DESC: An Automated Strategy to Efficiently Account for Dynamic Environment Effects in Solution

Albert Masip-Sánchez, Josep M. Poblet, and Xavier López*

Departament de Química Física i Inorgànica, Universitat Rovira i Virgili (URV), Marcel·lí Domingo 1, 43007 Tarragona (Spain)

E-mail: javier.lopez@urv.cat

Abstract

The properties and dynamic behavior of molecules in liquid solutions depend critically on the solvent and other species, or co-solutes, including electrolytes (if present), especially when molecular association or pairing occurs. In Quantum Mechanical (QM) calculations, the electronic structure of molecules in liquid solution is typically obtained with implicit solvent models (ISMs). However, ISMs cannot differentiate between, for example, cation types (e.g., Cs^+ versus $n Bu_4N^+$), leading to limited accuracy in capturing possible solute-specific interactions. Addressing this issue in QM calculations often requires an explicit treatment of the co-solute, typically a counterion, a challenging approach due to the definition of representative co-solute positions, numerical convergence, and high computational cost for bulky species.

A new computational strategy called Dynamic Environment in Solution by Clustering (DESC) is herein presented, which leverages classical Molecular Dynamics (MD) data to feed QM calculations, enabling the inclusion of counterion-specific effects with greater detail and efficiency than ISMs. DESC is particularly advantageous in cases where ion pairing/aggregation is significant, offering chemically representative QM results at a small fraction of the computational cost associated to the explicit inclusion of counterions in the model. This work presents MD data on polyoxometalate-counterion-solvent systems, introduces the philosophy behind DESC and its operational details, and applies it to polyoxometalate solutions and other relevant systems, comparing outcomes with benchmark QM/ISM calculations.

1 Introduction

Chemical processes mostly occur in liquid solutions, including all bio-activity in aqueous media. Consequently, the detailed chemical comprehension of solution effects, along with the appropriate application of solvent models in computational chemistry (both in classical molecular mechanics and in quantum mechanical calculations) has been subject of extensive research since its theoretical inception in the 1920s[1] and continues to be an area of intensive study today.[2–6]

The increase of computational power, combined with more efficient algorithms, allow for the inclusion of solvent effects and other factors related to the surroundings of target molecules into calculations and simulations, primarily through the use of solvent models. This progress has dramatically improved theoretical results, such as the accurate calculation of thermodynamic properties of processes in solution.[7,8] The well-known implicit solvation models (ISMs), such as the Conductor-like Screening Model (COSMO)[9] and the Polarizable Continuum Model (PCM),[10,11] are applied routinely to introduce the stabilizing effect of a liquid solution acting on some solute molecule (the *target*). The underlying idea of ISMs is to embed a molecule (the solute) within a cavity[12] that includes the effects of the surrounding liquid solution, the theoretical treatment of which mimics that of a homogeneous polarizable continuum material. The charge distribution of the solute and the solution mutually interact via the cavity surface, inducing a mutual charge polarization. In the end, changes are induced in the electron density of the solute upon incorporation of the solvent model.

Accounting for the solution effects implicitly provides correct results in general, this approach being capable of reproducing some general features of reactive processes, including energetic aspects at a qualitative or semi-quantitative level, and predicting solvation Gibbs energies. This is possible because, in many cases, the portion of the solution in contact with the target solute does not present specific localized inhomogeneities. That is, it behaves as a dielectric regular bulk that can be characterized by the few parameters of an ISM, being the solvent's relative permittivity, $\varepsilon_{\text{bulk}}$ and its molecular radius the most relevant. Nevertheless,

this behavior can change significantly for an ionic solute because its close environment can easily include, besides solvent molecules, interacting counterions in an ion pairing fashion with no mediating solvent molecules. By applying an ISM, the system's total charge is compensated by a set of *mirror* charges, not related to the actual nature of the counterion, and only influenced by the characteristics of the solute and by the relative permittivity of the solvent. These limitations do not permit capturing the actual effects of ion pairing onto the electronic structure of the target molecule. Consequently, regular ISMs cannot distinguish between cations (for example Li^+, K^+, Me_4N^+ or $n Bu_4N^+$), which can affect differently the solute properties if ion pairing takes place. Literature has shown that the cationic component of a salt in solution can have an impact on various properties, including shifts of several hundred mV in electrochemical waves.[13–15] Of course, counterions attached to the solute's surface can influence its reactivity and the stability of excited states,[13,16] among other consequences. Hence, a proper computational model accounting for these interactions is mandatory for a correct description of the system characteristics.

Polyoxometalates (POMs) constitute a large family of molecular polynuclear metal-oxides, typically anionic in nature, carrying charges ranging from very low, -1 or -2, to formally very high, -15 or even greater. For this reason, they often form ionic pairs in solution, such as the widely used quaternary ammonium cations, R_4N^+ in organic solvents. POMs have been largely studied in our and other groups without (in the earliest stages) and with ISMs since the early 2000s.[17] Indeed, it has been systematically shown that accounting for the effects of the target surroundings as part of the QM calculation is required for a reasonably good description of its molecular properties, among which the electron distribution via the molecular orbital shapes and energies,[18,19] although some general features can sometimes be equally nicely reproduced within the gas phase approximation.[20]

The effect of the environment on POMs has been thoroughly reported by our group.[17,21–24] In a recent study, we presented a combined experimental and computational investigation of a Wells-Dawson species functionalized with light-harvesting organic antennas.[13] In liquid

solution, certain electronic properties, such as excited-state lifetimes and redox potentials of the target system, exhibit a clear dependency on the solvent (CH_2Cl_2, MeCN, and DMF) and/or the counterions (Me_4N^+, Et_4N^+, and nBu$_4N^+$). These variations determine, at the timescale of excited-state relaxations, whether a given system is a promising candidate for generating a technologically relevant charge-separated state. That QM analysis explores the effect of cation substitution on the electronic properties of the POM in solution Figure 1(left)). Given that an ISM cannot internally distinguish between cations, we defined models with explicit counterions (taken from MD simulations) in proximity of the POM for our calculations. Depending on the solvent/counterion combination, they included 4 to 6 medium- to large-sized organic cations, entailing long computational times, considerable human effort and, in some cases, poor numerical convergence.

Another case studied by our group involves a triiron Keggin-type POM ($SiFe_3W_9O_{37}$) capable of binding and activating nitrogen with tetrahydrofuran as solvent, but only in the presence of lithium ions. The process consists of two consecutive reductions: the first one reduces the Fe^{III} centers to Fe^{II}, while the second one is attributed to the catalytic activity, i.e. the formation of NH_3 from N_2 in the presence of a proton source. Initial MD simulations on the solution evidenced notable Li^+-POM ion pairing. The following DFT calculations demonstrated that countercations must be explicitly included in the model to accurately reproduce the experiments. Thus, when ion pairing occurs it can play a crucial role, exerting a significant effect that cannot be described solely with an ISM model (Figure 1, right).

The incorporation of explicit ions in DFT calculations requires a detailed atomistic description of the solute environment.[26] The most reliable *in silico* method to obtain such ion arrangements is through molecular dynamics (MD) simulations, which provide the trajectories of all species in solution. The next step involves selecting several snapshots (nuclear positions at different simulation times), and selecting the coordinates for the desired POM and counterions units. Finally, these are used into the corresponding DFT calculations. To

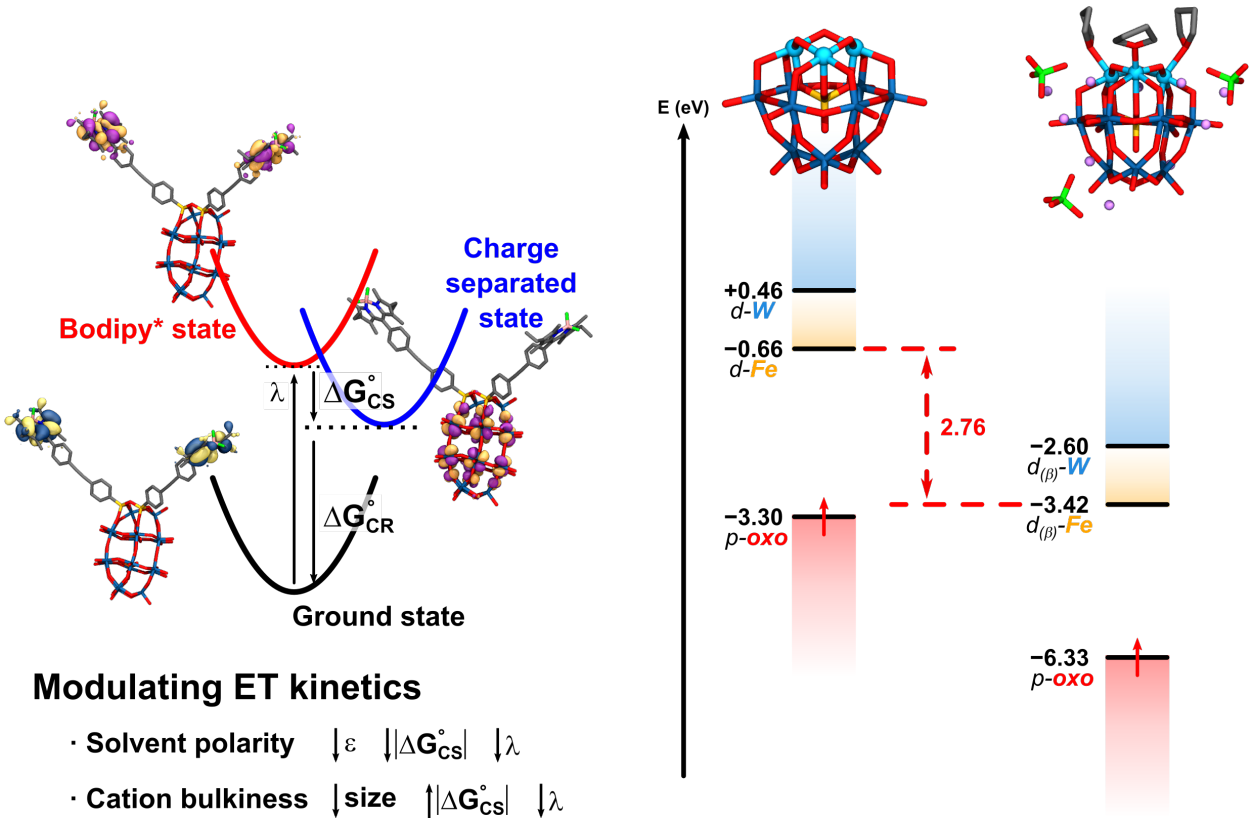


Figure 1: (Left) Influence of the solvent and cations on the electronic properties of a hybrid polyoxometalate.[13] (Right) Influence of the explicit environment on the frontier molecular orbital energies of a mixed-metal polyoxometalate.[25] In both cases, the explicit inclusion of counterions in the QM calculations was essential to rationalize the experimental observations.

ensure statistically representative results, multiple DFT runs on the selected snapshots are necessary. This methodology to include the counterion effects is, therefore, time consuming since it requires significant manual intervention and extensive computational time. It is also important to note that processing MD trajectory data and preparing QM input files are tedious tasks that, without some level of automation, can become particularly labor-intensive and cause of error.

