was used for the QM/MM-MD studies of the dynamics and energetics of the 4D sites.

We started from the QM/MM-MD pre-equilibrated structure of *Tc*AChE. The protein structure was treated with Charmm-GUI (Brooks et al., 2009; Jo et al., 2008) and solvated with TIP3P water molecules, forming a 100 Å × 100 Å × 100 Å cubic simulation box. To mimic the experimental conditions, the protein was neutralized by adding Na^+ and Cl^- ions to a concentration of 200 mM NaCl. Two systems with different concentrations of $MgCl_2$ were created: one with 102 Mg^{2+} ions, equal to the number of the *Tc*AChE carboxylate residues, and the second one with 306 Mg^{2+} ions, thus providing an excess of Mg^{2+} ions.

To adjust parameters for the MTD simulations, 50 ns of a classical unbiased MD trajectory was obtained with Gromacs 2023.3 (Abraham et al., 2015), under conditions similar to those for the above-mentioned NAMD simulations.

MTD simulations were performed using Gromacs 2023.3 patched with Plumed 2.9.0 (Bussi & Tribello, 2019; Tribello et al., 2014). The collective variable was chosen as the number of contacts between the protein and Mg^{2+} ions, calculated as the `COORDINATION` function (White & Voth, 2014), with parameters of the switch function selected from the initial unbiased MD run radial function distribution (Bussi & Tribello, 2019; White & Voth, 2014). The other MTD parameters selected from the test runs were the bias factor of 10, hills height 10, and $\sigma = 2.0$.

For analysis of the MTD simulations, the free energy profiles (FEPs) and surfaces, as a function of the number of contacts between individual residues and sites of interest, and Mg^{2+} ions, were produced upon reweighting the MTD simulations using the PLUMED driver tool (Bussi & Tribello, 2019). The MTD simulations were continued during 1.5 μs, until the shapes of the calculated profiles and surfaces did not change, and block analysis established that the error value with increasing block size had reached a plateau.

3 | RESULTS AND DISCUSSION

3.1 | Structure and sequence analysis

3.1.1 | Metal cations in TcAChE crystal structures

Currently, there are 122 *Tc*AChE crystal structures deposited in the PDB. Only 10 of them, listed in Table 1, have metal cations included in a PDB file.

Other metal-binding sites

In addition to the 4D site, metal cations have been shown to interact with His264 and adjacent Glu residues, referred to as the H264 site, or the 2nd site, and with His471 and adjacent Glu residues, referred to as the H471 site, or the 3rd site, both located on the protein surface and distant both from each other and from the 4D site (Figure 1).

In the 2nd site, metal cations are bound in two alternative ways. In one case, the metal ion interacts with His264, Glu260, and either Glu163 or Glu261 (Figure 2a–c). In the other case, the metal ion is located between His264 and Glu268 (Figure 2d–f).

In the 3rd site, metal cations are also seen to be bound in two alternative ways. Thus, the metal cation may be located either between His471 and Glu139 (Figure 3a,b) or between His471 and Glu484 (Figure 3c).

Electron densities

Assignments of some of the metal ions bound to the 2nd and 3rd sites, but not the 4D site, need to be re-evaluated due to large deviations between the metal B-factors and environmental *B*-factors obtained from the CheckMyMetal web server (Dauter, 2006; Handing et al., 2018) and listed in Table 1. Careful inspection of the electron density maps of the 112 *Tc*AChE crystal structures suggests that some additional structures may also contain bound metal ions. These include:

TABLE 1 *Torpedo californica* acetylcholinesterase crystal structures containing metal cations.

			His264—2nd site		His471—3rd site
PDB ID	**Resolution, Å**	**4D**[a]	**Glu268**	**Glu260/Glu163**	**Glu139/Glu484**
1W6R (Greenblatt et al., 2004)	2.05	Mg^{2+} 48.3 (42.7)			
2J3Q (Harel et al., 2008)	2.80	Mg^{2+} 18.9 (25.7)	Mg^{2+} 15.7 (41.0)		
7B8E (Silman et al., 2021)	2.23	Ca^{2+} 46.5 (43.8)	Zn^{2+} 65.5 (57.5)		
7B2W (Silman et al., 2021; Sussman et al., 1991)	2.65	$(UO_2)^{2+}$ 48.8 (43.5)		$(UO_2)^{2+}$ 108.1 (78.1)	
2J3D (Harel et al., 2008)	2.60	Mg^{2+} 13.8 (16.4)		2Mg^{2+}[b] 11.2 (47.9) 40.7 (48.8)	2Mg^{2+} 1.0 (32.3) 6.5 (28.9)
7B38 (Silman et al., 2021)	1.85	Mg^{2+} 26.6 (31.0)	Mg^{2+} 59.5 (46.0) Zn^{2+} 46.0 (47.4)		2Mg^{2+} 51.4 (49.0)
6G4O (Santoni et al., 2018)	2.78		Na^{+}[c] 63.3 (71.1)		
1FSS[d] (Harel et al., 1995)	3.00			Zn^{2+}[e] 19.8 (26.5) 19.2 (32.6)	
2XI4 (Sanson et al., 2011)	2.30			K^{+} 36.4 (50.9)	
2J4F[f] (Kreimer, Dolginova, et al., 1994)	2.80				Hg^{2+} 30.4 (19.6)

[a] Metal ion *B* factor, with valence-weighted environmental average *B* factor in parentheses, provided by data obtained from the CheckMyMetal web server (Handing et al., 2018; Zheng et al., 2014).

[b] Interacting with Glu261 rather than Glu163.

[c] Chain *B* only.

[d] In this structure, an additional Zn^{2+} ion interacts with Glu49, His29, and Asp285, an interaction which is not seen in the other structures.

[e] The Zn^{2+} listed in the PDB file, which interacts with Glu350, has a symmetrically related copy in the 2nd site of the asymmetric unit.

[f] This structure also contains Hg(II) bound to Cys231S$^{\gamma}$.

1. 1U65: in the 4D site (Figure S1A); in the 2nd site, between His264 and Glu268 (Figure S1B); and in the 3rd site, between His471 and Glu484 (Figure S1C).
2. 6G4O: in the 2nd site, near Glu268 (Figure S1D).
3. 2V96: in the 2nd site, between His264, Glu163 and Glu260 (Figure S1E).
4. 2XI4: in the 2nd site the difference electron density map suggests that a higher atomic number cation might be assigned, rather than K^{+}, as designated in the PDB file (Figure S1F), which is consistent with the results of the CheckMyMetal web server validation.

Crystallographic packing

We examined crystal packing contacts in the crystal structures listed in Table 1. In all except two cases, the metal ions in the 2nd and 3rd sites interact with symmetrically related residues (Table 2, Figure S2). These are non-physiological interactions since the physiological *Tc*AChE structure is composed of two monomers interacting through a four-helix bundle (Novichkova et al., 2019). Only in two crystal structures are the metal ions bound to the 2nd site not involved in crystal packing, viz. 7B2W containing a uranyl cation, and 7B38, containing Mg^{2+} and Zn^{2+}. It should be noted that the 4D motif was never seen at a crystal packing interface, and thus appears to be *self-contained*.

These observations raise questions as to whether the binding of metal cations to the 2nd and 3rd sites observed in the crystal structures are crystallographic artifacts, or whether similar binding also occurs in solution. Also, as to whether they may correspond to the weak binding sites for which experimental evidence had been presented previously (Silman et al., 2021), although without any structural assignment. These questions have been

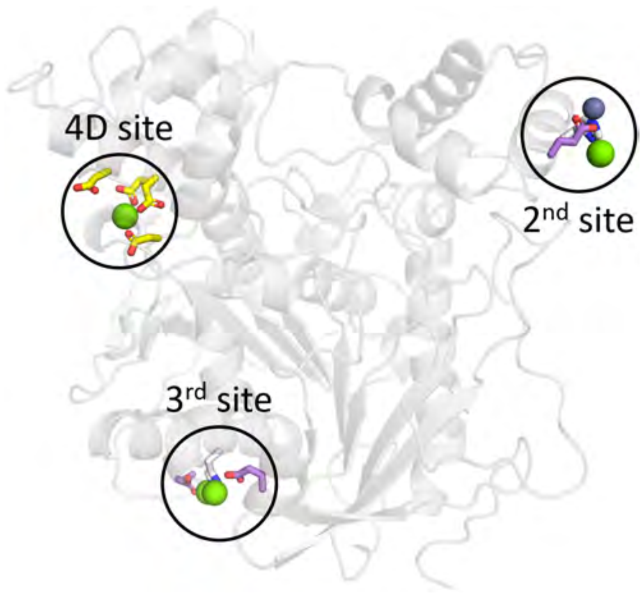


FIGURE 1 The three major metal-binding sites in the *Tc*AChE crystal structures. The 4D site consists of Asp326, Asp389, Asp392, and Asp393; the 2nd site consists of His264 and surrounding Glu residues; the 3rd site consists of His471 and surrounding Glus. The structure shown is that of 7B38, in which all three sites are occupied (Silman et al., 2021).

explored via the classical MD method, using a MTD approach, as described in Section "Classical MD MTD simulations of binding of Mg^{2+} to the surface of *Tc*AChE."

3.1.2 | The 4D site in the *Tc*AChE crystal structures

In addition to the six X-ray structures listed in Table 1, in which a cation is bound at the 4D site, and one structure in which the electron density difference map suggests the presence of a metal cation, there are nine *Tc*AChE structures in the PDB in which a PEG fragment, i.e., ethylene glycol or longer, is located near the 4D site (Table 3). These fragments do not enter the 4D site itself, except in the case of 5E4T (Dym et al., 2016). Superimposing it with a metal-bound structure, PEG comes as close as 1.8 Å to the metal site (minimal distance from the overlayed metal to the PEG fragment atom, Figure S3A). Further analysis of electron densities suggests that there might be more structures in which PEG fragments are adjacent to the 4D site (Table 3). These include structures in which unattributed elongated blobs penetrate the 4D site. The 5NUU structure (Chalupova et al., 2019) shown in Figure S3B is one such case.

