Supporting Information

Vibrational and Molecular Properties of Mg^{2+} Binding and Ion Selectivity in the Magnesium Channel MgtE

Tetsunari Kimura[a, b, 1], *Victor A. Lorenz-Fonfria*[c, d], *Shintaro Douki*[e], *Hideyoshi Motoki*[f], *Ryuichiro Ishitani*[e], *Osamu Nureki*[e], *Masahiro Higashi*[f] & *Yuji Furutani*[a, b*]

[a]Department of Life and Coordination-Complex Molecular Science, Institute for Molecular Science, National Institutes of Natural Sciences, 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan

[b]Department of Structural Molecular Science, the Graduate University for Advanced Studies (SOKENDAI), 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan

[c]Institute of Molecular Science (ICMol), Universitat de València, Catedràtic José Beltrán Martínez 2, 46980 Paterna, Spain

[d]Department of Biochemistry and Molecular Biology, Universitat de València, Carrer Doctor Moliner 50, 46100 Burjassot, Spain

[e]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan

[f]Department of Chemistry, Biology and Marine Science, University of the Ryukyus, 1 Senbaru, Nishihara, Nakagami, Okinawa 903-0213, Japan

*Corresponding author: furutani@ims.ac.jp

[1] The present address is as follows. Department of Chemistry, Graduate School of Science, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan

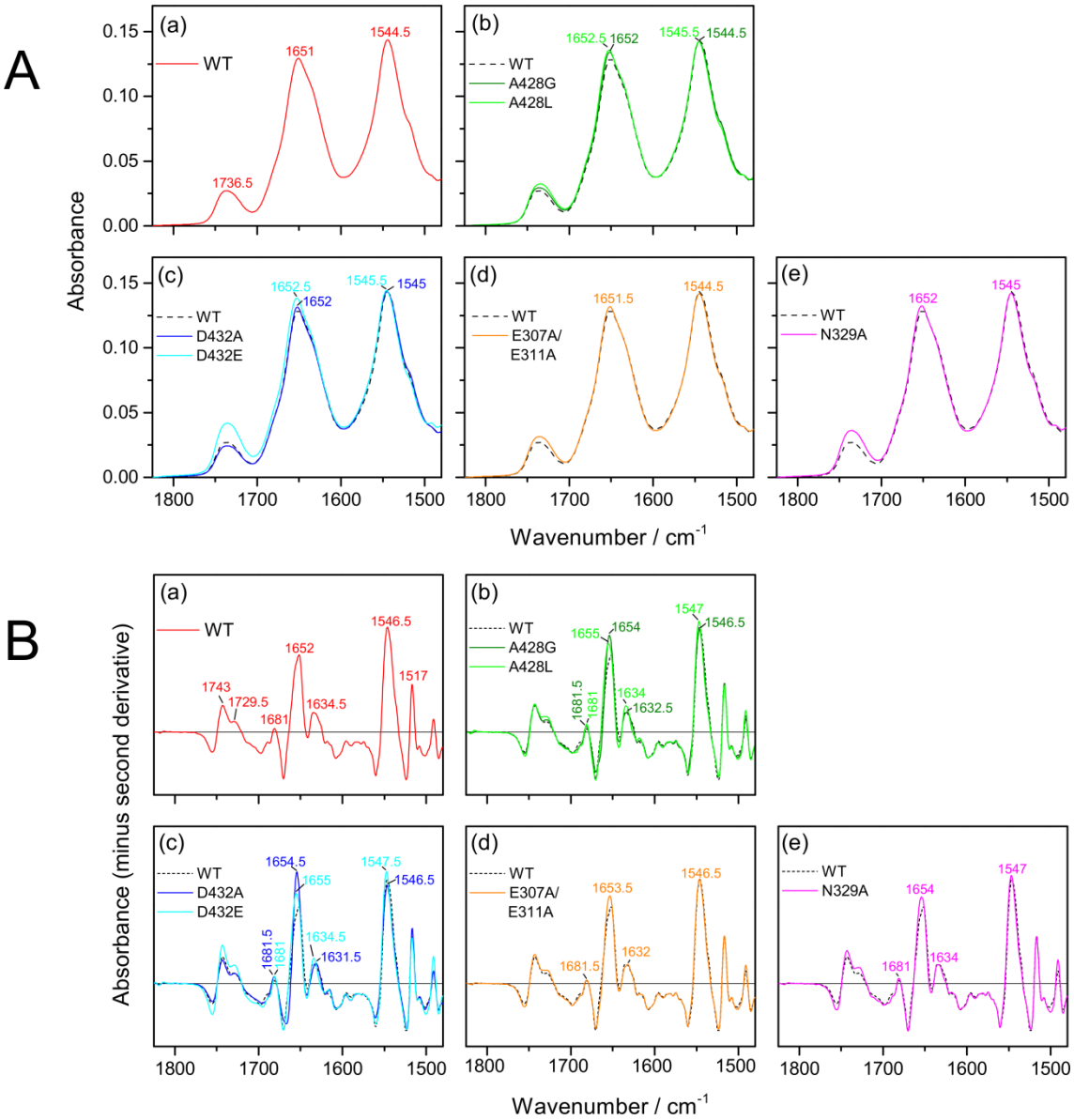

Figure S1. Absolute absorption IR spectra (A) and their second derivative spectra (B) of MgtE WT (a) and mutants (b; A428G and A428L, c; D432A and D432E, d; N329A, e; E307A/E311A) in the 1825-1480 cm^{-1} spectral region. In this spectral region we find main contributions from the C=O of phospholipids (~1737 cm^{-1}), amide I and amide II vibrations from the peptide backbone (~1651-53 and ~1544-45 cm^{-1}, respectively), and a vibration from tyrosine side chain at 1517 cm^{-1}. The second derivative spectra was computed in the Fourier domain at 4 cm^{-1} resolution as described.[1-2] Some additional bands are resolved, in particular bands in the amide I region at ~1681, ~1652-55 and ~1631-35 cm^{-1}, characteristic of loops, helices and beta sheet secondary structures, respectively.[3] The spectra of WT are reproduced in panels (b-e) of both (A) and (B) for comparison (dotted line). The absolute spectra of mutants were normalized to that of WT using the amide I and II bands (1700-1520 cm^{-1} region).