Besides POMs, many other systems present aggregation as herein described. Among the compounds that have indisputable technological interest are fullerenes which, at different reduction states, gain tendency to form ionic pairs in solution. Also, transition metal coordination compounds, radical organic systems, etc. are susceptible of presenting some degree of association with co-solutes.

It is worth mentioning that efforts to tackle this topic differently can be found in the literature. For example, in 2013, Matsui et al.[27] proposed a protocol to correct the ISM-based redox potential values by adding a charge-dependent correction term for the counterions around charged target molecules. As a result, they considerably reduce the mean absolute errors but not capturing the structural features of solute-solute aggregation. More recently, an innovative study by Liu and coworkers[28] has employed Machine Learning techniques to correct errors in redox potential calculations arising from implicit solvent models (primarily C-PCM) and explicit solvent models (using QM/MM calculations to treat the solute at the QM level and the solvent at the MM level). The authors use simple solute and solvent parameters to train an algorithm capable of mitigating errors associated with conventional methods, specifically systematic bias and large-error outliers, which are primarily attributed to the large uncertainties in the solvation free energy of charged species. Furthermore, the approach proposed by Liu et al. reduces the dependence of the results on the choice of the DFT functional, a well-documented issue extensively discussed by various authors.[29,30]

To address these challenges, we herein propose a novel strategy that combines MD and QM calculations in a highly automated and efficient manner. This approach is specifically designed to account for co-solute (counterion) effects in solution at a significantly reduced computational cost, functioning as an extension to existing ISMs. Our method becomes particularly relevant when a co-solute and the target solute exhibit some degree of aggregation, as these interactions can lead to significant differences in the local environment of the solute—and thus its electronic structure—depending on the counterion. This scenario is common in various chemical processes, including concentrated electrolyte solutions, reactivity, and ion pairing in poorly soluble species, among others. If no molecular pairing takes place, the proposed approach naturally reproduces the ISM results.

2 Models and Methods

2.1 Computational Details

Classical Molecular Dynamics. Atomistic molecular dynamics (MD) simulations with explicit solvent molecules were performed with the GROMACS 2019.3 code[31,32] to determine the behaviour in solution of a set of POM anions together with the quaternary ammonium cations namely Me_4N^+, Et_4N^+, and nBu_4N^+. Simulations make use of a modified AMBER 14 force field, which has been satisfactorily employed to study the aggregation of POMs in different environments.[33] The force field provides the potential energy of the system as the sum of bond, angle and dihedral deformation energies and non-bonding terms. The latter consist of pair-wise additive 1-6-12 electrostatic and van der Waals potentials that account for interactions between atoms that are separated by more than three bonds.

Force field parameters for POMs were obtained following the procedure by López et al.[34] using the TopoMOx code.[35] Atomic charges to reproduce the molecular electrostatic potential (CHelpG) were obtained at the DFT level with the Gaussian 16 package[36] using the B3LYP functional.[37] Basis sets of double-zeta quality were used: LANL2DZ(f) for W and a Pople-type 6–31+G(d,p) for the rest. For the MD trajectories, cubic boxes were used with 3D-periodic boundary conditions, containing one POM anion, the number of cations required to neutralize the system and the embedding solvent for ca. 30 mM POM concentration. Solvent molecules were described by the full-atom model provided by van der Spoel et al.[38]

For 1-4 van der Waals interactions we applied an interatomic distance cutoff of 14 Å, and for Coulombic interactions of 14 Å corrected for long-range electrostatics by using the particle-particle mesh Ewald (PME) summation method. All bonds were restrained by the LINCS algorithm. Production trajectories were performed on a NVT canonical ensemble during 20 ns, collecting data from the trajectories every 1 ps. Simulations were carried out at 298 K, controlling the temperature by coupling the system to a thermal bath using the

velocity-rescaling algorithm. Before production runs, all systems were equilibrated by an initial 1 ns run at constant NPT to readjust the box size and a final 1 ns run at constant NVT with a relaxed solute.

Single point calculations. DFT calculations were carried out with the ADF 2022 program package[39] using the GGA BP86 exchange-correlation functional[40,41] and a Slater basis set of TZP quality. The solvent effects were introduced via the conductor-like screening model (COSMO) with different solvents.[9,42] The relativistic effects via the zeroth-order regular approximation (ZORA) were included. Initial structures were fully optimized, while the rest of the calculations (explicit and DESC model) were single point calculations.

In this work we put the focus on molecular orbital energies, which are typically linked to many macroscopic properties such as redox potentials or chemical reactivity, among many others. Thus, in some instances, having a more detailed and realistic picture of the nearby region of the solute is necessary for a better accuracy in the computational description of the electronic properties.

2.2 Target systems

The selected target species are the Lindqvist-type isopolyanion, $[W_6O_{19}]^{2-}$ (hereafter, L), and the Keggin-type aluminate heteropolyoxoanion, $[\alpha{-}AlW_{12}O_{40}]^{5-}$ (hereafter, AlK), in two reduction states: the fully oxidized and the two-electron reduced forms, with correspondingly different molecular charges. Quaternary ammonium cations of different size, that is Me_4N^+ (TMA), Et_4N^+ (TEA), and nBu$_4$N$^+$ (TBA), were used as counterions in solution. These were simulated in two solvents with differing polarity: CH_3CN (MeCN) and CH_2Cl_2 (DCM), with relative permittivities $\varepsilon = 37.5$ and 8.9, respectively. These combinations not only exhibit strong POM···cation ion-pairing but are also common in studies of POMs for various applications, such as catalysis and energy storage,[13,43,44] where the solute environment is crucial. The cases analyzed represent a broad range of the chemical spectrum explored here, from small, rigid cations to large, flexible ones, displaying diverse coordination patterns

depending on the POM's charge density, size, solvent polarity, and its capacity to solvate the anionic cluster. A graphical representation of the studied systems can be seen in Figure 2.

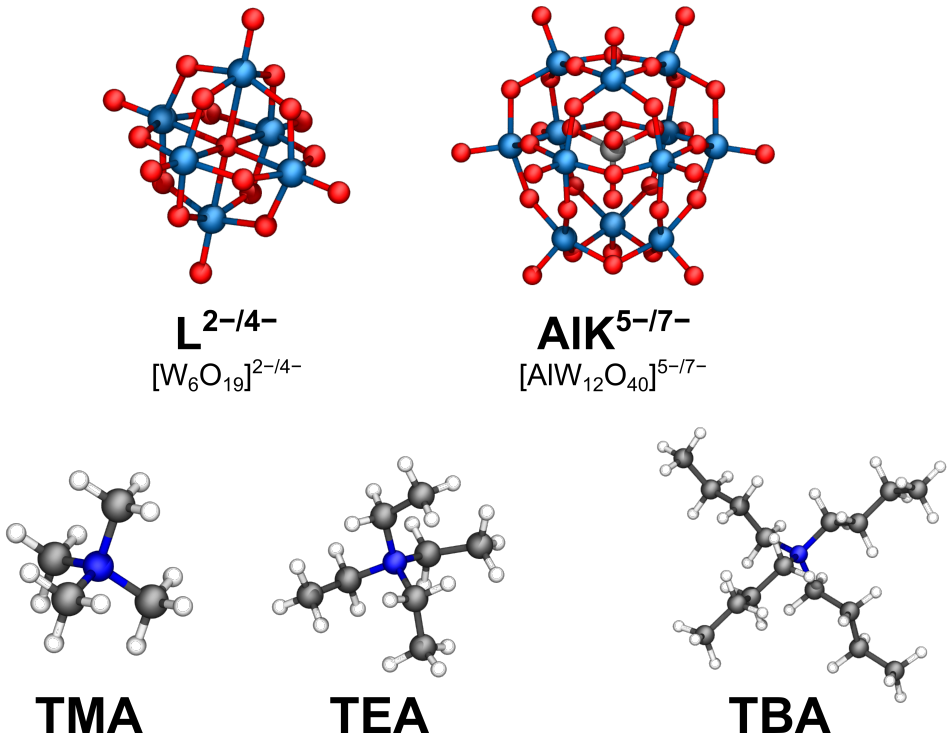


Figure 2: Ball-and-stick representations for the main POM (anionic, top) and counterion (cationic, bottom) species of the study. Anion-cation combinations with MeCN and DCM solvents were simulated. Color palette: W - blue, O - red, Al - silver, N - dark blue, C - gray, H - white.

3 The DESC methodology

3.1 Dynamic environment of POMs in solution

To obtain a detailed description of the solute environments, we conducted fully atomistic classical MD simulations in different media defined by the solvent and the counterions (see Section 2). The analysis presented here emphasizes the spatial distribution of the solvent and the nearby counterions in solutions containing a POM molecule in various reduction states. This analysis aids in understanding how the solute's surroundings can vary depending on

the counterions present and highlights the importance of accounting for this phenomenon to achieve reliable QM-based results.

Table 1 presents the most relevant parameters of the model, derived from the resulting MD trajectory: the average aggregation number, the corresponding maxima of the highest peak of the radial distribution function (RDF) (1^{st}), and the integration distance (I).