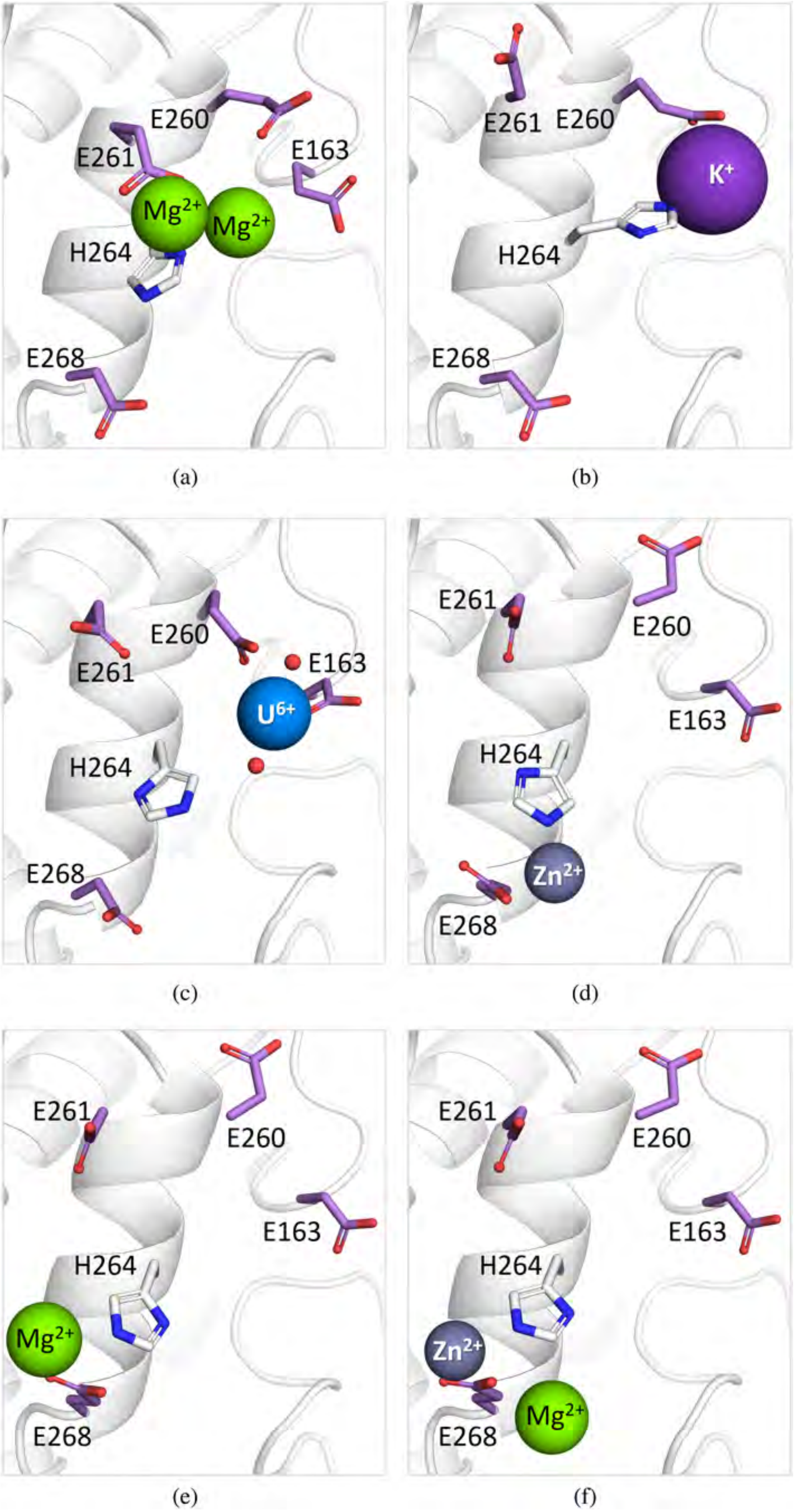
Overlaying the metal-bound 4D motifs of the available crystallographic structures shows that their geometries are essentially identical (Figure 4a), with that containing the $(UO_2)^{2+}$ cation (PDB ID 7B2W) being slightly different. The Ca^{2+} and Mg^{+2} ions embedded in the 4D motif display a square pyramid coordination shell (Figure 4b), which is not typical for complexes of these ions, which are usually octahedral (Katz et al., 1996). The first coordination shell consists of two Asp residues bound in an *anti*-monodentate mode (Rardin et al., 1991). The other two Asps are in the second coordination shell, interacting with the metal ion through bridging water molecules.

We compared the deviations of the carboxyl groups of the 4D motif in all the *Tc*AChE crystallographic structures available in the PDB, taking PDB ID 7B8E, with Ca^{2+} within the 4D site, as the reference structure (Table 3).

As might be expected, the top entries in Table 3 are the X-ray structures in which a divalent ion is present within the 4D motif, except for PDB ID 7B2W, in which an uranyl oxycation is present (Figure 4a). However, the next entry is 1U65, which may have an unattributed metal ion within the 4D motif. It is followed by 1HBJ, which also displays an unattributed electron density peak adjacent to the 4D site, most likely a PEG fragment, as discussed above. The next structure, 1VOT (Raves et al., 1997), is devoid of any suspicious electron density peaks in or near the 4D site and also displays very low deviations of the aspartate carboxylates from the reference structure (Figure 5a). This is rather surprising, as mentioned above, since the unoccupied 4D motif contains four aspartates in close proximity; in the absence of a cation, one might expect that electrostatic repulsion would cause them to move away from each other. At the same time, much higher displacements in the structures located at the bottom of Table 4 suggest that the dynamics of an unoccupied 4D site differs from that of the site when occupied by a metal cation.

Inspection of the coordination shell surrounding the 4D motif revealed that it contains three positively charged residues, Lys325, Arg388, and Arg517, which form salt bridges with three of the four Asp residues in the 4D motif—Lys325 with Asp326, Arg388 with Asp389, and Arg517 with Asp393 (Figure 5b). Apparently, these three cationic residues stabilize the 4D motif in the absence of a metal cation, and the 4D motif can better be described as three ionic pairs and one Asp residue. Table 3 shows that Asp392, which is not stabilized by a cationic residue, adopts a broader range of conformations than the other three residues of the 4D motif.

A comparison of the sequences of AChEs and BChEs from various species shows that the 4D motif occurs frequently in both enzymes (Silman et al., 2021). The same type of analysis for the three surrounding cationic



(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 2 Examples of binding of metal cations to the 2nd site in the *Torpedo californica* acetylcholinesterase crystal structures: (a) 2J3D, (b) 2XI4, (c) 7B2W, (d) 7B8E, (e) 2J3Q, (f) 7B38.

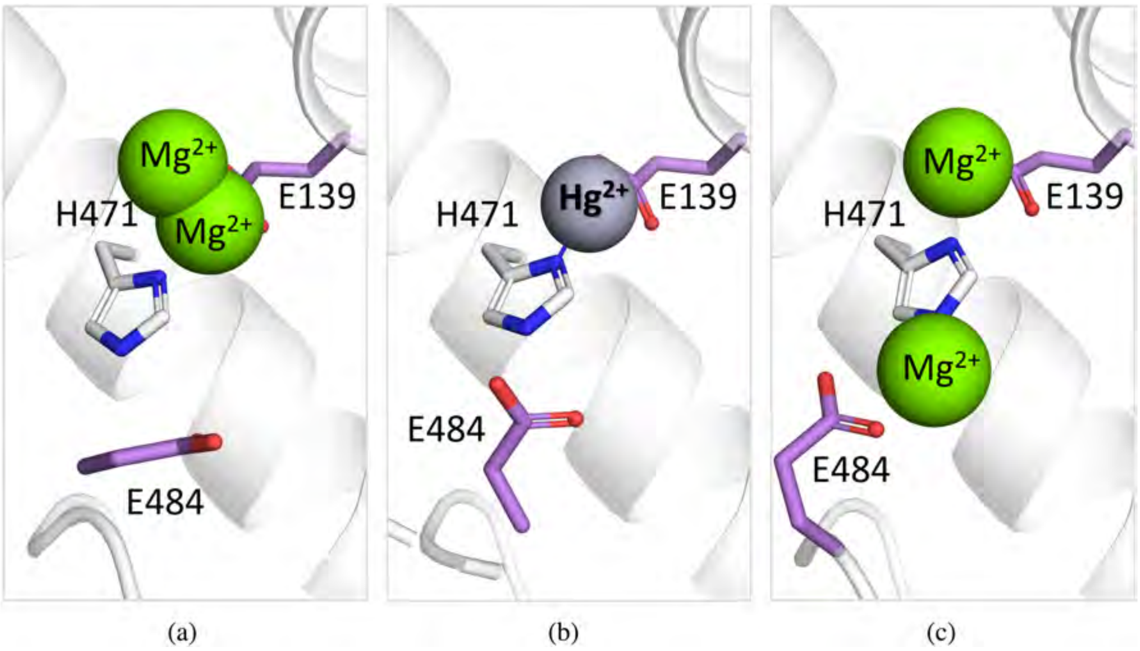


FIGURE 3 Examples of binding of metal cations to the 3rd site in the *Torpedo californica* acetylcholinesterase crystal structures: (a) 2J3D, (b) 2J4F, (c) 7B38.

TABLE 2 Symmetric copy residues interacting with the metal cations bound to the site.