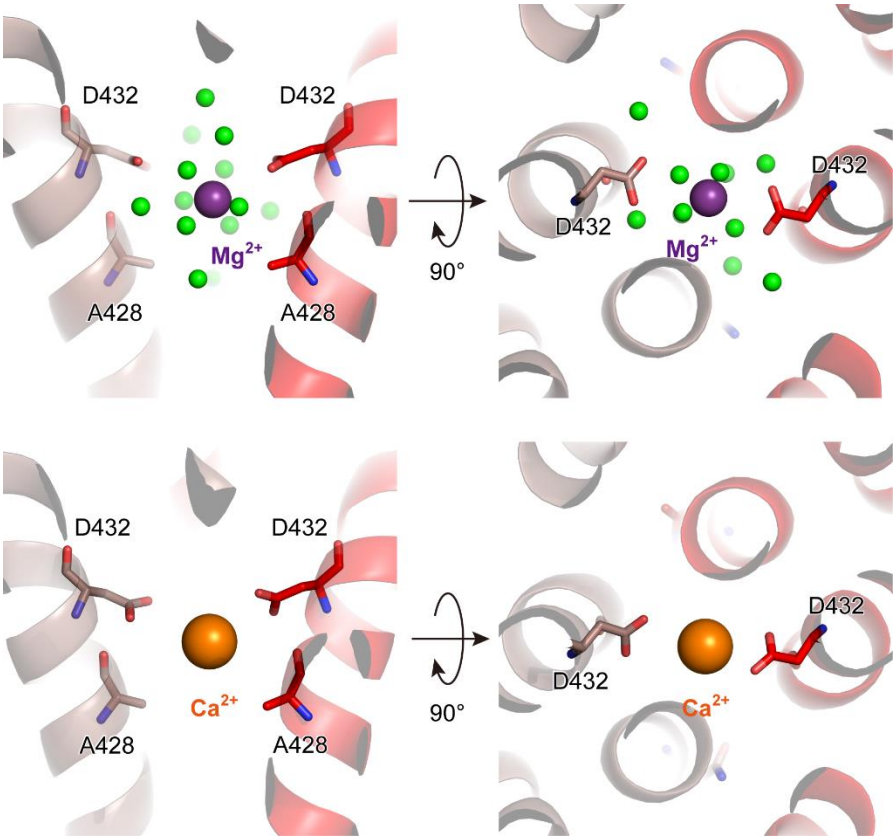

Figure S2. Molecular structure around the Mg^{2+} and Ca^{2+} binding site in the transmembrane region of MgtE (MgtE-TM), as described by the X-ray structures 4U9L for the Mg^{2+}-bound form and 4WIB for the Ca^{2+}-bound form, respectively.[4] (left; side view, right; top view). Mg^{2+} and Ca^{2+} ions are colored purple and orange, respectively. Green colored balls are oxygen atoms of water molecules. Asp432 and Ala428 residues are shown as stick model.

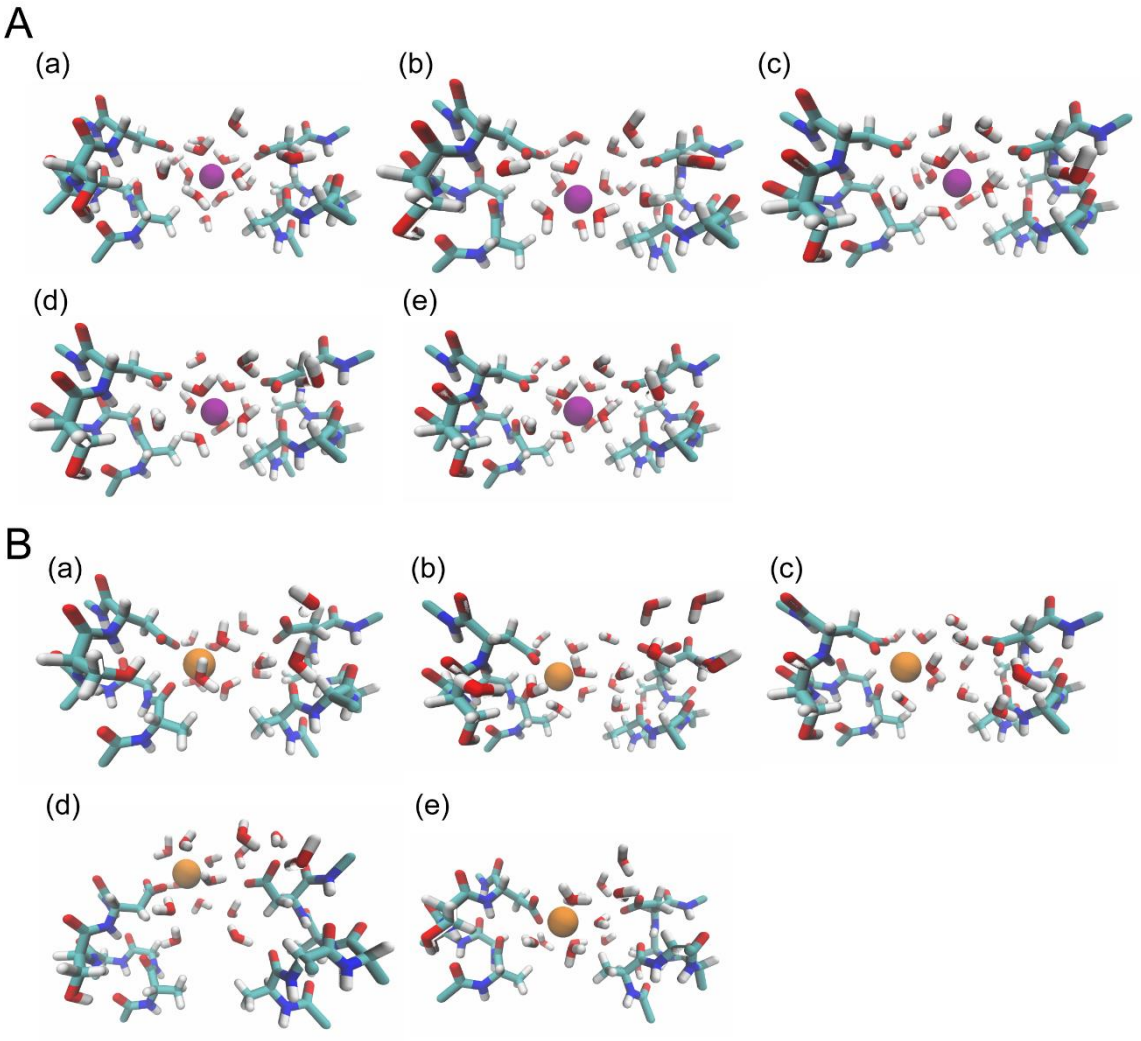

Figure S3. Snapshots utilized for vibrational analysis of the carboxyl groups of Asp432 interacting with Mg^{2+} (A) or Ca^{2+} (B). These snapshots were obtained after equilibration of molecular dynamics simulations for 1 μs (see Figure 6 in the manuscript). The snapshots of (a)-(e) correspond to the equilibration of the trajectories 1-5, respectively.