This exhaustive table lists the relevant data obtained from the aforementioned MD simulations. Key observations on POM···cation ion-pairing across the analyzed systems can be summarized as follows:

- Smaller cations consistently exhibit closer and stronger aggregation due to their size and hardness: **TMA(N, 1^{st}, I) > TBA(N, 1^{st}, I)**.
- In more polar media, such as MeCN, POMs and cations are better solvated, resulting in fewer ion pairs: **DCM(N, 1^{st}, I) > MeCN(N, 1^{st}, I)**.
- POM···cation association reaches saturation more rapidly for larger cations, which may eventually interfere with one another sterically or electrostatically.

These observations are well established within the scientific community and are consistent with previous studies[13,33,45,46] and with the findings presented in this work. Nevertheless, their significance remains undiminished, as they provide an exceptionally detailed description of counterion positioning around the *target* solute. Additionally, they enable precise tracking, at the angstrom scale, of the evolution of both ionic pairs and solvent molecules throughout the solution. Two examples of the simulation results are shown in Figure S1: the $(TMA)_5AlK$ salt in MeCN and the $(TBA)_4L$ salt in DCM, which illustrate the observations discussed above.

The perspective on the dynamic environment of a solute (in this case, a POM anion), which varies significantly based on numerous parameters, provides a foundation for developing an alternative computational methodology. By directly extracting data from MD trajectories in an ordered and automated way, we construct a model that incorporates this

Table 1: Relevant data taken from 24 MD simulations of the model systems: combinations of two POMs, two reduction states, three countercations, and two solvents. The distance of the first peak maximum (1^{st}) and the integration distance (I), both in Å, and the averaged coordination number (N) are given. Data taken from the last 10 ns of the trajectory, sampled every 1 ps.

POM	**q**	**Solvent**	**Cation**	**N**	**1^{st}**	**I**
L	−2	MeCN	TMA	1.04	6.370	9.480
			TEA	1.93	6.659	10.365
			TBA	1.13	6.664	9.370
	−4		TMA	4.00	6.071	8.480
			TEA	3.99	6.071	7.777
			TBA	1.71	6.483	9.297
	−2	DCM	TMA	1.99	6.269	9.479
			TEA	1.93	6.577	9.690
			TBA	1.72	6.761	10.267
	−4		TMA	4.00	6.076	7.884
			TEA	4.00	6.056	8.158
			TBA	2.99	6.264	9.171
AlK	−5	MeCN	TMA	4.33	6.976	8.582
			TEA	4.99	7.070	9.978
			TBA	3.71	8.057	10.659
	−7		TMA	7.00	6.960	8.362
			TEA	7.00	6.975	9.383
			TBA	3.36	7.955	10.857
	−5	DCM	TMA	5.00	6.965	9.971
			TEA	5.00	7.078	8.684
			TBA	4.19	7.860	10.964
	−7		TMA	7.00	6.955	8.356
			TEA	7.00	6.965	9.069
			TBA	5.92	7.678	10.287

information in a fast, straightforward, and user-friendly manner, integrating it into routine QM calculations to account for the effect of the environment on the target solute properties.

3.2 Influence of the explicit environment on the electronic structure of the solute

With a detailed description of the environment, the effect of counterions on the POM was analyzed with the traditional methodology: selecting N random MD snapshots (in this case, five) representing aggregation, obtaining the respective atomic coordinates, and performing QM (specifically, DFT) calculations. Then, the electronic properties were determined by averaging the results of the five DFT runs. A key point to note is that, in cases where the average aggregation values for the selected snapshots are not integers, and straightforward rounding is not feasible (i.e., values between X.22 and X.78, where $X \in \mathbb{N}$), the average was weighted according to the relative frequency of the observed aggregation values:

$$\begin{gathered} E_i = w_a \cdot E_i\left(N_a\right) + w_b \cdot E_i\left(N_b\right) \\ w_a + w_b = 1 \end{gathered} \tag{1}$$

In Equation 1, E_i denotes a property (e.g., total energy or MO energy), and a and b are the subscripts of the two nearest integer values to the actual fractional aggregation number.

It should be noted that when we refer to environment effects, we are specifically addressing the influence of explicit counterions. Solvent effects are incorporated via the ISM (COSMO in this study). This strategy has been successfully employed in previous studies.[13,47,48] Herein, special attention is given to molecular orbital energies, as these are often correlated with macroscopic properties such as redox potentials, spectroscopic properties, and chemical reactivity, among others.

Figures 3 and Figures S2 and S3 clearly highlight the significant role of explicit cations in the systems analyzed. It is evident that the molecular orbital energies can vary considerably depending on the cation being the co-solute. Consequently, taking the frontier

orbitals as representative of the effects taking place on the electronic structure, the range of absolute energies computed for the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) is substantial. For instance, in DCM, the HOMO of the oxidized AlK system shifts from −6.14 to −6.71 eV when TBA is replaced by TMA, showing a difference of over half an eV. In contrast, the corresponding HOMO energy at the COSMO level is −5.51 eV. In MeCN (Figure S3), the range of energies is narrower but remains significant; for the oxidized AlK system with TMA cations, the LUMO is found at −4.14 eV, while the pure COSMO calculation gives −3.86 eV. Lastly, in the reduced AlK system in DCM (Figure S2) with TMA, the HOMO and LUMO lie at −3.19 and −2.94 eV, respectively, while the COSMO values are −1.68 and −1.45 eV. For systems containing L (reduced system in Figures S2 and S3), the observations are similar. The discrepancies in orbital energies are, in some cases, greater than 1.5 eV, rendering pure ISM results unreliable in systems where aggregation is present.

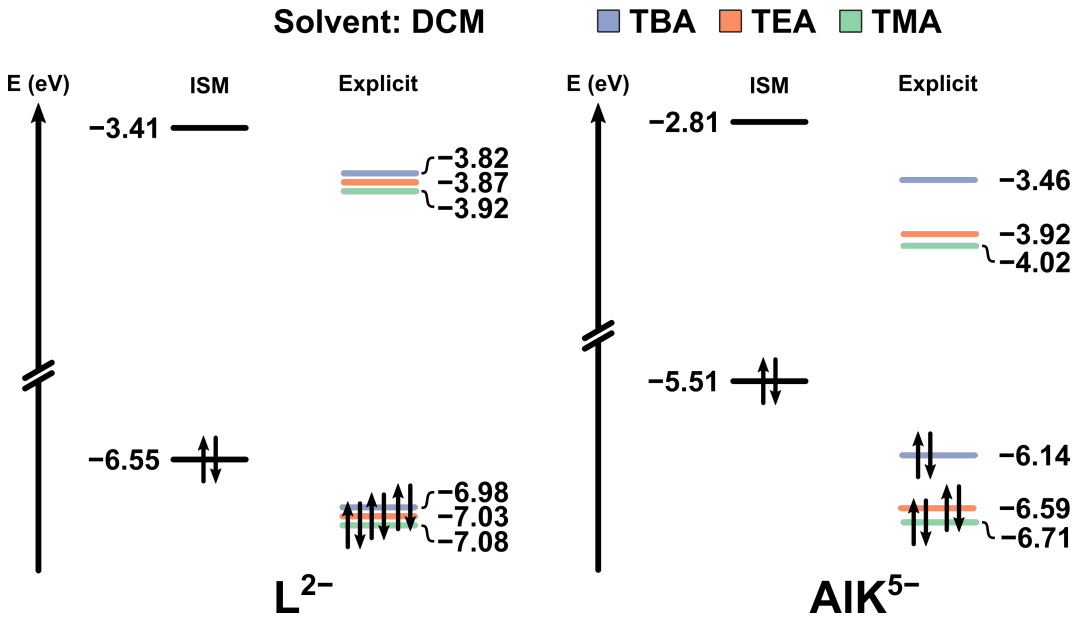


Figure 3: Molecular orbital diagram for L and AlK salts of TMA, TEA and TBA in fully oxidized state, in DCM calculated following two methodologies: ISM (COSMO, $\varepsilon = 8.9$) and Explicit+ISM (average of 5 snapshots).

It is clear that when a pure ISM is used to simulate solution effects, the calculations cannot distinguish between different cations. Moreover, the pure COSMO results hardly resemble

any of the explicit models used as benchmarks. However, the HOMO-LUMO gaps across all explicitly treated cases are quite similar (differing by approximately 20–30 meV), suggesting that certain properties dependent on these gaps may still be reasonably reproduced by the ordinary solute + COSMO approach. It is worth mentioning that the electronic energies obtained with COSMO are not always the least stable, particularly in solvents with high relative permittivity (such as MeCN). This aligns with expectations for ions in solution, where electrostatic interactions are influenced by $1/\varepsilon$. Conversely, in DCM, the electronic structure description provided by COSMO alone is poor and deviates considerably from that of the explicit aggregated solute model.

The next step is: how can we combine the environmental effects introduced by COSMO with those of the aggregated co-solute? At this point, it is important to stress that a given new strategy must be computationally cheaper and more efficient than the method of explicit cations mentioned above.

3.3 A new strategy to account for the environment in solution: DESC

When analyzing the differences between an ISM calculation and a calculation with explicit counterions, two primary factors can be distinguished:

- Electrostatic effects: The co-solute surrounding the target molecule may be a charged species that influences its electronic structure. Additionally, since the solvent cavity is generated by point charges that model ISM polarization, the presence of an adjacent species alters the ISM solvent cavity itself, leading to modified charge distributions in certain regions of the cavity and, consequently, affecting solvation.

- Steric effects: Co-solutes aggregated around the solute occupy physical space, reducing ISM interactions with that part of the target molecule.

With the results from the previous section and these foundational considerations in mind,

we developed the new methodology coined Dynamic Environment in Solution by Clustering (DESC), which is described in detail below. The computational strategy employed by DESC has been implemented in a Python code[49] that, given a MD trajectory in PDB (Protein Data Bank) format and a simple input file with basic system information, analyzes the MD simulation and generates an ADF input file ready for submission, thus incorporating the effects of aggregated ions on the solute. This model is broadly applicable, with no restrictions regarding molecular charge, nature or number of constituent molecules, or solvent type. The use of ADF software is not limiting; the concepts behind DESC are adaptable to other computational software. However, at this stage, testing has been conducted exclusively with the ADF program. The DESC code can be easily used by installing it from the GitHub platform (`https://github.com/qcgurv/DESC`).