		2nd site		3rd site	
PDB ID[a]	**Space group**	**Residues in the 2nd site (H264)**[b]	**Symmetry-related residue(s)**	**Residues in the 3rd site (H471)**[b]	**Symmetry-related residues**
1FSS, A	P 21 21 2	Zn^{2+}; E260	E350		
2J3D, B	C 1 2 1	$2Mg^{2+}$; E260, E261	H471, E139, E484 (3rd site)	$2Mg^{2+}$, E139	H264, E260, E163 (2nd site)
2J3Q, C	P 32 2 1	Mg^{2+}; E268	H181, D217, K315		
7B8E, D	P 32 2 1	Zn^{2+}; E268	H181		
2XI4, E	P 21 21 21	K^{+}; E260, E163	E508, K511		
6G4O, F	P 21 21 21	Na^{+}; E268	E508, K511		
7B38, G	P 32 2 1	Mg^{2+}, Zn^{2+}; E268	H181, D217, K315	$2Mg^{2+}$; E139, E484	H471, E139, E484, E489 (3rd site)
7B2W, H	P 32 2 1	$(UO_2)^{2+}$; E260, E163			

[a] Figure S2 displays the crystallographic packing interface. The upper-case letter following each PDB entry indicates the corresponding panel in the supplementary figure.

[b] The metal cation bound to the site and the Glu residues participating in the binding of the metal ion together with the corresponding His residue are listed.

residues (Table 4) reveals that they are similarly, or, perhaps, even better, conserved. Thus, in human, mouse and bovine AChE, and in human BChE, Asp392 (*Tc*AChE numbering), which lacks ionic interactions with any surrounding positively charged residues, is replaced by Ser or Gly, which affects the overall charge of the whole motif. In the designed thermostable human AChE D4 variant, this replacement is retained (Goldenzweig et al., 2016). In the mouse sequence, there is one more neutralizing substitution of Asp393 by Ala, and in *Anopheles gambiae* AChE it is replaced by a positively charged residue. In both these cases, their structural ionic partner, conserved Arg525, is retained. In several cases in which Asp389 (*Tc*AChE numbering) is replaced by Glu, viz., in human, mouse and *Bungarus* AChE, and in human, bovine, horse, sheep, rabbit, *Macaca mullata* (rhesus monkey) and cat BChE, the corresponding Arg residue is retained.

TABLE 3 Relative conformations of the 4D motifs in the *Torpedo californica* acetylcholinesterase (*Tc*AChE) X-ray structures.

PDB ID	Chain	D326		D389		D392		D393		Max	Ligand	Electron density peak, σ
		$O^{\delta1}$	$O^{\delta2}$	$O^{\delta1}$	$O^{\delta2}$	$O^{\delta1}$	$O^{\delta2}$	$O^{\delta1}$	$O^{\delta2}$			
7b8e	A	0	0	0	0	0	0	0	0	0	Ca^{2+}	14.9
7b38	A	0.26	0.26	0.25	0.06	0.18	0.07	0.09	0.09	0.26	Mg^{2+}	14
2j3q	A	0.33	0.31	0.15	0.23	0.15	0.22	0.23	0.22	0.33	Mg^{2+}	12.5
2j3d	A	0.36	0.34	0.17	0.17	0.12	0.23	0.26	0.29	0.36	Mg^{2+}	11.4
1w6r	A	0.37	0.41	0.22	0.25	0.13	0.42	0.27	0.25	0.42	Mg^{2+}	9.5
1u65	A	0.31	0.38	0.34	0.13	0.3	0.43	0.13	0.19	0.43	*PEG*	8.3
1hbj	A	0.44	0.47	0.11	0.12	0.3	0.2	0.18	0.24	0.47		6.6
1vot	A	0.45	0.46	0.22	0.19	0.4	0.47	0.41	0.41	0.47		3.6
2ack	A	0.45	0.49	0.23	0.21	0.38	0.05	0.37	0.47	0.49		3.5
2vjd	A	0.28	0.4	0.15	0.23	0.4	0.49	0.22	0.39	0.49		3.3
2vjd	B	0.26	0.36	0.16	0.21	0.34	0.5	0.17	0.35	0.50		3.5

Note: RMS deviations of the oxygen atoms of the carboxylate groups of the Asps forming the 4D motif for all the *Tc*AChE structures deposited in the PDB. The deviations (values in Å) are obtained based on alignment to the corresponding residues in the 4D motif of the 7B8E reference structure by the C^{γ} atoms (the first entry with zero distances). The entries are color-coded from the lowest values in green to the highest values in red. The "Max" column refers to the largest deviation for an individual carboxylate oxygen atom. The table is sorted in ascending order in this column. Only the first 10 rows are shown in the main text. The full version is provided in Table S1. Ligands in the 4D site are listed according to the PDB files; in some cases, electron density within or adjacent to the 4D motif suggests the presence of a PEG fragment (Dym et al., 2016). In such cases, PEG is italicized, and colored gray. When the 4D site is occupied only by water molecules, and inspection of the electron density does not suggest the presence of a PEG fragment, the corresponding cell is left blank. The last column lists the highest electron density peak observed in the 4D site, based on F_o–F_c difference maps obtained for protein models with all ions and water molecules omitted. The columns with deviations of the oxygen atoms and the column with the electron density peaks are colored independently.

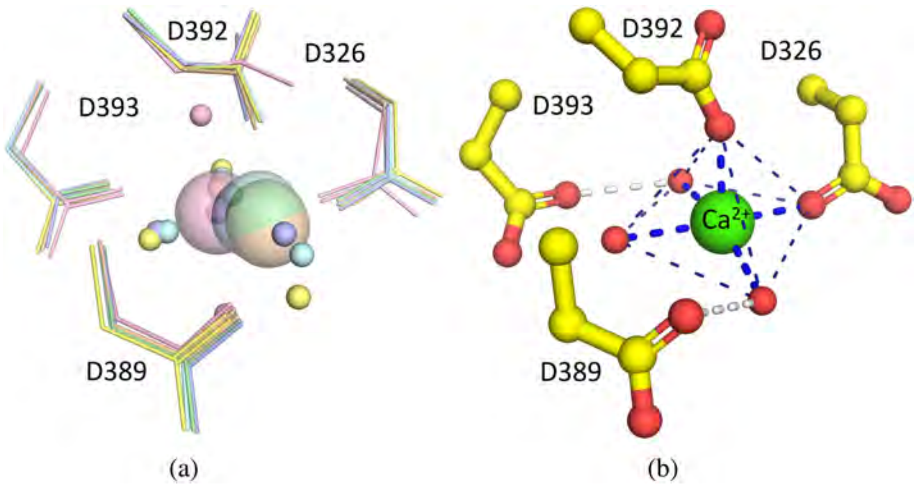


FIGURE 4 (a) Overlay of the metal-bound 4D motifs in several *Tc*AChE crystal structures: 1W6R, atoms colored blue; 2J3D, green; 2J3Q, beige; 7B2W, pink; 7B38, cyan; 7B8E, yellow. In 7B2W, no waters are detected, the oxygen atoms displayed being those of the $(UO_2)^{2+}$ cation. The side chains of the Asp residues are shown as sticks, the cations as semi-transparent spheres, and oxygens of waters as red spheres of a smaller radius. (b) Structure of the *Tc*AChE 4D motif (PDB 7B8E) with coordination geometry of Ca^{2+} shown with black dashes. Oxygens of the crystallographic waters are displayed as isolated spheres. Blue dashes show direct interactions with the ion, and white dashes show the water-bridged interactions.

Positively charged residues vary in the following three cases: the structural ionic pair Arg325–Asp326, in electric eel AChE, is replaced by the neutral residues, Gln and Asn, and the residue homologous to Arg517 is replaced by a Lys. In *Danio rario* AChE there is another neutralizing replacement of Lys325 by Gln.

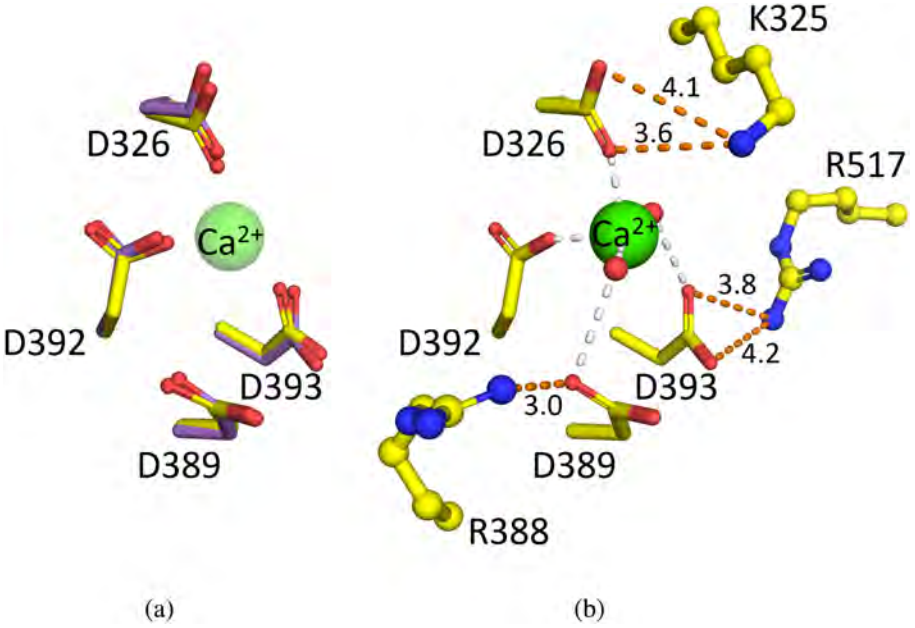


FIGURE 5 Charge stabilization in the *Tc*AChE 4D motif: (a) Overlay of the aligned 4D motifs in the X-ray structures of the Ca^{+2}/*Tc*AChE complex (PDB ID 7B8E, carbon atoms in yellow) and the apo enzyme (PDB ID 1VOT, carbon atoms in violet); (b) Close up of the 4D site in 7B8E. Interactions within the motif are shown as white dashed lines; ionic interactions between the motif's Asp residues and Arg and Lys residues surrounding the 4D site are shown as orange dashed lines.