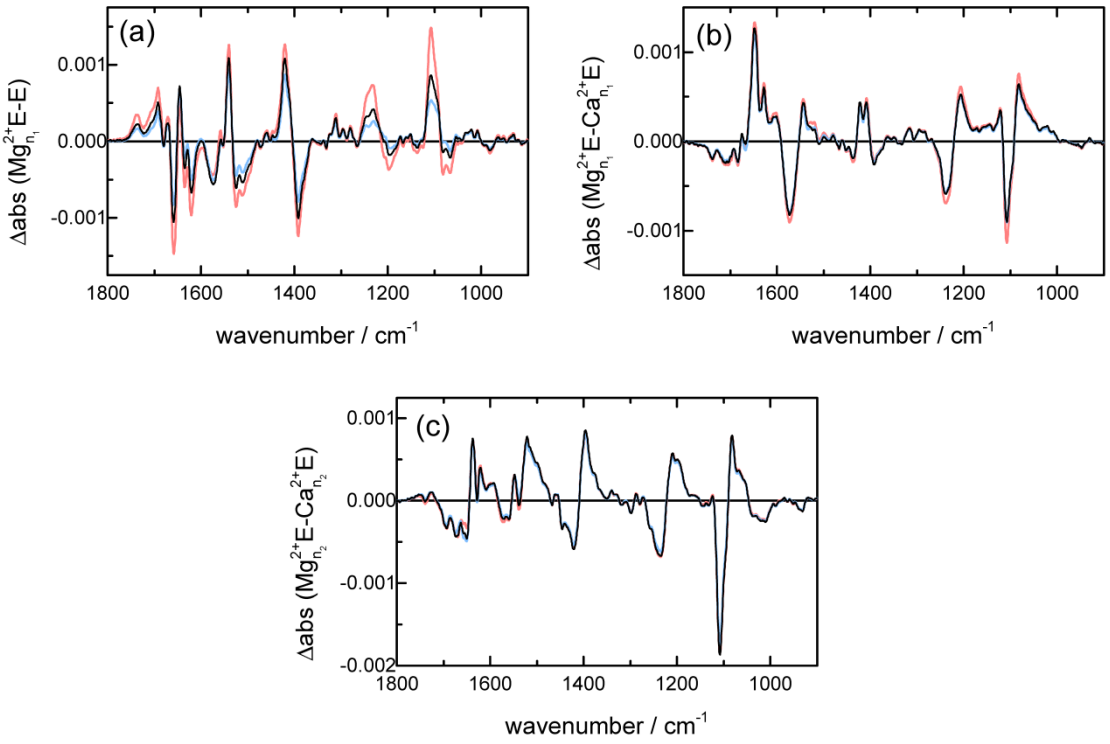

Figure S4. Illustration of how the uncertainty in the K_d for Ca^{+2} binding at site 1 (Table 1 of the MS) has actually a little effect on the estimated associated spectra for the binding of Mg^{2+} and Ca^{+2} at sites 1 and 2. We display estimated associated spectra using optimized binding parameters (black lines, reproducing spectra in Fig. 5 of the MS), as well as associated spectra obtained with the K_d for Ca^{+2} binding to site 1 restricted to its low and high confidence value (red and blue spectra, respectively) following by optimization of the rest of parameters (see Table S2 for final parameter values). (a) Associated spectra for Mg^{2+} binding at site 1. (b) Associated spectra for the exchange of Ca^{+2} with Mg^{2+} at site 1. (c) Associated spectrum for the exchange of Ca^{+2} with Mg^{2+} at site 2. Spectra in (a), (b) and (c) correspond to $\Delta Abs_{Mg^{2+}_{n1}\mathrm{E-E}}$, $\Delta A_{Mg^{2+}_{n_1}E-Ca^{2+}_{m_1}\mathrm{E}}$, and $\Delta A_{Mg^{2+}_{n_2}E-Ca^{2+}_{m_2}\mathrm{E}}$ as described in the Materials and Methods of the MS and the SI.