One of the most unique aspects of DESC is how it accounts for the counterion effects by introducing them through: 1) a cloud of point charges positioned at the locations of counterions aggregated during the MD simulation, and 2) a set of ghost atoms with explicit spherical volumes simulating the presence of cations aggregated to the target solute, placed at the most representative positions. With appropriately selected radius and charge parameters (described below), this methodology replicates the effects of a calculation with explicit counterions without double-counting their impact on the ISM. While COSMO already implicitly introduces the average effect of counterions contained in the bulk, the introduction of ghost atoms with volume, which restrict the ISM from reaching certain solute regions while providing an adjusted surrounding charge distribution, reproduces the localized effects of explicit counterions.

Previous to applying DESC, it is essential that the target solute retains consistent orientation (both rotational and translational) with respect to the other molecules surrounding it throughout the MD trajectory. This consistency ensures that all frames processed by DESC are equivalent and comparable, thereby accurately reproducing the co-solute environment. Trajectories can be easily converted using specific software like GROMACS, Python libraries

such as MDAnalysis,[50,51] or VMD scripts.[52] These algorithms employ a frame from the MD simulation as a reference, minimizing the root-mean-square deviation (RMSD) between frames by fixing one residue (herein the target molecule) and adjusting the rest accordingly. Figure S4 shows two example snapshots from a simulation of $(TBA)_7AlK$ in MeCN, before and after reorientation. Effectively, this process aligns the classical trajectory by fixing the position of one residue (the target solute, i.e., the POM) while repositioning everything else, including the simulation box. Original energies and interactions remain unaltered during this process.

The complete DESC workflow is illustrated in Figure 4. The following section will detail each step to explain how we transform a MD trajectory into a ready-to-run QM input file.

3.3.1 Input file

Once the trajectory has been reoriented based on the position of the solute, the next requirement for DESC to function is basic information about the system. To maintain the simplicity of the method, the input file requires only:

- The name of the trajectory file (in PDB format)
- The number of frames to be analyzed
- The name of the solute residue
- The name of the aggregated co-solute residue, its most representative atom, and its formal charge

An example of the input file for a TMA-AlK system in MeCN is provided in the ESI.

3.3.2 Splitting frames

Even though all frames in a dynamic simulation are equally important, not all are necessary in this approach; only a selection is taken and processed. In the 'standard' method for explicit QM calculations, the first step implies selecting a collection of well-spaced snapshots and,

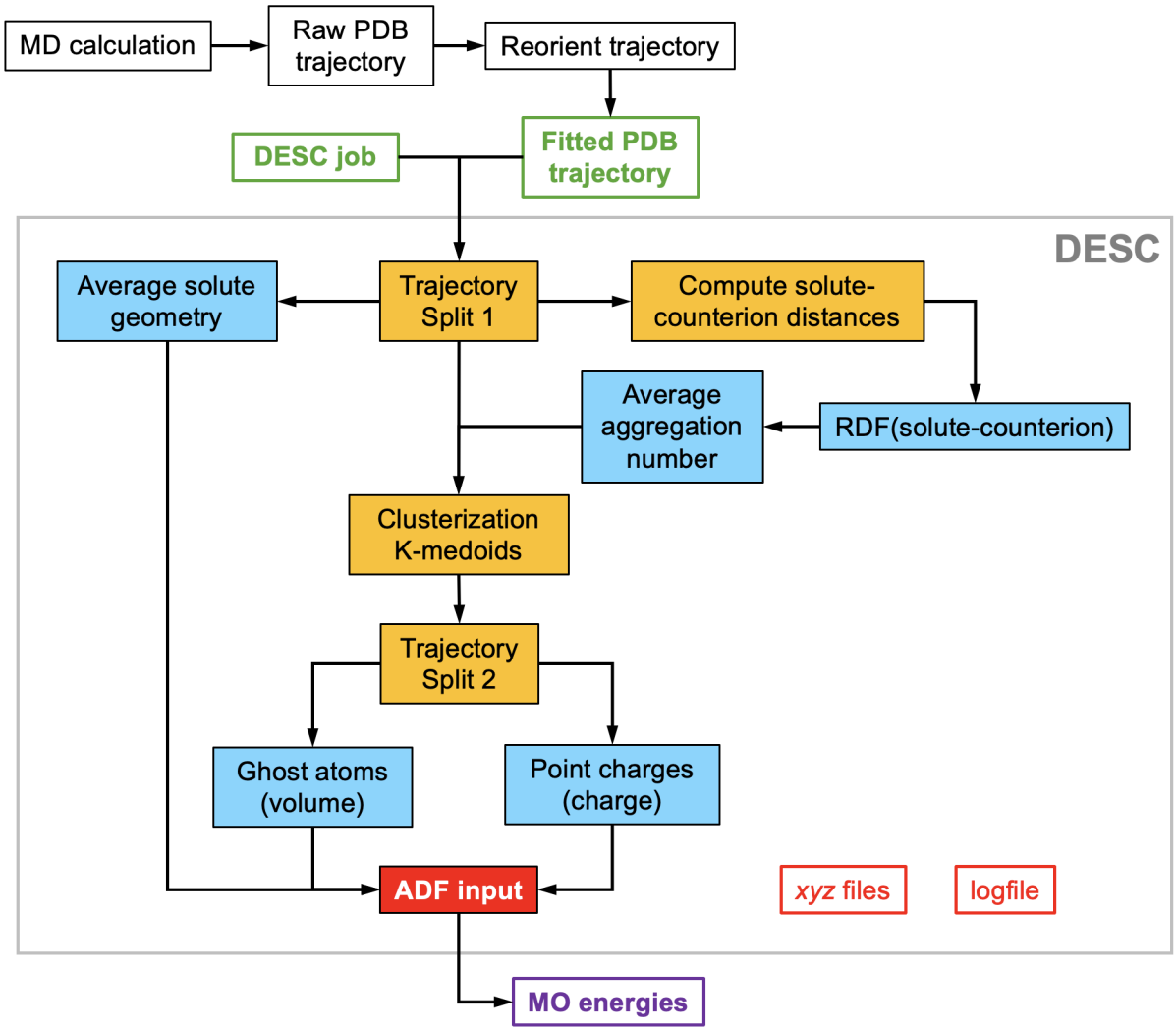


Figure 4: Schematic view of the main data treatment associated with DESC. Taking the raw PDB data from a MD run, an initial transformation must be done to generate the fitted PDB which, together with a job-DESC input file, feed the program (green boxes). A simplification of the DESC workflow is represented inside the thin box, with tasks and outputs represented in orange and blue.

then, QM calculations are run. Post-processing of these results generate an average of the electronic property under study, assumed representative. Therefore, in a small to medium trajectory containing 10,000–50,000 frames (10-50 ns), at least 5-10 snapshots (and their QM runs) should be involved, despite the sizeable computational and human cost associated. Instead, DESC extracts firstly a large number N of frames from the MD trajectory to work with. A relatively large N is crucial to ensure good representativeness of the DESC result. We have determined that $N = 400$ for this initial split is largely sufficient. Using a lot less frames may yield a poor description of the solute environment, while a lot more would slow down the analysis without improvement of the results. From this initial split the radial distribution function (RDF, $g(r)$) between the centers of mass (COMs) of the solute and the counterions is calculated and, then, the associated average number of ions aggregated around the solute and their distances. Therefore, the number of frames is critical for achieving a well-resolved RDF (see Figure S5 for further information).

From the set of 400 frames selected, a subset of 100 frames is taken (either one every four consecutive frames or randomly, the result has proven to be consistent), from which the charge distribution surrounding the target molecule is captured and used as input data for the last QM calculation step. This selection influences the ease of convergence in the self-consistent field (SCF) procedure during the QM calculation, its speed, and how accurately the ionic environment of the solute is described, depending on the number of point charges included. The rationale for selecting 100 frames is illustrated in Figure S6, which shows how the energy of the frontier molecular orbitals of an L^{2-} evolves based on the number of frames in this second split. Importantly, if only point charges are included, the orbital energies remain different from the ones of the snapshot.

3.3.3 Average Solute Geometry

Another feature of DESC is that it provides the user with a nearly optimized structure of the target solute for carrying out QM calculations. From the first split, the 400 solute structures

are collected and an average structure is generated. This calculation benefits from the initial reorientation, where all solute structures analyzed by DESC share the same orientation and minimal internal RMSD. The resulting structure is considered nearly optimized because it is derived from a dynamic simulation where, in principle, the solute geometry is optimized, and the force field applied does not significantly distort the molecule.

An example is provided for AlK^{5-} in MeCN, where the electronic energy difference (ΔE) between the DFT-optimized structure and the DESC-generated structure is 1.96 kcal mol^{-1}, with the frontier molecular orbitals showing a mean absolute error (MAE) of 20 meV (-6.56 eV vs. -6.54 eV for the HOMO, and -3.86 eV vs. -3.83 eV for the LUMO in the optimized vs. DESC structures, respectively). Regarding molecular structure, the RMSD between the DFT-optimized and DESC-generated structures is 0.0034 Å, supporting the consistency of the electronic parameters. Reflecting this low RMSD, Table S1 shows the average bond distances and their deviations for the various bond types in AlK in both the optimized and DESC-generated structures. This feature can be highly advantageous for DESC applications, simplifying the study of environmental effects on solute reactivity or spectroscopy by assuming the solute structure is effectively "optimized."

3.3.4 Automated analysis of the RDF

From the previously described solute-counterions RDF, whose generation and analysis are incorporated into DESC without further user input, the peak integration distance and the average number of counterions around the solute are determined. Prior to this, the $g(r)$ function is smoothed using a Savitzky-Golay filter, which preserves peaks and curvatures by fitting a low-degree polynomial to a series of data points and then calculating the smoothed value of the signal. This filter is implemented in the code via the SciPy library,[53] which also enables calculation of the function's first derivative. With these tools and the SciPy's *find_peaks* function, peak maxima (*find_peaks*) and minima (when the first derivative is zero and changes from negative to positive) are identified to establish the integration distance for

each peak. If multiple peaks are present, the integration distance of the peak farthest from the solute is used. The robustness of this method was tested and confirmed during code development.