TABLE 4 Residues homologous to the 4D motif Asp residues and to the adjacent cationic residues of *Tc*AChE in acetylcholinesterases (AChEs) and butyrylcholinesterases (BChEs) from other species.

Enzyme	Species	UniProt accession number[a]	Residues						
AChE	*Torpedo californica*[b]	P04058	K325	D326	R388	D389	D392	D393	R517
	Human	P22303	K332	D333	R395	*E396*	**S399**	D400	R525
	Designed human (Goldenzweig et al., 2016)		K332	D333	R395	D396	**S399**	D400	R525
	Mouse	P21836	K332	D333	R395	D396	**S399**	**A400**	R525
	Bovine	P23795	K331	D332	R394	*E395*	**S398**	D399	R524
	Electrophorus electricus	O42275	**Q327**	**N328**	R390	*E391*	D394	D395	*K552*
	Danio rario (zebrafish)	Q9DDE3 (-23)	**Q350**	D351	R413	D414	D417	D418	R576
	Bungarus fasciatus (krait	Q92035	K325	D326	R388	*E389*	D392	D393	R517
	Anopheles gambiae (mosquito)	Q869C3	**T484**	*E485*	R548	D549	D552	**K553**[c]	R679
BChE[d]	Human	P06276	K323	D324	R386	*E387*	**G390**	D391	R515
	Pig	A0A4X1VAZ7 (-28)	K351	D352	R414	D415	D418	D419	R543
	Mouse	Q03311 (-29)	K352	D353	R415	D416	D419	D420	R544
	Dog	A0A8C0NAH8 (-28)	K351	D352	R414	D415	D418	D419	R543
	Rhegmatorhina hoffmannsi (white-breasted antbird)	A0A973YFB2 (-28)	K351	D352	R414	D415	D418	D419	R543
	Horse	P81908	K323	D324	R386	*E387*	D390	D391	R515
	Bovine	P32749 (-28)	K351	D352	R414	*E415*	D418	D419	R543
	Rabbit	P21927 (-7)	K330	D331	R393	*E394*	D397	D398	R522
	Macaca mulatta (Rhesus macaque)	H9FXZ3 (-28)	K351	D352	R414	*E415*	D418	D419	R543

Note: Charge conserving variations are italicized, neutralizing changes are shown in bold, and opposite charge changes in red.

[a] The signal sequence length is given in brackets, and needs to be subtracted so as to generate the residue numbering of the expressed protein.

[b] Structural ion pairs: K325-D326, R388-D389, D393-R517.

[c] In the *Anopheles gambiae* AChE crystal structure K553 interacts with E485, D549 and D552; R679 interacts with D557, which is not homologous to any of the *Tc*AChE 4D residues, but is located near the 4D site.

[d] Only the X-ray structure of human BChE is available in the PDB. For the other sequences, AlphaFold (Jumper et al., 2021) predicts the same configuration of the 4D site and of the surrounding residues.

3.2 | Molecular modeling of the binding of a metal cation to the *Tc*AChE 4D site

The fact that the configuration of the 4D motif is very similar in the *apo* enzyme and when it is occupied by a divalent metal ion was, as already stated, an unanticipated finding, which appears to be largely explained by the presence of adjacent positively charged amino acids. However, several issues need to be addressed. First, does the presence of a metal cation affect the dynamics of the 4D motif? Second, does the absence of a metal cation cause proton redistribution between the closely located carboxylic groups and water molecules? Third, how do the surrounding cationic residues affect its stability, compared to the other metal cation binding structural elements?

Accordingly, we searched for a protein containing four Asp residues coordinating a metal cation, which were devoid of adjacent Arg/Lys residues. We sought metal-binding sites composed of relatively distal Asp residues in the sequence, as is the case in the 4D motif of *Tc*AChE. The active human fibrin-stabilizing factor (fXIIIa) PDB 4KTY (Stieler et al., 2013) fulfills these criteria. It connects two loops that are separated by a β-strand. The Ca^{2+} ion that occupies this site has an octahedral coordination shell (Figure 6) that is typical for this ion (Katz et al., 1996), but differs from the square pyramidal coordination shell seen in *Tc*AChE. However, as in the case of *Tc*AChE, the ion is bound to the carboxylic groups in a monodentate mode.

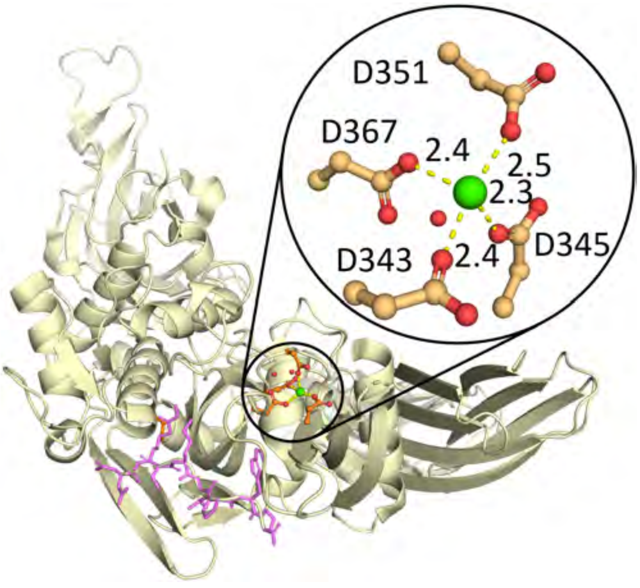


FIGURE 6 Structure of fXIIIa 4KTY and its 4D Ca^{2+}-binding domain, which is shown in greater detail in the blowup. The numbers denote H-bond lengths in Å. A covalent inhibitor present in the structure is shown as pink sticks.

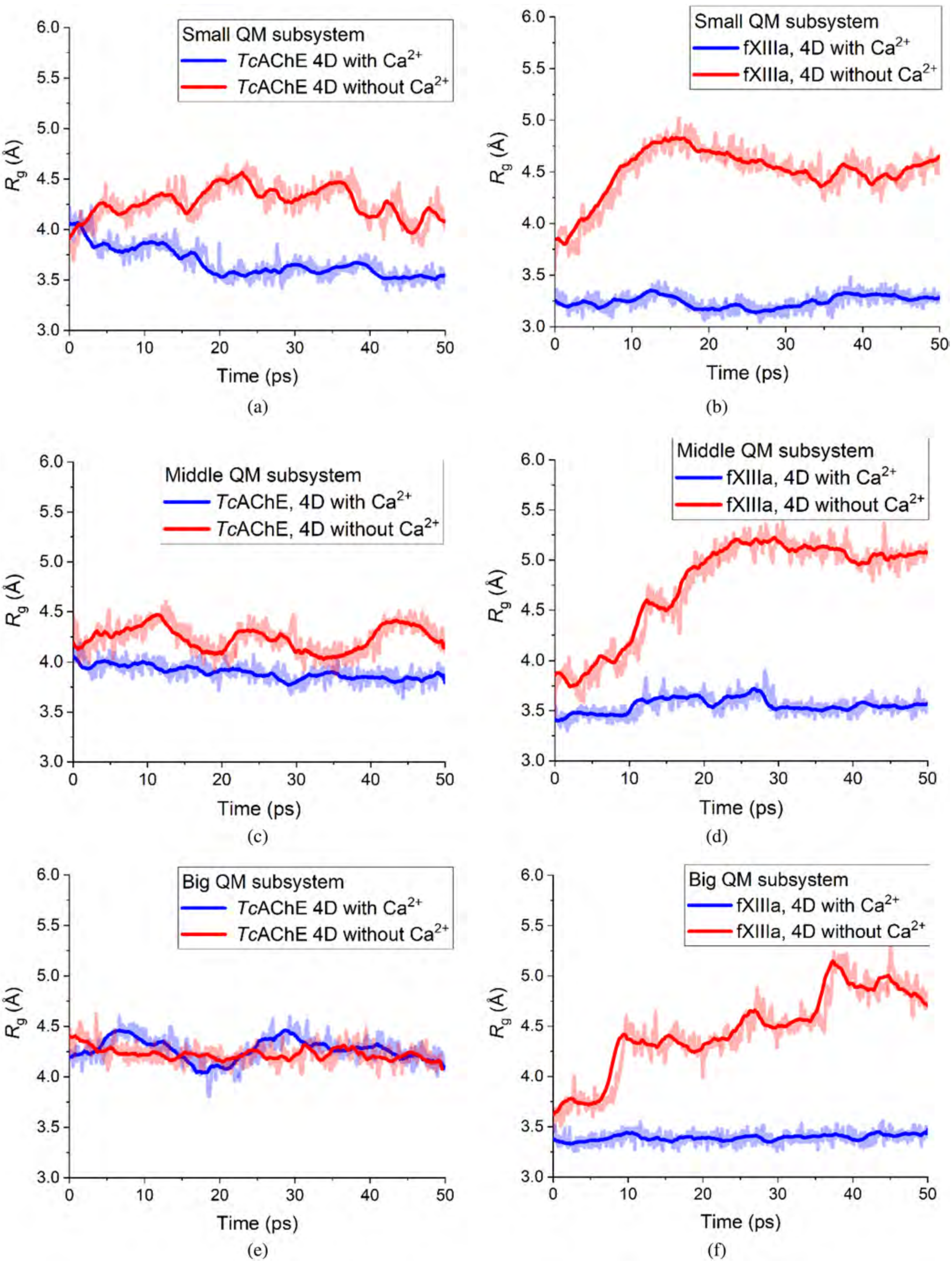
3.2.1 | Unbiased QM/MM-MD simulations

Unbiased MD simulations with quantum potentials were performed for the *Tc*AChE and fXIIIa monomers in the presence and absence of a Ca^{2+} ion in the 4D site. The flexibility of the *Tc*AChE 4D site was observed to increase in the absence of the cation. Unlike the more rigid structure of the metal-cation-occupied site, it displays periodic breathing-like motions (Figure 7a). However, the absence of a divalent cation does not result in a significant change in structure or proton redistribution on the timescale of the simulation. The low amplitude periodic motions in the absence of the metal cation can be attributed to the ability of a particular structure to assume the range of conformations of the Asp carboxylates observed in the repertoire of X-ray structures with unoccupied 4D sites displayed in Table 3.