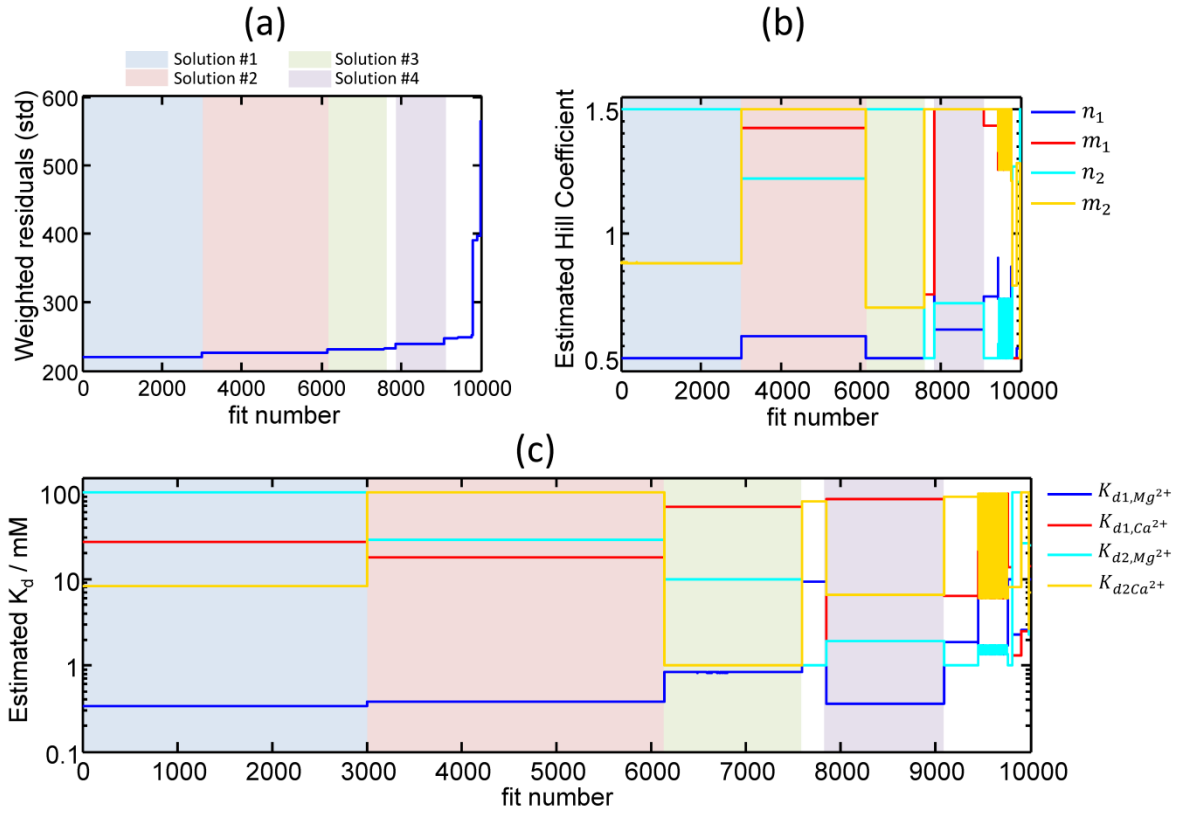

Figure S5. Exploration of multiple solutions when performing a global fitting of the experimental data set. The binding data set was globally fitted minimizing the standard deviation of the weighted residuals between the data and the model. As a model we used two independent binding sites to which both Mg^{2+} and Ca^{2+} compete. We used 10,000 random values for the binding parameters as starting points for the global fitting, restricting the Hill coefficients to be between 0.5 and 1.5, and the binding constants to be between 1 mM and 100 mM, except for the binding constant for Mg^{2+} to site 1 ($K_{d1,Mg^{2+}}$), restricted to be between 0.01 mM and 10 mM. Only solutions where $K_{d1,Mg^{2+}}$ was the smaller binding constant were considered acceptable. (a) Fitting solutions sorted by increasing standard deviation of the weighted residuals. (b) Fitted Hill coefficients. (c) Fitted binding constants. Four acceptable solutions were found, color-shaded in (a-c). See Table S3 for numerical values of the binding parameters for these four solutions. Associated spectra for these four solutions are presented in Fig. S6.

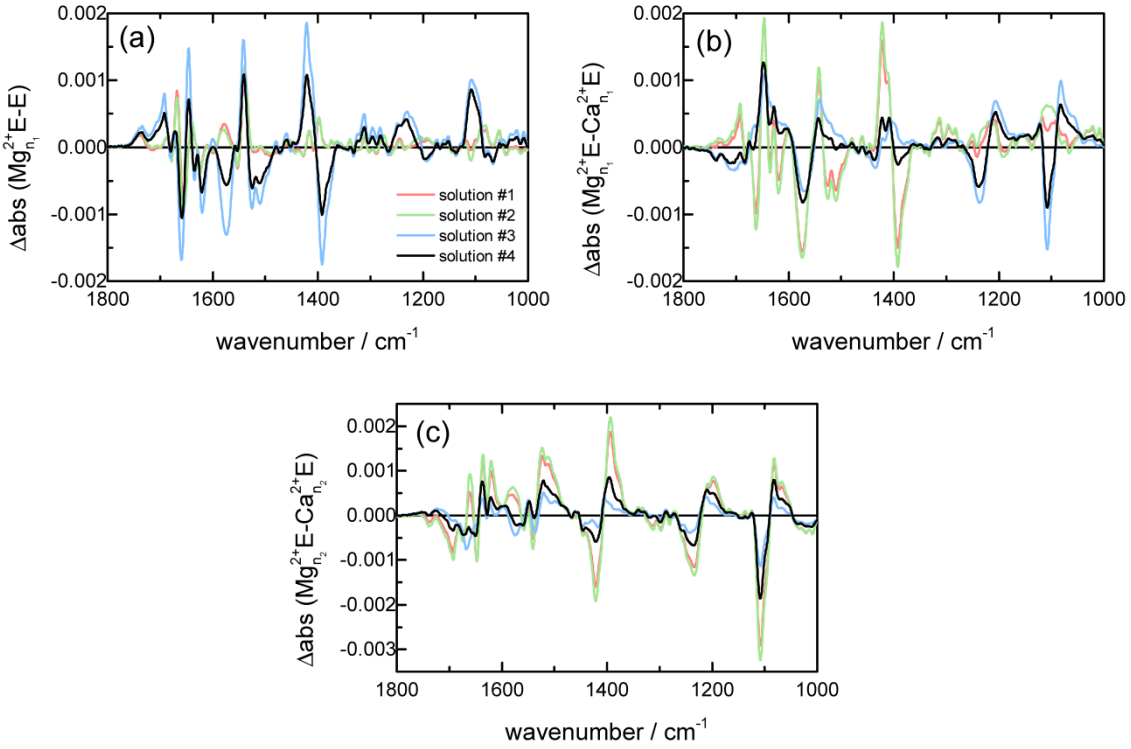

Figure S6. Associated binding spectra for the four possible solutions found by global fitting (see Fig. S5). Solution #4 corresponds to that presented in Fig. 5 of the MS. (a) Associated spectra for Mg^{2+} binding at site 1. (b) Associated spectra for the exchange of Ca^{+2} with Mg^{2+} at site 1. (c) Associated spectrum for the exchange of Ca^{+2} with Mg^{2+} at site 2. Spectra in (a), (b) and (c) correspond to $\Delta Abs_{Mg_{n1}^{2+}\mathrm{E-E}}$, $\Delta A_{Mg_{n_1}^{2+}E-Ca_{m_1}^{2+}\mathrm{E}}$, and $\Delta A_{Mg_{n_2}^{2+}E-Ca_{m_2}^{2+}\mathrm{E}}$ as described in the Materials and Methods of the MS and the SI.

Table S1. Initial and final concentrations of Mg^{2+} and Ca^{2+} in the titration experiments of the four data sets displayed in Figure 3 of the manuscript. The concentrations are given in mM.