When solute and counterions do not present aggregation, the RDFs may show distant, weak peaks or integration distances much further than typically considered to be the coordination sphere of the solute. Since DESC can be applied to any simulation, regardless of the presence of ion pairs, an aggregation cutoff distance to prevent these spurious peaks to be considered later. Thus, if a peak in the RDF is found not accomplishing this criterion, it is discarded and no aggregation is considered. The cutoff distance is system-dependent and is calculated as:

$$D_{cutoff} = \left(\frac{d_{max,solute}}{2} + 1.5\right) + \left(1.5 \times R^{v}_{g,counterion}\right) \tag{2}$$

where $d_{max,solute}/2$ is half the largest interatomic distance within the solute, and R^{v}_{g} is the radius of gyration of the counterion, explained in detail in a later section. The interatomic distance is increased by 1.5 Å to approximate a sufficiently generic "effective van der Waals radius". Additionally, R^{v}_{g} is scaled by a factor of 1.5 to ensure that D_{cutoff} encompasses an adequate pairing distance. Although this distance may seem excessive or lack physical meaning, its purpose is to prevent unrealistic interpretations of the RDF that could lead to inexistent aggregation.

An example of RDF calculated by DESC is shown in Figure 5, which highlights the original RDF, the smoothed RDF, the peak maximum, the first derivative function, and the integration distance.

3.3.5 Clustering of counterions' positions

At this point, DESC has collected the solute structural data and the associated distribution of aggregated ions, represented as points located at the centers of mass of the explicit co-solute molecules (or, if unavailable, at the most representative atoms, such as nitrogen atoms

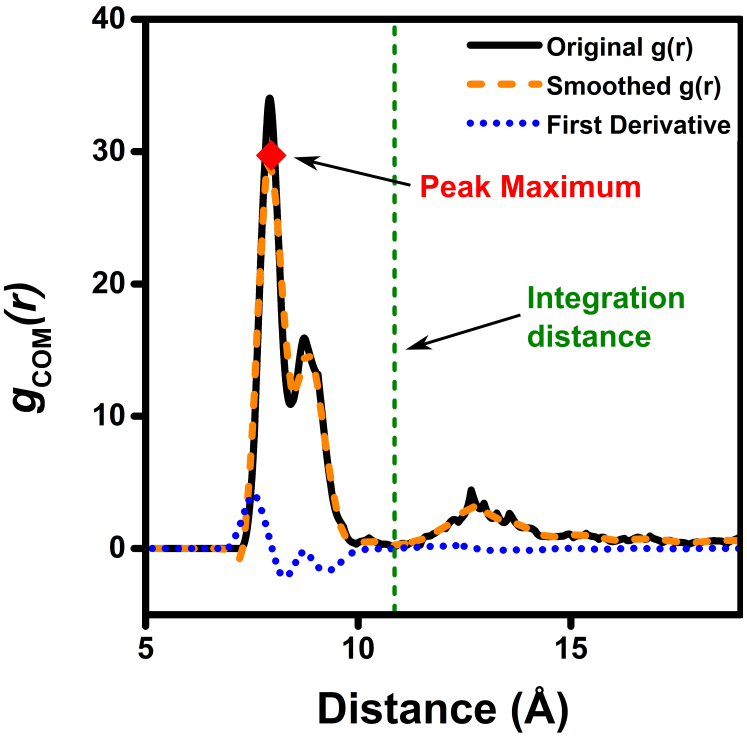


Figure 5: Radial distribution function and associated features, calculated by the DESC code.

for $n\mathrm{R_4N^+}$). The next step is to group these points by spatial proximity argument(s), that is, to make *clusters* of points. This step applies a clustering method, specifically the well-known K-Medoids algorithm.[54,55] The choice of clustering algorithm is based on the need to find as many clusters as the RDF analysis has previously identified as aggregated counterions, (N). Among clustering algorithms that depend on a set number of clusters (such as K-Means, agglomerative clustering, etc.), the selection of K-Medoids is not deliberate. Next, once clusters have been generated, a special point per cluster is identified and, hence, converted into a ghost atom (with no charge and with volume). While similar to K-Means, K-Medoids is applied by DESC. It minimizes the sum of the distances between each point and the medoid of its cluster, rather than the sum-of-squares within the cluster. Unlike a centroid, at the end of the procedure one medoid is identified: an existing point within the original dataset, not an averaged location that may fall outside the actual collected data. Figure 6 displays the clustering results for the system $(\mathrm{TBA})_5\mathrm{AlK}$ in $\mathrm{CH_2Cl_2}$ ($N = 4.19$), with four clusters differentiated by color. The medoid of each cluster is highlighted with a larger

sphere and will ultimately be treated as a ghost atom in the final QM calculation.

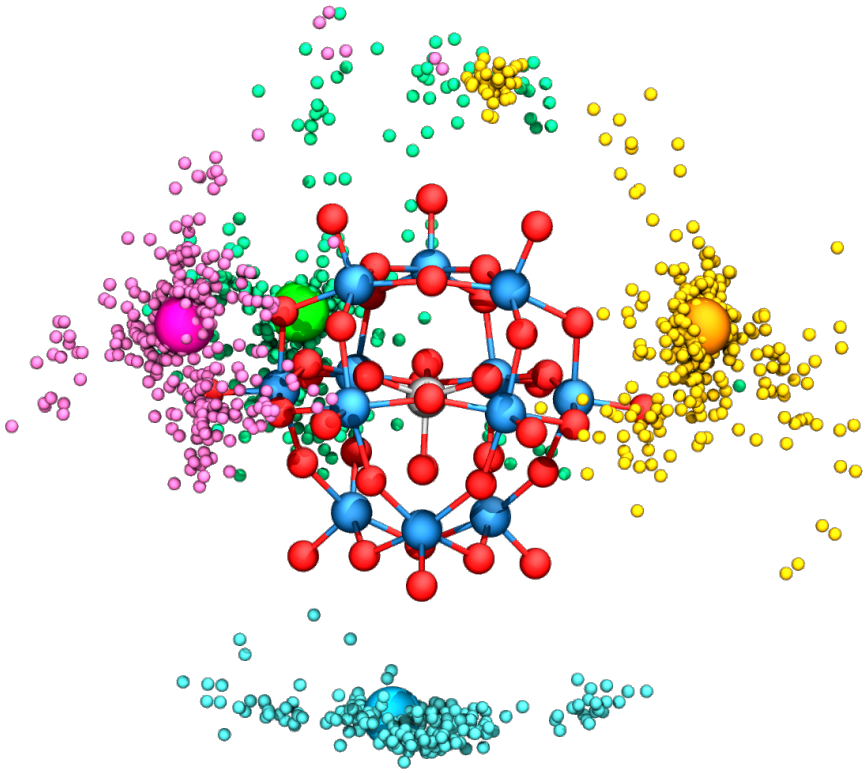


Figure 6: Representation of the four clusters of counterions generated by DESC (colored fuchsia, green, orange, and blue) for $(TBA)_5AlK$ in CH_2Cl_2. The four corresponding medoids are highlighted as larger spheres.

The average aggregation number obtained by DESC may be fractional. An interesting feature of DESC, similar to the explicit approach for including counterions, is that when the value obtained falls between X.22 and X.78 ($X \in \mathbb{N}$), the clustering algorithm is applied twice, once for each of the two nearest integers (X and X+1), generating two QM input files, one for each integer. Subsequently, the electronic properties obtained from the final QM calculations are weighted according to Equation 1. These weights are provided to the user in the DESC logfile.

With these components defined, we can precisely identify ghost atoms and point charges within the model based on the RDF analysis. The ghost atoms correspond to the medoids, representing the statistically N most representative positions of the ion groups (clusters) near the solute. The point charges are the remaining points lying within integration distance. We

now illustrate the treatment applied to each component.

3.4 Ghost atoms and Point charges

3.4.1 Ghost Atoms: Volume-weighted Radius of Gyration

The role of ghost atoms and how they are created is a major innovation in this model. Currently, computational chemistry programs use spherical atomic surfaces to create the cavity surrounding molecules to simulate the ISM. This approach naturally extends to using spherical volumes for ghost atoms in DESC, regardless of the morphology of the molecule they represent. The question then arises: what radius best characterizes the residue represented by the ghost atoms? Various metrics can describe a molecular shape, such as the distance from the COM to the farthest point, the sum of van der Waals radii, or the radius of the smallest sphere encompassing the entire molecule, among others. There is no definitive answer as to which is ideal; the choice largely depends on the context and the goal of the chosen approximation. To best represent the associated counterion treated as a ghost ion with volume, we have determined that the volume-weighted radius of gyration (denoted as R_g^v) is the most suitable descriptor for defining the volume of each ghost atom:

$$R_g^v = \sqrt{\frac{\sum_i v_i \left(r_i - r_{COM}\right)^2}{\sum_i v_i}} \tag{3}$$

In Equation 3, summations run over all the atoms in the molecule, v_i is the van der Waals volume of the i-th atom, r_i is the position vector of the i-th atom, and r_{COM} is the position vector of the molecule's COM. Here, the van der Waals radii are consistent with those used by COSMO in the ADF program, derived by dividing the MM3 method radii calculated by Allinger[56] by 1.2.

The radius of gyration provides an indication of molecular size, accounting for shape, flexibility, and conformation by representing the root-mean-square (RMS) distance of each atom from the COM. Traditionally, in polymer chemistry, each RMS distance is weighted

by atomic mass. For our purposes (defining the volume occupied by a molecule as a single point at its COM), weighting each RMS distance by van der Waals volume is more suitable. DESC calculates the R_g^v of each target molecule in every selected frame of each trajectory and internally computes the average, that is, it determines the R_g^v directly from the simulation, thus accounting for the flexibility and deformation of the large counterions aggregated with the solute, implicitly including the effect of the solvent. This significant feature distinguishes DESC from other models. The volume of the ghost atoms prevents the ISM from forming a cavity around those regions of the solute, thereby avoiding over- or under-stabilization effects from the solvent. Figure 7 shows how DESC assigns volumes to ghost atoms, modifying the cavity around the solute to mimic the explicit model. This approach also provides a straightforward way to distinguish between counterions. The ghost atoms, however, are neutral. Although this may seem counterintuitive, particularly when the coordinated molecules are ions, the primary role of the ghost atoms in this model is to occupy the physical space occupied by the aggregated co-solute, as the explicit model does. The electrostatic influence of the surrounding ions, as explained below, is entirely provided by the cloud of point charges derived from the selected MD frames.