In the absence of Ca^{2+} ions in the *Tc*AChE 4D site, its size does not depend on the size of the QM subsystem (Figure 7a,c,e, Figure S7A). For the metal-bound 4D site, the R_g difference for the three QM subsystems of different sizes is rather pronounced (Figure 7a,c,e, Figure S7A). When only Asp residues are included in the QM subsystem, the 4D site substantially decreases relative to the X-ray structure (Figure 7a). When all three positively charged residues are included in the QM-subsystem, the 4D site sizes in the presence and absence of a bication become very close, supporting our hypothesis (Figure 7e). This also suggests that interactions between positively and negatively charged residues are described similarly when the positively charged residues are treated by the QM DFT method or by classical force fields (Figure S7C). Meanwhile, in the presence of Ca^{2+}, differences in the 4D site size for the QM subsystems become apparent (Figure S7D). This may be attributed to a polarization effect of Ca^{2+} on surrounding Asp residues, and to the second shell positively charged residues.

In sharp contrast, in the case of fXIIIa, the replacement of Ca^{2+} by water leads to a major disruption of the structure of the 4D site, which is irreversible on the timescale of the simulation (Figure 7b,d,f, Figure S7B). The influence of the QM subsystem sizes on the 4D site size is much less pronounced than in the case of *Tc*AChE (Figure S7A,B).

In the X-ray structure of the 4D site in fXIIIA, Asp343 and Asp345 are located on a loop, Asp351 on the beginning of the following β-strand, and Asp367 on a loop following the β-strand. All these elements are held in place by the coordinating Ca^{2+} ion (Figure 8a). Without the Ca^{2+}, the loop carrying Asp343 and Asp345 is displaced and becomes flexible. All four Asps form new H-bonds with neighboring residues to compensate for the



(a) (b)

(c) (d)

(e) (f)

FIGURE 7 Legend on next page.

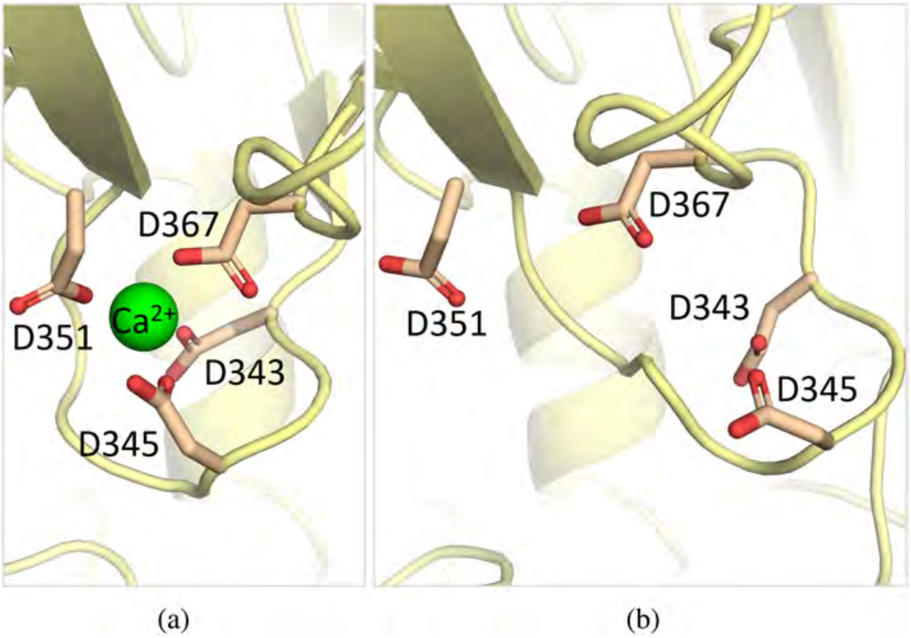


FIGURE 8 Comparison of the starting and final structures of the fXIIIa 4D site in the absence of a metal cation, after 50 ps molecular dynamics with quantum potentials (QM/MM-MD) simulations. (a) An initial model structure after the preliminary optimization. The full QM subsystem is shown in Figure S5. (b) The deformed structure of the 4D site in the absence of the Ca^{2+} ion, and the loop on which it is partly located, taken from a snapshot at the end of the MD trajectory.

electrostatic repulsion (Figure 8b). However, formation of a new stable structure was not observed.

3.2.2 | Energy profiles of unbinding of Ca^{2+} from the 4D motifs of *Tc*AChE and fXIIIa

Comparing the PMF profiles obtained by QM/MD umbrella sampling for binding/unbinding of Ca^{2+} to the 4D sites of the two proteins, a significant difference in energy was observed (Figure 9).

The values obtained could be compared to analogous unbinding energy barrier calculation data reported in the literature. Thus, for unbinding of Ca^{2+} from an acetate anion in water, using a classical MD umbrella sampling method, energy barriers were obtained in a range 4.7–5.8 kcal/mol (Iskrenova-Tchoukova et al., 2010; Jafari et al., 2024). The unbinding of Ca^{2+}/Mg^{2+} ions from the catalytic site of ribonuclease H was modeled by Babu et al. (Babu et al., 2013) using a similar classical MD method. In this active site, an ion is coordinated by two Asps, one of which is a catalytic residue, and one Glu. The calculated unbinding energy barrier for Ca^{2+} from this site is 19.7 kcal/mol (Babu et al., 2013).

Although in the present study we used the MD umbrella sampling method with QM potentials, whereas in both quoted studies (Babu et al., 2013; Iskrenova-Tchoukova et al., 2010), classical force fields were utilized, the results that we obtained are consistent with theirs. The calculated energy barrier for unbinding of Ca^{2+} from the 4D site of *Tc*AChE is much lower than that for the sites consisting of 3 or 4 carboxylates that are not surrounded by positively charged side chains. However, it is higher than that calculated for a single negative charge of an acetate ion, which is a simple molecule, not a protein. Thus, the binding of a metal cation to the *Tc*AChE 4D site cannot be attributed solely to interaction with the only residue in the site not forming an ion pair with a cationic amino acid, namely Asp392 (Figure 5b).

Relatively weak binding of a cation to the 4D site of *Tc*AChE raises the question: how competitive is binding to the *Tc*AChE 4D site, compared to other surface cation-binding areas.

Classical MD MTD simulations of binding of Mg^{2+} to the surface of Tc*AChE*

The above MD simulations with QM potentials were performed for the *Tc*AChE 4D site complexed with Ca^{2+}.

FIGURE 7 Results of 50 ps unbiased molecular dynamics (MD) simulations with quantum potentials for the three quantum mechanical (QM) subsystems described under the Methods: small (a, b), medium (c, d), and big (e, f). Evolution along the MD trajectory of the sizes of the 4D sites of *Tc*AChE (left column) and fXIIIa (right column) (with the size being defined as the R_g) in the presence and absence of a Ca^{2+} ion. For comparison, different combinations of these plots are shown in Figure S7. Values with a 1 fs timestep are shown as semi-transparent, and averaged over 2 ps as bold lines. Due to different configurations of the 4D sites, the R_g values of the Ca^{2+} bound states differ for *Tc*AChE and fXIIIa. R_g, radius of gyration.

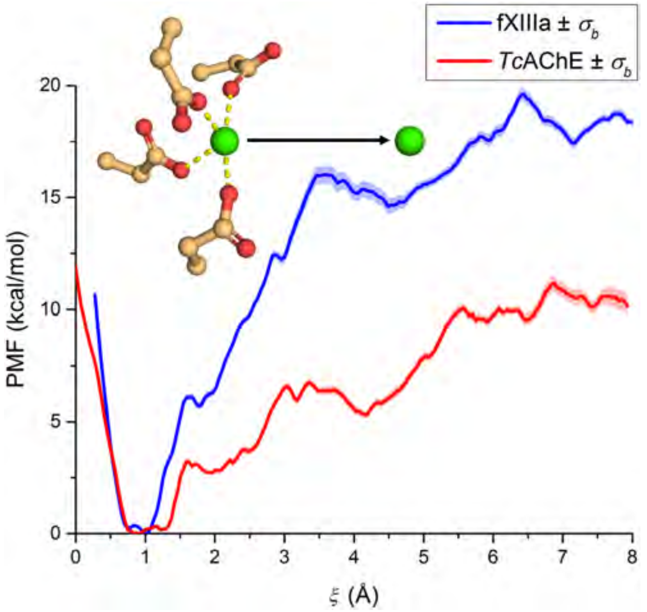


FIGURE 9 Quantum mechanical/molecular dynamics potential of mean force (PMF) profiles for pulling a metal cation out of the 4D sites of fXIIIa (blue) and *Tc*AChE (red). The ±SD range is shown as a semi-transparent area around the PMF value lines. The reaction coordinate, ξ, is the distance between Ca^{2+} and the COMs of the carboxylate oxygens of all four Asps forming the 4D sites. Due to the different configurations of the 4D sites of the two proteins, the location of the PMF profile minimum, corresponding to an optimal position of Ca^{2+} within the site, differs between them. The minima were superimposed at $\xi = 1$ Å for easier comparison of the two PMF profiles. Energy barriers for the initial displacement of the metal ion from its equilibrium positions in the bound states are 3.2 kcal/mol for *Tc*AChE, and 6.1 kcal/mol for fXIIIa. Maximal barriers at ~6 Å from the 4D site centers are 11.2 kcal/mol for *Tc*AChE and 19.6 kcal/mol for fXIIIa, which could be attributed to the general attraction of the electrostatic field of a protein surface, including the 4D site. Beyond this distance, the influence of the 4D sites' attraction is less pronounced.