Data set	$[Mg^{2+}]_f$	$[Mg^{2+}]_i$	$[Ca^{2+}]_f$	$[Ca^{2+}]_i$
#1	0	0	10	10
#1	0.2	0	9.8	10
#1	0.5	0	9.5	10
#1	1	0	9	10
#1	1.5	0	8.5	10
#1	2.5	0	7.5	10
#1	5	0	5	10
#1	7.5	0	2.5	10
#1	10	0	0	10
#2	0	10	10	0
#2	1.00E-03	10	9.999	0
#2	0.01	10	9.99	0
#2	0.02	10	9.98	0
#2	0.05	10	9.95	0
#2	0.1	10	9.9	0
#2	0.23	10	9.77	0
#2	0.36	10	9.64	0
#2	0.5	10	9.5	0
#2	1	10	9	0
#2	1.5	10	8.5	0
#2	2.5	10	7.5	0
#2	3.5	10	6.5	0
#2	5	10	5	0
#2	6	10	4	0
#2	7.5	10	2.5	0
#2	8.5	10	1.5	0
#2	10	10	0	0
#3	0.01	100	99.99	0
#3	0.1	100	99.9	0
#3	1	100	99	0
#3	10	100	90	0
#3	50	100	50	0
#3	100	100	0	0
#4	0.2	0	0	0.2
#4	0.5	0	0	0.5
#4	1	0	0	1
#4	1.4	0	0	1.4
#4	2	0	0	2
#4	2.75	0	0	2.75
#4	3.7	0	0	3.7
#4	5	0	0	5
#4	7	0	0	7

Table S2. Optimized parameters for the binding parameters (reproducing those in Table 1 of the MS), and two suboptimal sets of binding parameters obtained by fixing the K_d value for Ca^{2+} at site 1 to its lower and higher confidence limit. The rest of parameters were optimized, but constrained to take values within the confidence limits reported in Table 1 of the MS. The suboptimal parameters were used to visualize in Fig. S4 how the large uncertainties for the K_d value for Ca^{2+} at site 1 affects the estimated binding associated spectra.

			Optimized values	Suboptimal values (low $K_{d1,Ca^{2+}}$)	Suboptimal values (high $K_{d1,Ca^{2+}}$)
Site 1	Mg^{2+}	K_d (mM)	0.36	0.31	0.28
		n	0.61	0.59	0.70
	Ca^{2+}	K_d (mM)	85	14*	520*
		m	1.5**	1.0	2.1**
Site 2	Mg^{2+}	K_d (mM)	1.9	1.4**	2.0
		n	0.7	0.7	0.7
	Ca^{2+}	K_d (mM)	6.5	7.9	9.0**
		m	1.5**	1.8*	1.7

*Parameter fixed during the optimization process.

**Parameter hitting a lower or higher constrain during the optimization process.

Table S3. Binding parameters for the four solutions found by global fitting with random initial starting values (see Fig. S5). Confidence intervals (96%) from asymptotic errors are provided in the brackets. Solution #4 corresponds to the solution presented in Table 1 of the MS. Hill Coefficients (n and m) were constrained to take values between 0.5 and 1.5, while binding constants were constrained to take values between 1 mM and 100 mM, except for Mg^{2+} binding at site 1, constrained between 0.01 mM and 10 mM.

			Solution #1	Solution #2	Solution #3	Solution #4
Site 1	Mg^{2+}	K_d (mM)	0.34	0.38	0.83 [0.41-1.7]	0.36
			[0.19-0.60]	[0.21-0.69]		[0.16-0.83]
		n	0.5* [0.4-0.6]	0.6 [0.5-0.7]	0.5* [0.4-0.6]	0.6 [0.4-0.8]
	Ca^{2+}	K_d (mM)	27 [8.8-83]	18 [10-30]	68 [10- 470]	85 [14-520]
		m	1.5* [1.1-1.9]	1.4 [1.2-1.6]	1.5* [0.9-2.1]	1.5* [0.9-2.1]
Site 2	Mg^{2+}	K_d (mM)	100* [28-360]	28 [14-57]	9.8 [5.5-17]	1.9 [1.4-2.7]
		n	1.5* [1.0-2.0]	1.2 [1.0-1.5]	1.5* [1.2-1.8]	0.7 [0.6-0.9]
	Ca^{2+}	K_d (mM)	8.4 [4.0-18]	100* [33-300]	1.0* [0.6-1.7]	6.5 [4.7-9.0]
		m	0.9 [0.7-1.1]	1.5* [1.1-1.9]	0.7 [0.6-0.8]	1.5* [1.2-1.8]

*Parameter hitting a lower or higher constrain during the optimization process.

Supporting Methods

Derivation of Equation 1. We first considered that the protein contains one binding site, where Mg^{2+} and Ca^{2+} bind competitively:

$$\mathrm{n}Mg^{2+} + \mathrm{E} \overset{K_{d,Mg^{2+}}}{\longleftrightarrow} Mg_n^{2+}\mathrm{E} \tag{1a}$$

$$\mathrm{m}Ca^{2+} + \mathrm{E} \overset{K_{d,Ca^{2+}}}{\longleftrightarrow} Ca_m^{2+}\mathrm{E} \tag{1b}$$

The binding is characterized by two dissociation constants, $K_{d,Mg^{2+}}$ and $K_{d,Ca^{2+}}$, and two Hill coefficients, n and m:

$$K_{d,Mg^{2+}} = \frac{[Mg^{2+}]^n[\mathrm{E}]}{[Mg_n^{2+}\mathrm{E}]} \tag{2a}$$

$$K_{d,Ca^{2+}} = \frac{[Ca^{2+}]^m[\mathrm{E}]}{[Ca_m^{2+}\mathrm{E}]} \tag{2b}$$

where E represents the protein with a free (unoccupied) binding site, $Mg_n^{2+}\mathrm{E}$ the protein bound to Mg^{2+} and $Ca_m^{2+}\mathrm{E}$ the protein bound to Ca^{2+}.