DESC automatically generates an ADF-like ghost atoms input block, **Gh.H**, that includes the *xyz* coordinates previously identified as medoids. For the SCF calculation, the ghost atom is arbitrarily a pseudo-hydrogen atom with zero nuclear charge and minimal basis set, a choice not affecting the electronic structure of the solute as it contributes nothing beyond occupying space. An example of the ATOMS block for a $(TMA)_4L$ system, where $N = 4$, can be found in the ESI.

3.4.2 Point charges: Fitted point charges

If we assert that there are countless ways to estimate the ill-defined radius (or volume) of a molecule, attempting to mimic the electrostatic effects of an ionic co-solute with a volumeless point charge is equally challenging. During the development of this method we found that,

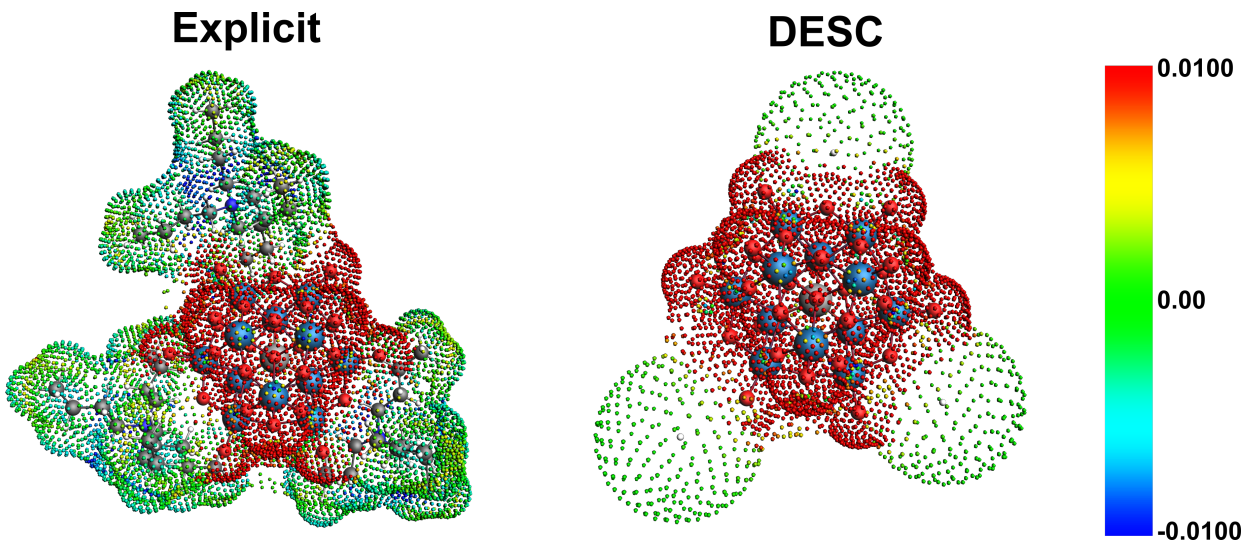


Figure 7: Comparison between COSMO charge density surfaces for the explicit system $(TBA)_7AlK$ (where N=3) and the corresponding DESC approach. Despite the differences in shape between the TBA cavities and the spherical ghost positions obtained via clustering, the amount of solvent excluded from the solute surface is comparable. A color bar indicating charge density is included.

to achieve a realistic effect on the solute, these point charges should not be equal to the formal charge of the ions that they represent (further details in the ESI). Instead, these charges must be significantly lower. This adjustment is necessary due to a combination of effects involving the surrounding solvent, which polarizes the charge distribution based on its relative permittivity, the internal structure of the co-solute and its electrostatic self-screening, and interactions with neighboring ions (mainly repulsion with co-ions and attraction to other species). Altogether, these factors scale the intensity of the electrostatic interaction between the solute and the point charges in a complex manner. Nevertheless, we have parameterized the following simple equation using the systems listed in Table 1, which effectively captures the screening effect on each point charge in a straightforward manner:

$$q_{fitt,i} = q_{formal} \frac{\langle N_{ion} \rangle}{\sum_i N_i} \left(0.025 + \frac{1}{\varepsilon_{bulk}} \right), \tag{4}$$

where q_{formal} is the formal charge of the ion (+1 for $n\mathrm{R}_4\mathrm{N}^+$), N_{ion} is the average coordination number of ions around the solute, $\sum_i N_i$ is the total number of point charges

determined by DESC, and the term in parentheses is an empirically fitted correction factor.

The above expression indicates that solvents with higher relative permittivity reduce the fitted effective charges $q_{fitt,i}$. This effect arises primarily because solvents with high relative permittivity can more extensively polarize the co-solute's charge distribution, thereby diminishing the intensity of the charge experienced by the solute. Interestingly, DESC automatically detects the solvent used in the MD run (from the input PDB file) and assigns the appropriate permittivity constant for internal data processing. The solvent is not specified by the user in the DESC input; rather, it is identified by locating the most abundant residue in the PDB file, processing one of these residues to construct its molecular formula, and comparing it to an extensive internal solvent library to obtain the permittivity constant. If the solvent is not found in the library, or if the formula processing fails, the user is prompted to enter the solvent's name and relative permittivity, which are then used to continue DESC's calculations.

For the set of point charges (one of the main DESC outputs) the code generates a *xyz* file and also incorporates an ADF-like input block containing the *xyz* coordinates of the non-medoids points, along with a column of charges corresponding to the q_{fitt} value in atomic charge units. The number of elements in this list varies depending on the degree of ion-solute aggregation. All information generated by DESC is included in a ready-to-run ADF input file. An example of the point charge block in ADF format is provided in the ESI.

3.4.3 Combining sterics and electrostatics

Before discussing the performance of the DESC model, it is essential to emphasize the importance of the two key components of this approach: the distribution of fitted point charges and the neutral ghost atoms.

Using the $(TMA)_7AlK$ system in MeCN as an example, where $N = 7$, Figure 8 illustrates the contribution of each component to the electronic structure of the salt, spanning from the highly destabilized gas-phase approach to the DESC model, in excellent agreement with the

benchmark explicit model results. This demonstrates not only the accuracy of DESC but also the individual impact of the two main components of this model. Applying ghost atoms only, the ISM surface has reduced contact with the solute, leading to destabilized orbital energies compared to the ISM-only model. Only when the cloud of fitted point charges is added to the final QM calculation do the energies align with those obtained using the explicit model, significantly improving the ISM-only results.

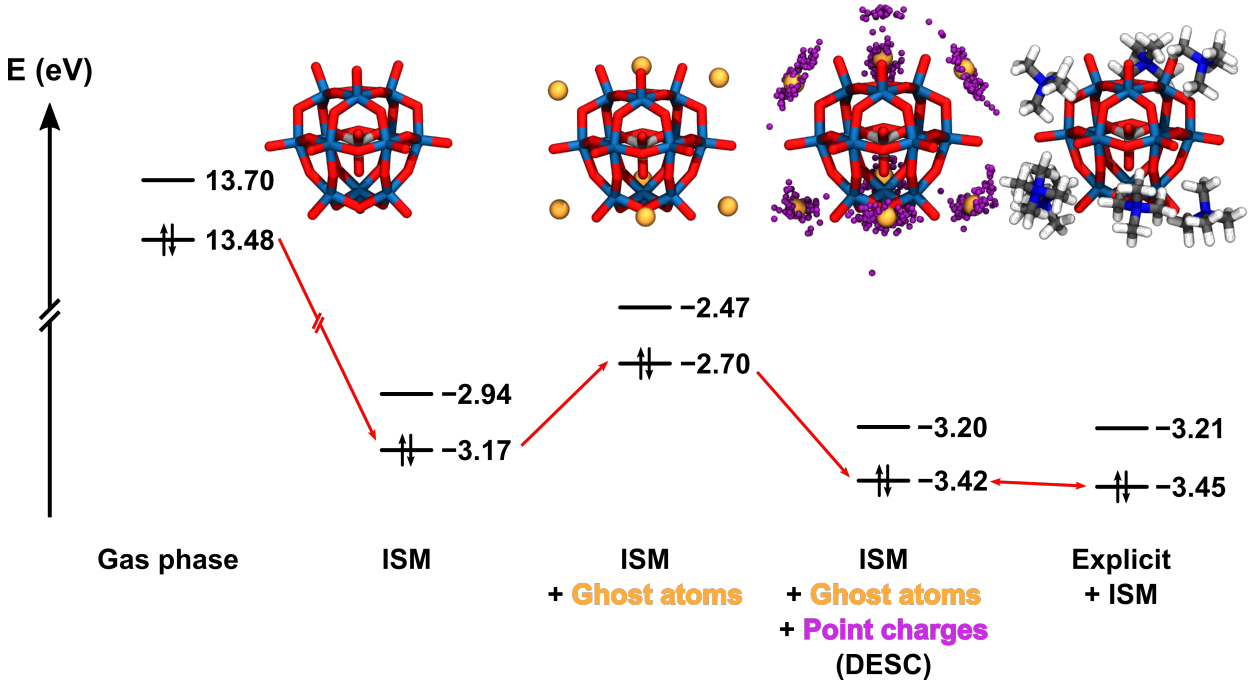


Figure 8: Frontier molecular orbital energies of $(TMA)_7AlK$ in MeCN calculated using different models. ISM corresponds to COSMO with $\varepsilon = 37.5$. A structural representation of the POM for each method is included. Color palette: W - blue, O - red, Al - silver, Ghost atoms - yellow, Point charges - purple, N - dark blue, C - gray, H - white.

The final output of DESC is a logfile summarizing the main results obtained during the process starting by reading a PDB file of the MD simulation. An example for $(TMA)_5AlK$ in MeCN, with $N = 4.33$, is also presented in the ESI. In this case, where the average aggregation number is non-integer and lies within the X.22-X.78 range, DESC provides the weights of the two nearest integers in the logfile for application in each QM calculation.