However, it was experimentally determined that Mg^{2+} thermostabilizes *Tc*AChE significantly more strongly (Silman et al., 2021). Furthermore, Ca^{2+} is never seen in the other binding sites in the available X-ray structures, whereas Mg^{2+} is present in those sites in several of them (Table 1). Accordingly, the following classical MD MTD calculations of the energetics of binding of a divalent metal cation to the *Tc*AChE surface were performed with Mg^{2+}.

Using the MTD approach, we estimated relative energy levels for different numbers of contacts between Mg^{2+} ions and either the whole protein, or one or more specific groups. For contacts between the entire protein and Mg^{2+} ions, a single energy minimum of the FEP was observed at the level of 21 contacts (Figure 10, insert). This corresponds to 13–17 ions within 4 Å of *Tc*AChE.

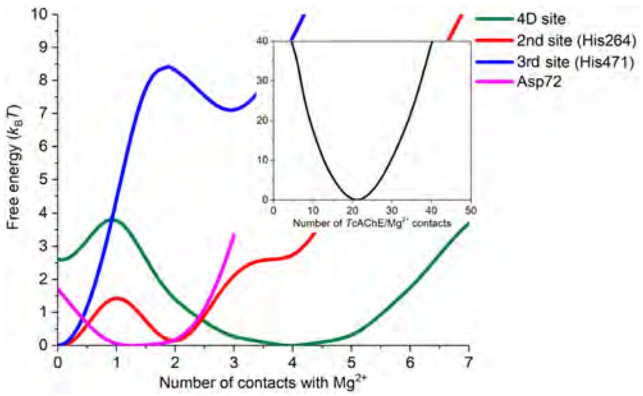


FIGURE 10 The calculated free energy profiles for the number of contacts between Mg^{2+} ions and the entire *Tc*AChE molecule (insert) or its groups (main figure): The 4D site (green); the 2nd site (red); the 3rd site (blue); Asp72 carboxylate (magenta).

For the 4D site, the FEP global minimum is at the level of 3–4 contacts. For the 2nd (His264) site, two very close energy minima were obtained at levels of 0 and 2 contacts (Figure 10). An increase in the number of Mg^{2+} ions in the simulation system led to a shift of the energy minimum to four contacts for the 4D site, and to a minimum at the level of two contacts for the 2nd site, which becomes considerably deeper than at the level of 0 contacts (Figure S8). Two individual residues on the protein surface, Asp1 and Asp72, also had energy minima at the level of 1–2 contacts. Upon increasing the number of Mg^{2+} ions in the simulation system, Asp285 and Glu350 also display energy minima at the level of 1–2 contacts. The Asp residues comprising the 4D site for both low and high numbers of Mg^{2+} ions also had global minima at 1–2 contacts, except for Asp393.

The FEPs for the 3rd site, either for His471 alone, or together with neighboring Glu139/Glu484 in different combinations, have global minima at 0, even for the system with the larger number of Mg^{2+} ions (Figure 10 and Figure S8).

A comparison of energy levels for binding of a Mg^{2+} ion to the 4D site and to the 2nd site shows the 4D site to be the primary binding site (Figure 11a). With increasing numbers of Mg^{2+} ions, simultaneous binding occurs more frequently (Figure 11b). For the individual anionic residues, Asp1 and Asp72, simultaneous binding both to them and to the 4D site is the most likely scenario.

Distribution of the number of contacts at the 4D site becomes notably narrower upon increasing the number of Mg^{2+} ions. This indicates the stabilization of the bound state for the 4D site, and a concomitant decrease in the number of binding-unbinding events in the course of the simulation. In contrast, for the 2nd site, the distribution becomes wider with the increase in the number of

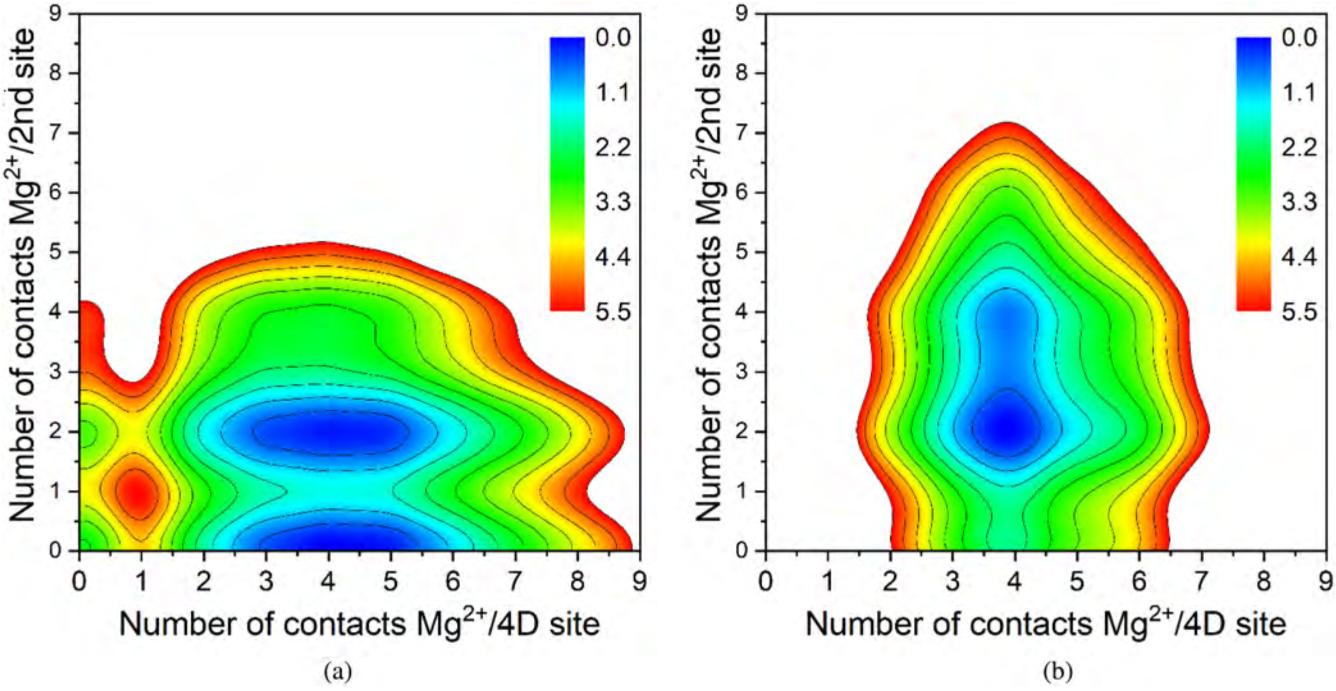


FIGURE 11 Free energy surfaces of a Mg^{2+} ion binding to the 4D site vs the 2nd site. (a) 122 Mg^{2+} ions in the simulation system; (b) 307 Mg^{2+} ions. The unit of free energy is k_BT.

Mg^{2+} ions in the system, and the distinct second local minimum appears. This is in good agreement with the experimental crystallographic observations that metal cations bind to the 2nd site in different modes (Figure 2).

The MTD simulations support the solution studies indicating the presence of the 4D site and a small number of weaker sites in *Tc*AChE (Silman et al., 2021). Whereas cation binding to the 2nd site observed in the metal-containing X-ray structures of *Tc*AChE may be expected to occur in solution, binding to the 3rd site appears to be due to crystallographic packing. Although the 2nd site is weaker than the 4D site, both would be occupied simultaneously in the presence of an excess of Mg^{2+} ions. Concentrations of both Mg^{2+} and Ca^{2+} at neuromuscular junctions are in the low millimolar range (Del Castillo & Engbaek, 1954; Wang et al., 2010). Our previous experimental work on this subject (Silman et al., 2021) was also performed with concentrations of the divalent cations of interest in the millimolar range. For the computational part of the present study, we built model systems mimicing those conditions. Thus, our data appear to be physiologically relevant.

3.3 | Is the *Tc*AChE 4D motif unique?

Metal-binding structural elements are abundant within the proteome. Tools such as ASSAM (Nadzirin et al., 2012), the PDB structural search (Guzenko et al., 2020), and pyScoMotif (Cia et al., 2023) permit template-based identification of small molecular motifs, e.g., of four negatively charged residues, independent of sequence order. Use of the IMAAAGINE web-server (Nadzirin et al., 2013) allows the identification of structural motifs without requirement for a structural template. This produces more geometrically diverse results than those generated by template-based tools. Several examples were provided in our earlier paper (Silman et al., 2021). Closer inspection of these 4D motifs revealed that they are a subset of a larger set of metal binding structures, including more than four carboxylate residues, both Asps and Glus, as well as His residues, and sometimes more than one metal ion.