The absorbance of the sample at a given Mg^{2+} and Ca^{2+} concentrations, $[Mg^{2+}], [Ca^{2+}]$, is:

$$Abs([Mg^{2+}], [Ca^{2+}]) = Abs_E \times f_E^{[Mg^{2+}],[Ca^{2+}]} + Abs_{Mg_{n1}^{2+}\mathrm{E}} \times f_{Mg_{n1}^{2+}E}^{[Mg^{2+}],[Ca^{2+}]} +$$

$$+Abs_{Ca_{m1}^{2+}\mathrm{E}} \times f_{Ca_{m1}^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} \tag{3}$$

where Abs_E, $Abs_{Mg_{n1}^{2+}\mathrm{E}}$ and $Abs_{Ca_{m1}^{2+}\mathrm{E}}$ are the absorbance of the different species present: free protein, protein bound to Mg^{2+}, and protein bound to Ca^{2+}; and f stands for their respective fractions:

$$f_E^{[Mg^{2+}],[Ca^{2+}]} = \frac{[\mathrm{E}]}{[\mathrm{E}]_T} \tag{4a}$$

$$f_{Mg_n^{2+}E}^{[Mg^{2+}],[Ca^{2+}]} = \frac{[Mg_n^{2+}\mathrm{E}]}{[\mathrm{E}]_T} \tag{4b}$$

$$f_{Ca_m^{2+}E}^{[Mg^{2+}],[Ca^{2+}]} = \frac{[Ca_m^{2+}\mathrm{E}]}{[\mathrm{E}]_T} \tag{4c}$$

In Eq. 4 $[\mathrm{E}]_T$ represents the total concentration of the binding site. Because all the fractions add up to one, we can express the fraction of the free enzyme in Eq. 3 as a function of the other two fractions:

$$Abs([Mg^{2+}], [Ca^{2+}])$$

$$= Abs_E \times \left(1 - f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} - f_{Ca_m^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]}\right) + Abs_{Mg_n^{2+}\mathrm{E}} \times f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} +$$

$$+Abs_{Ca_m^{2+}\mathrm{E}} \times f_{Ca_m^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} \quad (5)$$

Rearranging Eq. 5 we arrive to:

$$Abs([Mg^{2+}],[Ca^{2+}]) = Abs_E + \Delta Abs_{Mg_n^{2+}\mathrm{E-E}} \times f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} + \Delta Abs_{Ca_m^{2+}\mathrm{E-E}} \times f_{Ca_m^{2+}\mathrm{E}}^{[Mg^{2+}],[Ca^{2+}]} \quad (6)$$

where $\Delta Abs_{Mg_n^{2+}\mathrm{E-E}}$ stands for the difference in absorbance induced when Mg^{2+} binds to the free/empty binding site of the protein, and $\Delta Abs_{Ca_m^{2+}\mathrm{E-E}}$ when Ca^{2+} does it.

The difference in absorbance when changing the Mg^{2+} and Ca^{2+} concentration from a set of initial values $[Mg^{2+}]_i$ and $[Ca^{2+}]_i$, to a set of final values $[Mg^{2+}]_f$ and $[Ca^{2+}]_f$, will be:

$$\Delta Abs = Abs\left([Mg^{2+}]_f,[Ca^{2+}]_f\right) - Abs([Mg^{2+}]_i,[Ca^{2+}]_i)$$

$$= \Delta Abs_{Mg_n^{2+}\mathrm{E-E}} \times \left(f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) +$$

$$+\Delta Abs_{Ca_m^{2+}\mathrm{E-E}} \times \left(f_{Ca_m^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Ca_m^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) \quad (7)$$

When the initial and final Mg^{2+} and Ca^{2+} concentrations are such that the fraction of protein with unoccupied binding site is negligible, a condition presumably meet in most (but not all) of our experimental conditions, Eq. 7 can be simplified:

$$\Delta Abs \approx \Delta Abs_{Mg_n^{2+}\mathrm{E}-Ca_m^{2+}\mathrm{E}} \times \left(f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Mg_n^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) \quad (8)$$

where $\Delta Abs_{Mg_n^{2+}\mathrm{E}-Ca_m^{2+}\mathrm{E}}$ stands for the difference in absorbance induced when Mg^{2+} substitutes Ca^{2+} in the binding site. Although, this expression does not apply to all of our experimental conditions, it explains why for the binding site 2 of MgtE $\Delta Abs_{Mg_{n2}^{2+}\mathrm{E-E}}$ and $\Delta Abs_{Ca_{m2}^{2+}\mathrm{E-E}}$ are harder to estimate and more prone to errors than their difference $\Delta Abs_{Mg_{n2}^{2+}\mathrm{E}-Ca_{m2}^{2+}\mathrm{E}}$.

To be able to use Eq. 7 as a fitting model we need an expression of the fraction of protein bond to Mg^{2+}:

$$f_{Mg_n^{2+}E}^{[Mg^{2+}],[Ca^{2+}]} = \frac{[Mg_n^{2+}\mathrm{E}]}{[\mathrm{E}]_T} = \frac{[Mg_n^{2+}\mathrm{E}]}{[\mathrm{E}]+[Mg_n^{2+}\mathrm{E}]+[Ca_m^{2+}\mathrm{E}]} \quad (9)$$

where we have taken into account that:

$$[\mathrm{E}]_T = [\mathrm{E}] + [Mg_n^{2+}\mathrm{E}] + [Ca_m^{2+}\mathrm{E}] \quad (10)$$

We need to reexpress Eq. 9 as a function of the Mg^{2+} and Ca^{2+} concentrations and their binding parameters. Rearrenging Eq. 2a we can express [E] as:

$$[\mathrm{E}] = \frac{K_{d,Mg^{2+}}[Mg_n^{2+}\mathrm{E}]}{[Mg^{2+}]^n} \quad (11)$$

Rearrenging Eq. 2b we can express $[Ca_m^{2+}\mathrm{E}]$, after substituting for [E] using Eq. 11, as:

$$[Ca_m^{2+}\mathrm{E}] = \frac{[Ca^{2+}]^m[\mathrm{E}]}{K_{d,Ca^{2+}}} = \frac{[Ca^{2+}]^m}{K_{d,Ca^{2+}}} \times \frac{K_{d,Mg^{2+}}[Mg_n^{2+}\mathrm{E}]}{[Mg^{2+}]^n} = \frac{K_{d,Mg^{2+}}[Ca^{2+}]^m}{K_{d,Ca^{2+}}[Mg^{2+}]^n}[Mg_n^{2+}\mathrm{E}] \quad (12)$$

Taking into account Eq. 10 and Eq. 11, Eq. 8 can now be repressed as

$$f_{Mg_n^{2+}E}^{[Mg^{2+}],[Ca^{2+}]} = \frac{1}{1+\frac{K_{d,Mg^{2+}}}{[Mg^{2+}]^n}+\frac{K_{d,Mg^{2+}}[Ca^{2+}]^m}{K_{d,Ca^{2+}}[Mg^{2+}]^n}} = \frac{1}{1+\frac{K_{d,Mg^{2+}}}{[Mg^{2+}]^n}\left(1+\frac{[Ca^{2+}]^m}{K_{d,Ca^{2+}}}\right)} \quad (13)$$

Likewise, the fraction of protein bound to Ca^{2+} can be expressed as:

$$f_{Ca_m^{2+}E}^{x,y} = \frac{1}{1+\frac{K_{d,Ca^{2+}}}{[Ca^{2+}]^m}\left(1+\frac{[Mg^{2+}]^n}{K_{d,Mg^{2+}}}\right)} \quad (14)$$

Combining Eq. 7 and Eq. 13 and Eq.14 we have an explicit fitting model when Mg^{2+} and Ca^{2+} compite for one binding site.