3.5 Comparison between DESC and explicit model

Once the components and operating mode of DESC have been explained, let us evaluate its performance. Our benchmark values refer to the average energies of selected relevant molecular orbitals of the solute: in the present case the highest orbital of the oxo band, the HOMO, and the LUMO. In cases where the POM is fully oxidized, the first two coincide. To assess our new methodology, the orbitals obtained with DESC are compared to those from explicit model calculations, analogous to Figures 3 and Figures S2-S3, where COSMO was compared with the explicit model. The electronic structure data obtained with the two models are presented in Figure 9 and Figures S7-S8 for MeCN and DCM, respectively. In general, the solute molecular orbital energies obtained with DESC show semi- or quantitative agreement with the explicit model, i.e. differences no larger than 100 meV with almost no exception. Also noteworthy is the capability of DESC to differentiate between counterions, accurately reflecting the variations in the electronic structure of the POM introduced by strongly aggregated explicit counterions, which are now reproduced with a simpler model. Additionally, the energies obtained with the ISM-only model notably deviate from DESC and the explicit model deviate, with no possibility of not cation selection. Finally, execution of DESC is very fast, requiring only seconds in interactive mode for the examples presented here. This includes reading the entire PDB file, processing data, and generating the ADF input and other DESC output.

As shown in Figure 10, the mean absolute errors (MAEs, in meV) for DESC-based orbital energies relative to the explicit model (DESC vs. explicit, orange bars) are very low, mostly below 100 meV and very frequently below 50 meV. In contrast, the quality of ISM-derived orbital energies (ISM vs. explicit, purple bars) is generally much worse. Indeed, a substantial number of COSMO tests deviate as much as 500 meV, and some MAEs even exceed 1000 meV when the solvent is DCM. These results emphasize the effectiveness of the strategy and design of DESC, and bring to light the limited accuracy of a pure ISM approach for analyzing aggregated, complex systems in solution. The worst performances of DESC in the

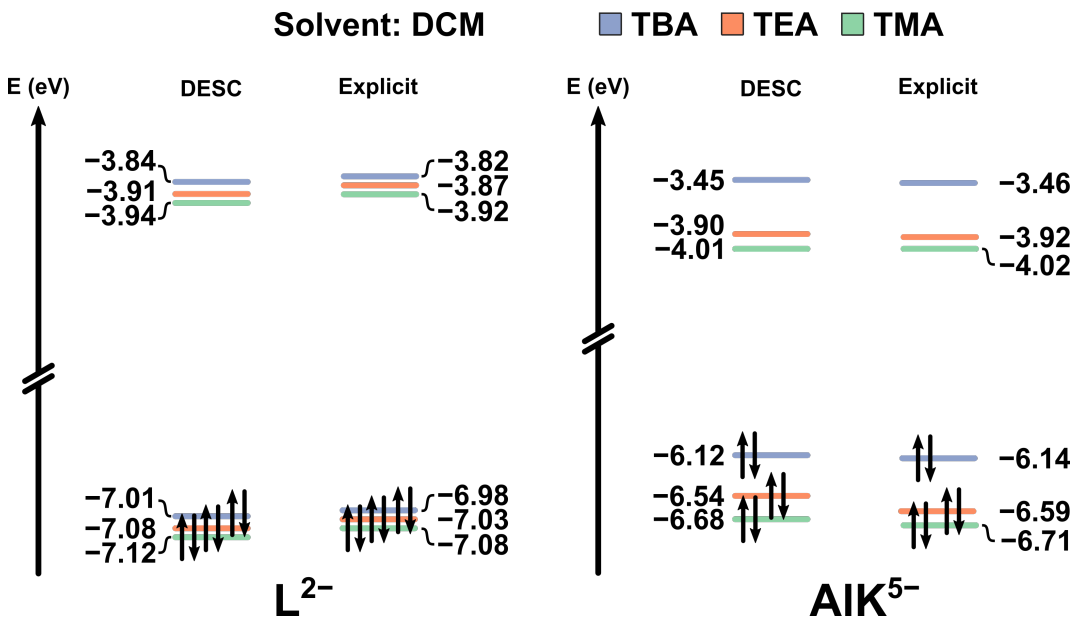


Figure 9: Molecular orbital diagram for **L** and **AlK** salts of TMA, TEA and TBA in fully oxidized state, in DCM calculated following two methodologies: DESC and Explicit+ISM (average of 5 snapshots).

present set of calculations——such as $(TEA)_7AlK$ and $(TBA)_7AlK$, especially in DCM——still largely outperform the homologous COSMO-based results.

Within the data presented in Figure 10, some differences are appreciable when comparing the performance of DESC in solutes with low or high charges. In the present set, fully oxidized POM states feature the lowest MAEs whereas slight discrepancies appear when calculating two-electron reduced states. The orbital differences become more pronounced, particularly in DCM, and may even lead to changes in orbital order for systems like $(TMA)_7AlK$ and $(TEA)_7AlK$. As noted earlier, this complexity likely arises because the effects that DESC simulates in a simple approach can be nonlinear and, thus, challenging to parameterize. Considering how the relative permittivity of the medium influences electrostatic interactions, it logically follows that DESC more accurately reproduces explicit calculations in less polar solvents, where each point charge retains a higher effective value, more perceptible to the solute.

Another advantage of DESC is its remarkable reduction in QM computational times with respect to explicit model calculations, which can be difficult to converge due to the large

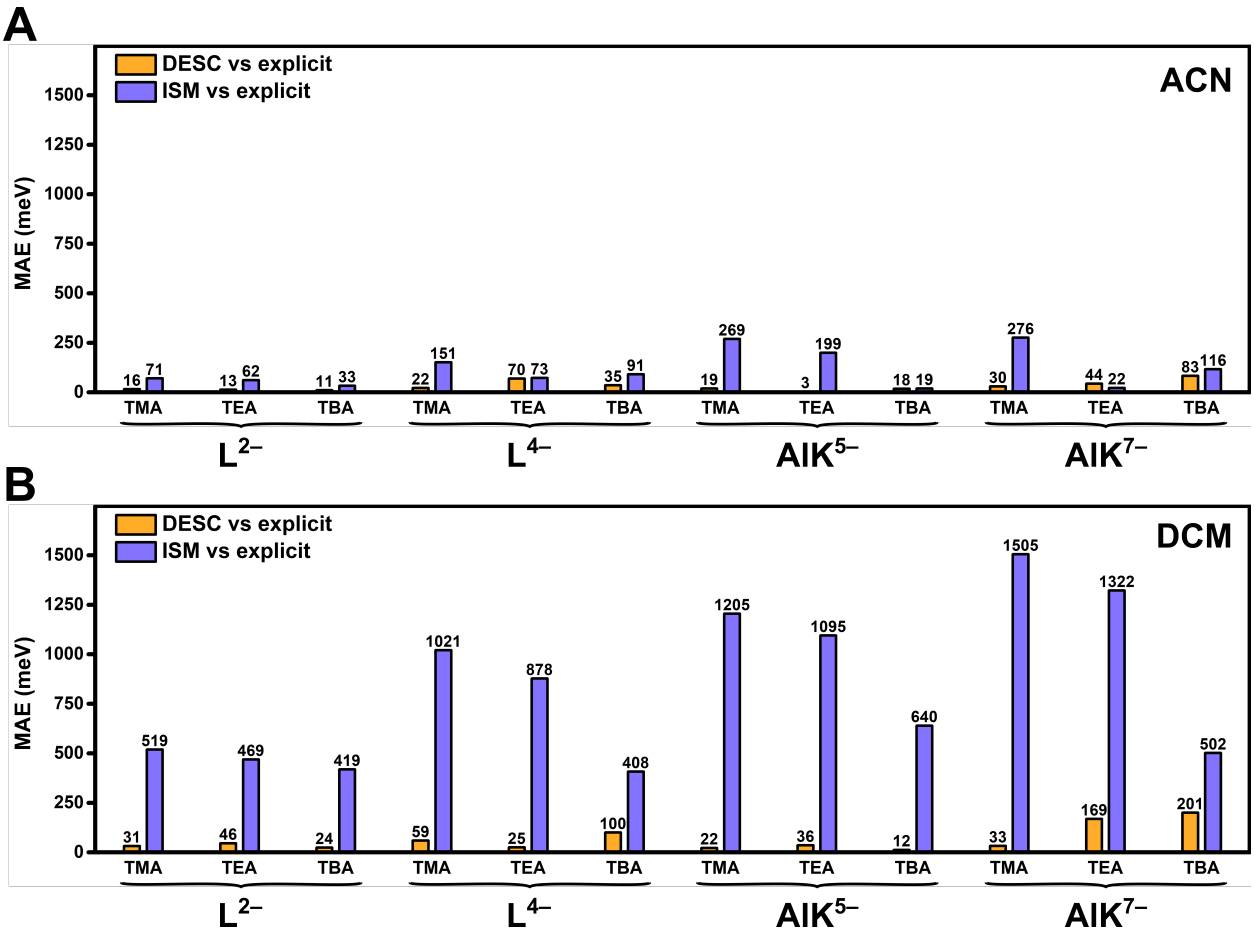


Figure 10: Performance of DESC for L and AlK systems. Orbital energy MAEs (vs explicit calculations) obtained with DESC (orange) and with ISM-only (purple) in (A) MeCN and (B) DCM solvents. Energy values for each group of MOs and each system are provided in Figures 3, 9, and Figures S2, S3, S7 and S8.

number of electrons and the complex, eccentric wave functions (or electronic densities) that may arise. In addition, because a series of explicit calculations (5-10) is recommended for representativeness of the average energy values. Thus, one of the main goals for developing DESC was to improve convergence by focusing solely on the solute electrons (excluding those of the co-solute), thereby reducing computational times.

Computational time analysis is shown in Figure 11. All results were obtained on a single node with 32-core EPYC 7282 processors, 128 GB of memory, and 960 GB SSD disks with Infiniband. Naturally, DESC largely improves the explicit model in terms of execution time, especially for bulky counterions such as TBA, since the explicit model requires QM

calculations on many more atoms (53 per cation, plus the solute). In general, the CPU times herein presented for the DESC methodology lie below 800 s. Only a few cases require up to 1000 s to execute. In contrast, most explicit calculations (for 5 snapshots) need much longer times to be completed, i.e. 3000 to 15000 s, with some cases even more expensive. Comparing these two methodologies, the speedup factors of DESC are typically 5 to 20, although for the largest counterions this factor can reach 40-50.