The key feature of the *Tc*AChE site, consisting of 4 Asp residues, that distinguishes it from other metal-binding motifs, is its ability to bind and unbind a metal cation without undergoing a drastic geometrical change, although this was not emphasized previously (Silman et al., 2021). This is in sharp contrast to the EF-hand, for example, which undergoes a large conformational change upon binding a metal cation (Efremov et al., 2015). It is plausible to ascribe this lack of conformational change of the 4D motif to the conserved cationic residues that surround and stabilize it (Figure 5a). Indeed, as we showed above, for fXIIIa, by use of QM/MM MD, a site including four Asps, and not surrounded by cationic residues

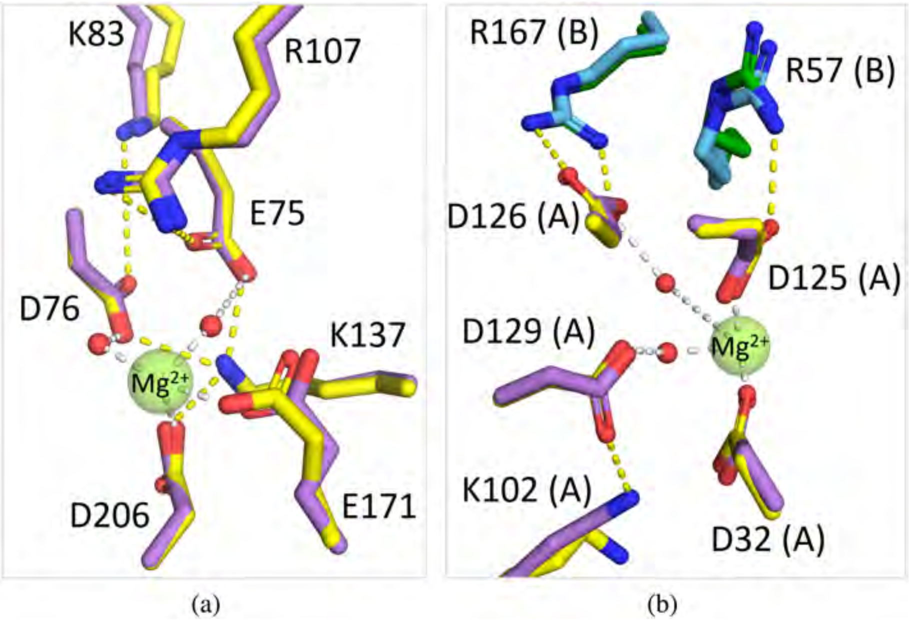


FIGURE 12 Overlay of crystallographic structures of 4D/E sites in presence and absence of Mg^{2+}: (a) 5VXC/5VXO, (b) 2R8E/2R8X. Carbon atoms in the structures containing metal cations are shown in yellow, and the *apo* structures in violet. In panel (b), for residues belonging to the second subunit coloring is in blue and green, respectively.

undergoes a dramatic conformational change upon removal of the metal cation (Figure 8), apparently due to electrostatic repulsion.

In a search for proteins containing a motif similar to the 4D motif of *Tc*AChE, we included both Asp and Glu residues in the search. Using the tools for structural motif search mentioned above, we found many proteins in the PDB that contain a site formed by four carboxylate residues closely surrounded by several cationic residues. In other words, a motif containing four acidic and three basic residues, 4A/3B. However, in most cases, either a structure in which the site is occupied by a metal ion or an *apo*- structure, is available in the PDB, but not both. This precludes experimental confirmation of the impact of the metal ion on the geometry of such a 4A/3B motif.

We did, however, find three pairs of crystal structures of proteins in which the 4A/3B motif was either occupied by a metal ion or devoid of it. One such protein is human citramalyl-CoA lyase, in which the 4A/3B site is occupied by Mg^{2+} (5VXC (Shen et al., 2017)). In the 5VXO structure, this site is not occupied by a metal cation, yet the conformation of the 4A/3B motif is practically identical to that in the 5VXC structure (Figure 12a).

Another example of a 4A/3B motif is present in the YrbI Mg^{2+}-dependent phosphatase, which is a physiological tetramer (Parsons et al., 2002). There are two pairs of structures with metal-containing and metal-free active sites in the PDB which show that absence of a metal cation does not affect the motif's conformation: from *Escherichia coli* with Mg^{2+}, 2R8X/2R8E (Biswas et al., 2009) (Figure 12b), and from *Haemophilus influenzae* with Co^{2+}, 1K1E/1J8D (Parsons et al., 2002), respectively. It is interesting to note that the 4A/3B motif of this latter enzyme is the active site. Four Asps coordinate a divalent cation either directly or through bridging water molecules. One cationic residue interacting with an Asp is located on the same subunit, whereas the other two cationic residues are on an adjacent subunit of the physiological tetramer (Figure 12b).

An additional example is *Thermotoga maritima* sodium pumping membrane integral pyrophosphatase. The conformation of the 4A/3B motif in the *apo*-enzyme (8B21) is very similar to that containing a Ca^{2+} ion (4AV3) (Kellosalo et al., 2012).

4 | CONCLUSIONS

The previously described 4D motif (Silman et al., 2021) of *Tc*AChE is more complex than was originally realized. It turns out that the four Asps are stabilized by three surrounding cationic residues. This can be also described as an assembly of three ionic pairs with an additional unpaired Asp. Thus, in more general terms, it consists of four acidic and 3 basic residues, which we call the 4A/3B motif. Unlike in the cases of other metal-binding anionic motifs, although the divalent cation is not crucial for the conformational stability of the protein, it significantly improves it. This is reflected in the pronounced increase in thermal stability (Silman et al., 2021). No significant conformational changes, and only minor differences in the *Tc*AChE 4A/3B site dynamics, were observed both in the crystallographic structures and using QM/MM-MD simulations. Contributions of individual residues to thermal stabilization could be investigated by site-directed mutagenesis (*in vitro* and *in silico*).

The 4A/3B motif's cation binding strength is ~10 kcal/mol weaker than in a site with several anionic residues devoid of charge-stabilizing residues surrounding it. This is seen, for example, in comparison of the 4D sites of *Tc*AChE and of fXIIIa, which was chosen for comparison since it is devoid of cationic residues in the vicinity of its 4D site. Nevertheless, the 4A/3B site is the primary metal cation-binding site in *Tc*AChE, relative to the other metal cation-binding sites seen in its crystal structures.

We conjecture that the 4A/3B motif, which has the ability to bind a metal ion without significant conformational change, may serve as a depot for controlling metal ion concentrations. It may also serve as an adaptive mechanism to increase protein thermal stability with increasing metal cation concentration.

AUTHOR CONTRIBUTIONS

Sofya Lushchekina: Formal analysis; conceptualization; investigation; writing – original draft; methodology; visualization. **Lev Weiner:** Investigation. **Yacov Ashani:** Investigation. **Reeki Emrizal:** Software; methodology. **Mohd Firdaus-Raih:** Methodology; software. **Israel Silman:** Conceptualization; supervision; writing – review and editing; investigation; formal analysis; writing – original draft. **Joel L. Sussman:** Conceptualization; investigation; funding acquisition; writing – review and editing; writing – original draft; validation; project administration; supervision; resources.

ACKNOWLEDGMENTS

Dr. Lushchekina's research is funded by the Miel de Botton and the Jacqueline and Marc Leland Foundation. SL has been supported by the Weizmann Institute's emergency program to host visiting scientists affected by the war in Ukraine, and by the Helen and Milton A. Kimmelman Center for Biomolecular Structure & Assembly. MF-R is funded by the Ministry of Higher Education Malaysia Translational Research Grant Scheme TRGS/1/2022/UKM/01/9/1. This study received support from the European Union's Horizon 2020 research and innovation program under grant agreement No. 824087. Financial support was also provided by Instruct-ERIC. We thank PetaChem, LLC for providing a demo version of the TeraChem software. Calculations were performed on the Weizmann Institute's WEXAC and ChemFarm supercomputers.

ORCID

Joel L. Sussman https://orcid.org/0000-0003-0306-3878

REFERENCES

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1-2:19–25.

Adamo C, Barone V. Toward reliable density functional methods without adjustable parameters: the PBE0 model. J Chem Phys. 1999;110:6158–70.

Andreini C, Cavallaro G, Lorenzini S. FindGeo: a tool for determining metal coordination geometry. Bioinformatics. 2012;28:1658–60.

Babu CS, Dudev T, Lim C. Differential role of the protein matrix on the binding of a catalytic aspartate to Mg^{2+} *vs* Ca^{2+}: application to ribonuclease H. J Am Chem Soc. 2013;135:6541–8.

Bazelyansky M, Robey E, Kirsch JF. Fractional diffusion-limited component of reactions catalyzed by acetylcholinesterase. Biochemistry. 1986;25:125–30.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res. 2000;28:235–42.

Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput. 2012;8:3257–73.

Biswas T, Yi L, Aggarwal P, Wu J, Rubin JR, Stuckey JA, et al. The tail of KdsC: conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase. J Biol Chem. 2009;284:30594–603.

Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30:1545–614.

Bussi G, Tribello GA. Analyzing and biasing simulations with PLUMED. Methods Mol Biol. 2019;2022:529–78.

Chalupova K, Korabecny J, Bartolini M, Monti B, Lamba D, Caliandro R, et al. Novel tacrine-tryptophan hybrids: multi-target directed ligands as potential treatment for Alzheimer's disease. Eur J Med Chem. 2019;168:491–514.

Cia G, Kwasigroch J, Stamatopoulos B, Rooman M, Pucci F. pyScoMotif: discovery of similar 3D structural motifs across proteins. Bioinform Adv. 2023;3:vbad158.

Dauter Z. Estimation of anomalous signal in diffraction data. Acta Crystallogr D Biol Crystallogr. 2006;62:867–76.

Del Castillo J, Engbaek L. The nature of the neuromuscular block produced by magnesium. J Physiol. 1954;124:370–84.

Di Palma F, Bottaro S, Bussi G. Kissing loop interaction in adenine riboswitch: insights from umbrella sampling simulations. BMC Bioinf. 2015;16(Suppl 9):S6.

Dimelow RJ, Bryce RA, Masters AJ, Hillier IH, Burton NA. Exploring reaction pathways with transition path and umbrella sampling: application to methyl maltoside. J Chem Phys. 2006;124:114113.

Doudou S, Burton NA, Henchman RH. Standard free energy of binding from a one-dimensional potential of mean force. J Chem Theory Comput. 2009;5:909–18.

Dvir H, Jiang HL, Wong DM, Harel M, Chetrit M, He XC, et al. X-ray structures of *Torpedo californica* acetylcholinesterase complexed with (+)-huperzine a and (−)-huperzine B: structural evidence for an active site rearrangement. Biochemistry. 2002;41:10810–8.

Dym O, Song W, Felder C, Roth E, Shnyrov V, Ashani Y, et al. The impact of crystallization conditions on structure-based drug design: a case study on the methylene blue/acetylcholinesterase complex. Protein Sci. 2016;25:1096–114.

Efremov RG, Leitner A, Aebersold R, Raunser S. Architecture and conformational switch mechanism of the ryanodine receptor. Nature. 2015;517:39–43.

Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of coot. Acta Crystallogr D Biol Crystallogr. 2010;66:486–501.

Goldenzweig A, Goldsmith M, Hill SE, Gertman O, Laurino P, Ashani Y, et al. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. Mol Cell. 2016;63:337–46.

Greenblatt HM, Guillou C, Guenard D, Argaman A, Botti S, Badet B, et al. The complex of a bivalent derivative of galanthamine with *torpedo* acetylcholinesterase displays drastic deformation of the active-site gorge: implications for structure-based drug design. J Am Chem Soc. 2004;126:15405–11.

Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys. 2010;132:154104.

Gucwa M, Lenkiewicz J, Zheng H, Cymborowski M, Cooper DR, Murzyn K, et al. CMM-an enhanced platform for interactive validation of metal binding sites. Protein Sci. 2023;32:e4525.

Gumbart JC, Roux B, Chipot C. Efficient determination of protein-protein standard binding free energies from first principles. J Chem Theory Comput. 2013;9:3789–98.

Guzenko D, Burley SK, Duarte JM. Real time structural search of the protein data Bank. PLoS Comput Biol. 2020;16:e1007970.

Handing KB, Niedzialkowska E, Shabalin IG, Kuhn ML, Zheng H, Minor W. Characterizing metal-binding sites in proteins with X-ray crystallography. Nat Protoc. 2018;13:1062–90.

Harel M, Kleywegt GJ, Ravelli RB, Silman I, Sussman JL. Crystal structure of an acetylcholinesterase-fasciculin complex: interaction of a three-fingered toxin from snake venom with its target. Structure. 1995;3:1355–66.

Harel M, Sonoda LK, Silman I, Sussman JL, Rosenberry TL. Crystal structure of thioflavin T bound to the peripheral site of *Torpedo californica* acetylcholinesterase reveals how thioflavin T acts as a sensitive fluorescent reporter of ligand binding to the acylation site. J Am Chem Soc. 2008;130:7856–61.

Hsin J, Arkhipov A, Yin Y, Stone JE, Schulten K. Using VMD: an introductory tutorial. Curr Protoc Bioinformatics. 2008;Chapter 5:Unit 5.7.

Iskrenova-Tchoukova E, Kalinichev AG, Kirkpatrick RJ. Metal cation complexation with natural organic matter in aqueous solutions: molecular dynamics simulations and potentials of mean force. Langmuir. 2010;26:15909–19.

Jafari M, Li Z, Song LF, Sagresti L, Brancato G, Merz KM Jr. Thermodynamics of metal-acetate interactions. J Phys Chem B. 2024;128:684–97.

Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem. 2008;29:1859–65.

Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB_REDO server for macromolecular structure model optimization. IUCrJ. 2014;1:213–20.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–9.

Katz AK, Glusker JP, Beebe SA, Bock CW. Calcium ion coordination: a comparison with that of beryllium, magnesium, and zinc. J Am Chem Soc. 1996;118:5752–63.

Kellosalo J, Kajander T, Kogan K, Pokharel K, Goldman A. The structure and catalytic cycle of a sodium-pumping pyrophosphatase. Science. 2012;337:473–6.

Kreimer DI, Dolginova EA, Raves M, Sussman JL, Silman I, Weiner L. A metastable state of *Torpedo californica* acetylcholinesterase generated by modification with organomercurials. Biochemistry. 1994;33:14407–18.

Kreimer DI, Szosenfogel R, Goldfarb D, Silman I, Weiner L. Two-state transition between molten globule and unfolded states of acetylcholinesterase as monitored by electron paramagnetic resonance spectroscopy. Proc Natl Acad Sci U S A. 1994;91:12145–9.

Kumar S, Nussinov R. Relationship between ion pair geometries and electrostatic strengths in proteins. Biophys J. 2002;83:1595–612.

Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem. 1992;13:1011–21.

Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. Acta Crystallogr D Struct Biol. 2019;75:861–77.

Luehr N, Jin AG, Martinez TJ. *Ab initio* interactive molecular dynamics on graphical processing units (GPUs). J Chem Theory Comput. 2015;11:4536–44.

Melo MCR, Bernardi RC, Rudack T, Scheurer M, Riplinger C, Phillips JC, et al. NAMD goes quantum: an integrative suite for hybrid simulations. Nat Methods. 2018;15:351–4.

Millard CB, Shnyrov VL, Newstead S, Shin I, Roth E, Silman I, et al. Stabilization of a metastable state of *Torpedo californica* acetylcholinesterase by chemical chaperones. Protein Sci. 2003;12:2337–47.

Nadzirin N, Gardiner EJ, Willett P, Artymiuk PJ, Firdaus-Raih M. SPRITE and ASSAM: web servers for side chain 3D-motif searching in protein structures. Nucleic Acids Res. 2012;40:W380–6.

Nadzirin N, Willett P, Artymiuk PJ, Firdaus-Raih M. IMAAAGINE: a webserver for searching hypothetical 3D amino acid side chain arrangements in the protein data Bank. Nucleic Acids Res. 2013;41:W432–40.

Novichkova DA, Lushchekina SV, Dym O, Masson P, Silman I, Sussman JL. The four-helix bundle in cholinesterase dimers: structural and energetic determinants of stability. Chem Biol Interact. 2019;309:108699.

Olsson MH, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical p*Ka* predictions. J Chem Theory Comput. 2011;7:525–37.

Park SJ, Kern N, Brown T, Lee J, Im W. CHARMM-GUI PDB manipulator: various PDB structural modifications for biomolecular modeling and simulation. J Mol Biol. 2023;435:167995.

Parsons JF, Lim K, Tempczyk A, Krajewski W, Eisenstein E, Herzberg O. From structure to function: YrbI from *Haemophilus influenzae* (HI1679) is a phosphatase. Proteins. 2002;46:393–404.

Pedretti A, Mazzolari A, Gervasoni S, Fumagalli L, Vistoli G. The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. Bioinformatics. 2021;37:1174–5.

Permyakov EA. Metal binding proteins. Encyclopedia. 2021;1:261–92.

Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. J Chem Phys. 2020;153:044130.

Rardin RL, Tolman WB, Lippard SJ. Monodentate carboxylate complexes and the carboxylate shift: implications for polymetalloprotein structure and function. New J Chem. 1991;15:417–30.

Raves ML, Harel M, Pang YP, Silman I, Kozikowski AP, Sussman JL. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (−)-huperzine a. Nat Struct Biol. 1997;4:57–63.

Rigden DJ, Woodhead DD, Wong PW, Galperin MY. New structural and functional contexts of the dx[DN]xDG linear motif: insights into evolution of calcium-binding proteins. PLoS One. 2011;6:e21507.

Rosenberry TL. Acetylcholinesterase. Adv Enzymol Relat Areas Mol Biol. 1975;43:103–218.

Roux B. The calculation of the potential of mean force using computer simulations. Comput Phys Commun. 1995;91:275–82.

Sanson B, Colletier JP, Xu Y, Lang PT, Jiang H, Silman I, et al. Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and molecular dynamics simulations. Protein Sci. 2011;20:1114–8.

Santoni G, de Sousa J, de la Mora E, Dias J, Jean L, Sussman JL, et al. Structure-based optimization of nonquaternary reactivators of acetylcholinesterase inhibited by organophosphorus nerve agents. J Med Chem. 2018;61:7630–9.

Shen H, Campanello GC, Flicker D, Grabarek Z, Hu J, Luo C, et al. The human knockout gene CLYBL connects itaconate to vitamin B12. Cell. 2017;171:771–782.e11.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal omega. Mol Syst Biol. 2011;7:539.

Silman I. The multiple biological roles of the cholinesterases. Prog Biophys Mol Biol. 2021;162:41–56.

Silman I, Shnyrov VL, Ashani Y, Roth E, Nicolas A, Sussman JL, et al. *Torpedo californica* acetylcholinesterase is stabilized by binding of a divalent metal ion to a novel and versatile 4D motif. Protein Sci. 2021;30:966–81.

Stieler M, Weber J, Hils M, Kolb P, Heine A, Buchold C, et al. Structure of active coagulation factor XIII triggered by calcium binding: basis for the design of next-generation anticoagulants. Angew Chem Int ed Engl. 2013;52:11930–4.

Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, Toker L, et al. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. Science. 1991;253:872–9.

Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: new feathers for an old bird. Comput Phys Commun. 2014;185:604–13.

UniProt Consortium. UniProt: the universal protein knowledgebase in 2023. Nucleic Acids Res. 2023;51:D523–31.

Valdez CE, Smith QA, Nechay MR, Alexandrova AN. Mysteries of metals in metalloenzymes. Acc Chem Res. 2014;47:3110–7.

van Beusekom B, Touw WG, Tatineni M, Somani S, Rajagopal G, Luo J, et al. Homology-based hydrogen bond information improves crystallographic structures in the PDB. Protein Sci. 2018;27:798–808.

Wang X, Pinter MJ, Rich MM. Ca^{2+} dependence of the binomial parameters p and n at the mouse neuromuscular junction. J Neurophysiol. 2010;103:659–66.

White AD, Voth GA. Efficient and minimal method to bias molecular simulations with experimental data. J Chem Theory Comput. 2014;10:3023–30.

Zheng H, Chordia MD, Cooper DR, Chruszcz M, Muller P, Sheldrick GM, et al. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat Protoc. 2014;9:156–70.

Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB, Minor W. CheckMyMetal: a macromolecular metal-binding validation tool. Acta Crystallogr D Struct Biol. 2017;73:223–33.

SUPPORTING INFORMATION

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

How to cite this article: Lushchekina S, Weiner L, Ashani Y, Emrizal R, Firdaus-Raih M, Silman I, et al. Why is binding of a divalent metal cation to a structural motif containing four carboxylate residues not accompanied by a conformational change? Protein Science. 2024;33(12):e5206. https://doi.org/10.1002/pro.5206