We know that in the protein MgtE there are more than a single binding site for Mg^{2+} and Ca^{2+}. Singular value decomposition indicates that the experimental absorption changes as a function of the Mg^{2+} and Ca^{2+} concentration contain at least three independent spectra, implying that at least a model with two binding sites is required to describe the experimental data. The result from singular value decomposition also implies that the data does not contain enough information to attempt to fit three or more binding sites. Besides considering that there are only two binding sites, for simplicity we assumed that these two sites are independent, so the occupancy of one of the binding sites does not affect the properties (K_d and Hill coefficient) of the other binding site. In other words, we ignore possible allosteric effects. The four binding events are:

$$\mathrm{n}_1 Mg^{2+} + \mathrm{E} \overset{K_{d1,Mg^{2+}}}{\longleftrightarrow} Mg_{n_1}^{2+}\mathrm{E} \quad (15a)$$

$$\mathrm{m}_1 Ca^{2+} + \mathrm{E} \overset{K_{d1,Ca^{2+}}}{\longleftrightarrow} Ca_{m_1}^{2+}\mathrm{E} \quad (15b)$$

$$\mathrm{n}_2 Mg^{2+} + \mathrm{E} \overset{K_{d2,Mg^{2+}}}{\longleftrightarrow} Mg_{n_2}^{2+}\mathrm{E} \quad (15c)$$

$$\mathrm{m}_2 Ca^{2+} + \mathrm{E} \overset{K_{d2,Ca^{2+}}}{\longleftrightarrow} Ca_{m_2}^{2+}\mathrm{E} \quad (15d)$$

Now, we can derive a similar equation as Eq. 7, but for two binding sites:

$$\Delta Abs = \Delta Abs_{Mg_{n1}^{2+}\mathrm{E}-\mathrm{E}} \times \left(f_{Mg_{n1}^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Mg_{n1}^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) + \Delta Abs_{Ca_{m1}^{2+}\mathrm{E}-\mathrm{E}} \times$$

$$\left(f_{Ca_{m1}^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Ca_{m1}^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) + \Delta Abs_{Mg_{n2}^{2+}\mathrm{E}-\mathrm{E}} \times \left(f_{Mg_{n2}^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Mg_{n2}^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right) +$$

$$\Delta Abs_{Ca_{m2}^{2+}\mathrm{E}-\mathrm{E}} \times \left(f_{Ca_{m2}^{2+}\mathrm{E}}^{[Mg^{2+}]_f,[Ca^{2+}]_f} - f_{Ca_{m2}^{2+}\mathrm{E}}^{[Mg^{2+}]_i,[Ca^{2+}]_i}\right)$$

(16)

Global fitting. Our fitting model is Eq. 16, with the fraction of the occupancy of the binding sites given by Eq. 13 and Eq. 14. We have 8 nonlinear parameters, four for the first binding site: $K_{d1,Mg^{2+}}$, n_1, $K_{d1,Ca^{2+}}$, and m_1; and another four for the second binding site: $K_{d2,Mg^{2+}}$, n_2, $K_{d2,Ca^{2+}}$, and m_2. As a linear parameters we have four spectra, $\Delta Abs_{Mg^{2+}_{n1}\mathrm{E-E}}$, $\Delta Abs_{Ca^{2+}_{m1}\mathrm{E-E}}$, $\Delta Abs_{Mg^{2+}_{n2}\mathrm{E-E}}$ and $\Delta Abs_{Ca^{2+}_{m2}\mathrm{E-E}}$.

As done in Eq. 8, Eq. 16 can be simplified considering that under the experimental conditions used the binding sites are almost ocupied by either Mg^{2+} or Ca^{2+}. In this case the data only depends on the spectral differences induced by Mg^{2+} substituting Ca^{2+} in either the binding site 1 or the binding site 2. The number of nonlinear parameters remains unchanged, but the linear parameters get reduced to two spectra: $\Delta Abs_{Mg^{2+}_{n1}\mathrm{E}-Ca^{2+}_{m1}\mathrm{E}}$ and $\Delta Abs_{Mg^{2+}_{n2}\mathrm{E}-Ca^{2+}_{m2}\mathrm{E}}$. Although such approximation simplifies the fitting function, we used the most general fitting functon in Eq. 16. Then, from the estimated $\Delta Abs_{Mg^{2+}_{n1}\mathrm{E-E}}$ and $\Delta Abs_{Ca^{2+}_{m1}\mathrm{E-E}}$ we obtained $\Delta Abs_{Mg^{2+}_{n1}\mathrm{E}-Ca^{2+}_{m1}\mathrm{E}}$, and from the estimated $\Delta Abs_{Mg^{2+}_{n2}\mathrm{E-E}}$ and $\Delta Abs_{Ca^{2+}_{m2}\mathrm{E-E}}$ we obtained $\Delta Abs_{Mg^{2+}_{n2}\mathrm{E}-Ca^{2+}_{m2}\mathrm{E}}$.

To globally fit the experimental data we first created and extended experimental data matrix, which combines our four data sets:

$$\mathbf{D} = [\mathbf{D1}\ \ \mathbf{D2}\ \ \mathbf{D3}\ \ \mathbf{D4}] \tag{17}$$

where the data matrix **D** contains n_w rows and n_c columns, where n_w is the number of wavenumbers and n_c the combined number of concentrations conditions.

The fitting function in Eq. 16 can be described as the multiplication of two matrices. One containing the spectral amplitudes and the other the occupancies of the two binding sites for Mg^{+2} and Ca^{+2}:

$$\mathbf{F} = \mathbf{A} \times \mathbf{C} \tag{18}$$

where **A** is a $n_w \times 4$ matrix, containing as columns the four "pure" difference spectra, and **C** is a $4 \times n_c$ matrix whose rows give the change in occupancy of the two binding sites to respect the initial and final Mg^{2+} and Ca^{2+} concentration as a function of the final Mg^{2+} and Ca^{2+} concentrations (see Eq. 13-14).

The nonlinear parameters ($K_{d1,Mg^{2+}}$, $K_{d1,Ca^{2+}}$, $K_{d2,Mg^{2+}}$, $K_{d2,Ca^{2+}}$, n_1, m_1, n_2, and m_2) as well as the elements of the matrix **A** can be estimated as those minimizing the square difference between the data and the fit:

$$\min\|\mathbf{D} - \mathbf{F}\| = \min\|\mathbf{D} - \mathbf{A} \times \mathbf{C}\| = \min\|\mathbf{D} - ((\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{D}) \times \mathbf{C}\| \tag{19}$$

where we used the fact that for a given matrix **C**, the matrix **A** that minimizes Eq. 19 is:

$$\mathbf{A} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{D} \tag{20}$$

Actually, we did not directly minimized Eq. 19, but we first performed SVD on the data matrix:

$$\mathbf{D} = \mathbf{U} \times \mathbf{S} \times \mathbf{V} \tag{22}$$

where the columns of **U** and the rows of **V** are orthonormal vectors, and **S** is a diagonal matrix with diagonal elements decreasing in amplitude. The data matrix was approximated by four singular

values, i.e., by the four first columns of **U** (size $n_w \times 4$), the first four rows of **V** (size $n_c \times 4$), and the first four elements of **S** (size 4×4). Instead of fitting the data matrix **D** ($n_w \times n_c$) we fitted the much smaller matrix $\mathbf{V}_{4\times n_c}$ ($4\times n_c$):

$$\min\left\|\mathbf{V}_{4\times n_c} - \mathbf{B}\times\mathbf{C}\right\| = \min\left\|\mathbf{V}_{4\times n_c} - \left((\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{V}_{4\times n_c}\right)\times\mathbf{C}\right\| \quad (22)$$

where **B** is a 4×4 matrix containing amplitudes. The nonlinear parameters were estimated minimizing Eq. 22, using the Matlab function "lsqnonlin", which implements the Levenberg-Marquardt method with a numerical calculation of the derivatives. The matrix **C**, corresponding to the best fitting parameters, was used to obtain the amplitude matrix **A** applying Eq. 20. We found, however, that this approach often leads to unrealistically intense amplitude spectra of opposed sign, a signature of over-fitting (caused by an attempt to fit the data more than necessarily given the noise level). To prevent this problem, we used a regularized version of Eq. 20, finding the amplitude matrix **A** of smaller second norm that best fit the data for a current value of **C**:

$$\mathbf{A} = (\mathbf{C}^T\mathbf{C} + \lambda\mathbf{I})^{-1}\mathbf{C}^T\mathbf{D} \quad (23)$$

where **I** is the identity matrix (4×4), and λ the regularization parameter. Consequently, the function to be minimized became:

$$\min\left\|\mathbf{V}_{4\times n_c} - \left((\mathbf{C}^T\mathbf{C} + \lambda\mathbf{I})^{-1}\mathbf{C}^T\mathbf{V}_{4\times n_c}\right)\times\mathbf{C}\right\| \quad (24)$$

The value for the regularization parameter was selected by trial and error, as the higher number giving a fit which does not significantly differs from the best fit in the absence of regularization. Confidence intervals were obtained from plus/minus two times the standard asymptotic errors. We used the logarithm of the dissociation constants instead of the dissociation constants themselves as fitting parameters. This choice ensured that all the dissociation constant remained positive during the iterative fitting and, furthermore, made the confidence intervals for the dissociation constants more realistic.

Because the estimated solution obtained from nonlinear least-squares can depend on the initial guess for the nonlinear parameters, in particular for a complex fitting problem as the present one, the fitting was conducted using 10,000 randomly generated initial guesses: (Hill coefficients from 0.5 to 1.5, and the K_d values from 1 to 100 mM, , except for the K_d of Mg^{2+} at site 1 where we used values between 0.01 mM and 10 mM). In addition, during the fit the dissociation constants and the Hill coefficients were constrained, not allowed to take values outside the above mentioned intervals. In addition, solutions where binding of Mg^{2+} at site 1 was not the one with higher affinity (lower K_d) were ignored. All the fittings converged to four acceptable solutions (Figure S4), whose parameters are provided in Table S3, and their associated binding spectra are displayed in Figure S5. From these four solutions, we selected the solution number 4 as the most soundly one given previous knowledge about the binding properties of MgtE as well as judging the shape and intensity of the estimated binding spectra. More specifically, the solution 1 and 2 hit the maximum value of K_d value at site 2 for Mg^{2+} and Ca^{2+}, respectively. Because MgtE chanel is closed at 10 mM Mg^{2+} or 20 mM Ca^{2+}, the

K_d values for site 2 are expected to be lower than 10 or 20 mM, respectively. Thus, we rejected these solutions. The solution 3 and 4 has K_d values of site 2 for Mg^{2+} and Ca^{2+} within this criteria, but the solution 3 has extremely low K_d value for Ca^{2+} (1.0 mM), which is not consistent with the electrophisological experiment. From these considerations, we chose the solution 4.

Legends for Supporting Movies

Movie S1. Molecular dynamics simulation of MgtE with a Mg^{2+} ion at Mg1 site

Movie S2. Molecular dynamics simulation of MgtE with a Ca^{2+} ion at Mg1 site

These movies are constructed from 500 snapshots retrieved in 2 ns interval and is played with the speed of 15 flames/s and for 32 s in total.

Supporting References

(1) Cameron, D. G.; Moffatt, D. J., *Appl. Spectrosc.* **1987,** *41*.

(2) Lórenz-Fonfría, V. A.; Padrós, E., *Spectrochim Acta A* **2004,** *60*.

(3) Barth, A., *Biochim. Biophys. Acta* **2007,** *1767*, 1073.

(4) Takeda, H.; Hattori, M.; Nishizawa, T.; Yamashita, K.; Shah, S. T.; Caffrey, M.; Maturana, A. D.; Ishitani, R.; Nureki, O., *Nature communications* **2014,** *5*.