What is more, the present discussion does not account for the human time required for the tasks exclusively associated with the explicit model data processing, i.e., snapshot selection from the MD trajectory, *xyz* processing, input generation of QM runs, molecular orbital energy reading and averaging. Combined, these tasks are estimated to increase the total time by ca. 30 minutes (1800 s) per case, unless some automatization is at work. Also, several sources of error present during file processing in the explicit model approach are ruled out from DESC.

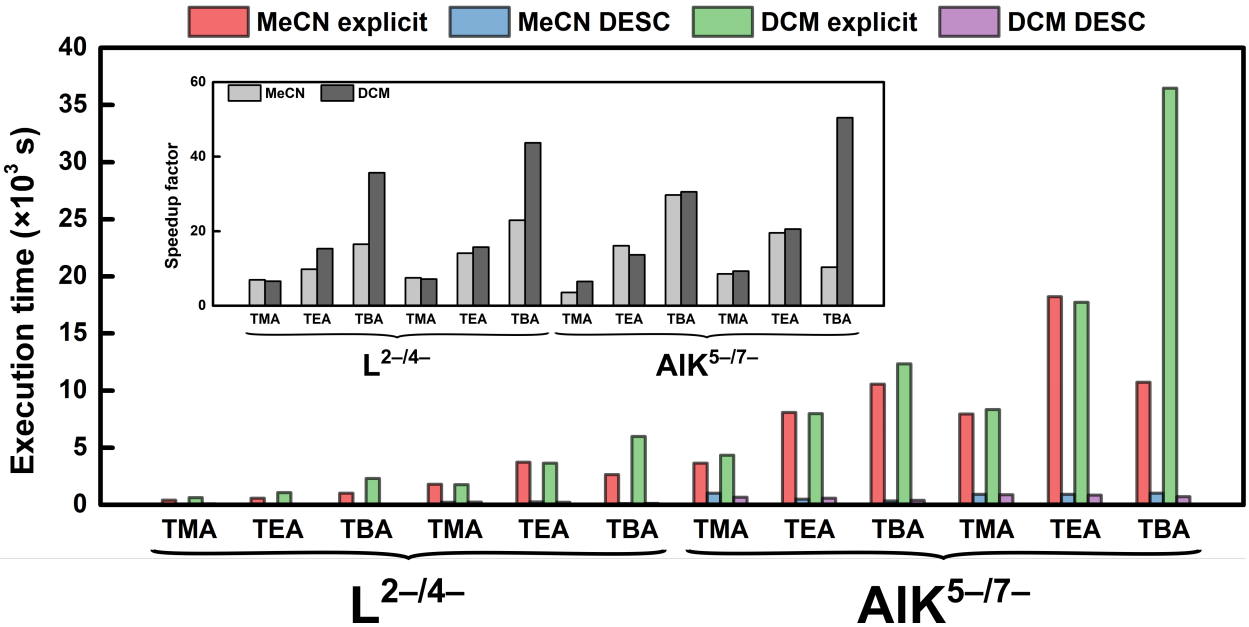


Figure 11: CPU times spent by DESC and explicit models for the $L^{2-/4-}$ and $AlK^{5-/7-}$ systems in MeCN and DCM. Inlay: speedup factor of DESC against explicit calculations. Detailed CPU times are provided in Table S2 in the ESI.

4 Testing DESC in other systems

The effectiveness of applying DESC to POMs has been demonstrated in the previous section, in terms of speedup and accuracy when compared to calculations performed with explicit cations, which constitutes our benchmark approach. It is important to note, however, that DESC was derived by taking the Lindqvist and the aluminum-Keggin tungstate as model systems. To assess its transferability to different compounds, we herein apply DESC to other POM species and to fullerenes, as well as with other solvents.

4.1 Polyoxometalates with lower symmetry in diverse solvents

Additional DESC calculations were performed on POMs with other morphologies and chemical compositions than those presented in Section 3. These include the non-spherical Wells-Dawson structure, $P_2W_{18}O_{62}$, and the vanadium-substituted Keggin phosphotungstate, $PV_2W_{10}O_{40}$ (Figure 12a). The latter compound carries the same molecular charge (5–) and is geometrically equivalent to $AlW_{12}O_{40}$ (AlK), but shows a distinct polarized charge distribution between its tungsten and vanadium regions, leading to differences in its electronic properties. To further validate the DESC approach, we also tested its performance in two other solvents: ethanol (EtOH, $\varepsilon = 24.55$) and ammonia (NH_3, $\varepsilon = 16.9$). Table 2 summarizes the combinations of POM-cation-solvent systems investigated and the average coordination numbers obtained (N), which highlight significant ion-pairing.

Table 2: Combinations of POM, cation and solvent simulated for application of DESC. The average aggregation of cations around the POM is indicated.

POM	**Cation**	**Solvent**	***N***
$[PW_{10}V_2O_{40}]^{5-/7-}$	TBA	MeCN	2.96 / 3.02
$[P_2W_{18}O_{62}]^{6-}$	TMA	MeCN	6.00
AlK^{7-}	TEA	EtOH	7.00
AlK^{7-}	TEA	NH_3	7.00

Figure 12 compares the orbital energies calculated with DESC and ISM approaches. DESC shows consistently low MAEs in the range 10–86 meV, significantly outperforming the regular ISM approach, which errors are >200meV in the majority of the cases, accurately reproducing the molecular orbital energies for $(TBA)_5PV_2W_{10}O_{40}$ in MeCN only. Despite its simplicity, Equation 4 proves remarkably robust in different solvents, as evidenced by its excellent performance in ethanol and ammonia (cases D and E in Table 2, respectively). Furthermore, DESC shows transferability to non-spherical solutes, which are traditionally more challenging for such analysis (case C).

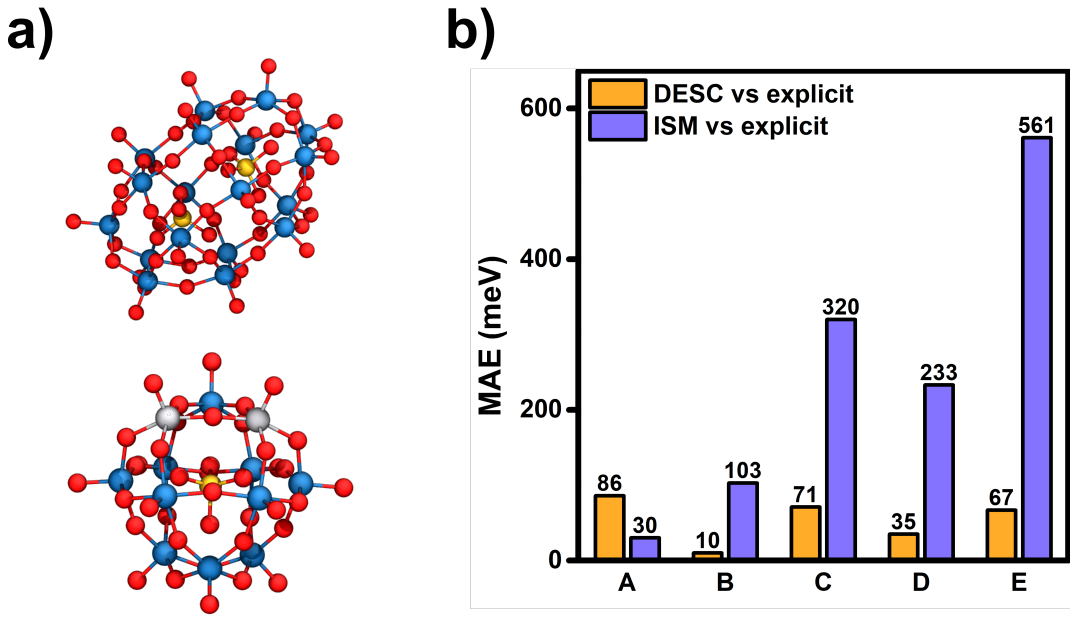


Figure 12: Performance of DESC in other POM systems: a) $[P_2W_{18}O_{62}]$ (top) and $[PV_2W_{10}O_{40}]$ (bottom). Color palette: W - blue, O - red, P - yellow, V - light gray; b) MAEs computed for orbital energies (vs explicit calculations) obtained with DESC (orange) and with ISM-only (purple) for: A: $(TBA)_5[PW_{10}V_2O_{40}]$ in MeCN, B: $(TBA)_7[PW_{10}V_2O_{40}]$ in MeCN, C: $(TMA)_6[P_2W_{18}O_{62}]$, D: $(TEA)_7$AlK in EtOH, and E: $(TEA)_7$AlK in NH_3.

4.2 Ion-pairing effects on reduced fullerenes

Fullerenes represent another important class of compounds, alongside POMs, where ion-pairing can play a critical role in some of their properties and applications. Although relatively underexplored, the effect of counterions on these carbon-based molecules is significant

due to their electrochemical relevance[57–59] and their applications as electron acceptors in photoactive systems.[60–62] Most experiments of fullerenes in solution involve supporting electrolytes (typically TBA in low-polarity or apolar solvents, such as dichloromethane, toluene or ortho-dichlorobenzene. Upon reduction, fullerenes are expected to associate strongly with cations in these media.

Although our group has focused for more than twenty years on investigating the structure and reactivity of endohedral metallofullerenes,[63–66] in this study we apply the DESC approach to investigate the multiple reduction states of the iconic C_{60}. Echegoyen and collaborators,[67] reported the six reductions of C_{60} in a 1:5.4 (% v/v) mixture of MeCN and toluene, with $TBAPF_6$ as supporting electrolyte. We have computationally mimicked these experimental conditions, focusing on the fully oxidized, singly reduced, and doubly reduced systems. For the fully oxidized state, no significant cation-fullerene association was observed, and as a result DESC and COSMO yielded identical molecular orbital energies for C_{60}(ox). In contrast, for the reduced species ${C_{60}}^{-}$ and ${C_{60}}^{2-}$, the analysis performed by DESC revealed strong ion pairing, with TBA more counterions surrounding closely the target molecule than the total negative charge of the fullerene, i.e. $N = 2.53$ and $N = 4.72$, respectively. This behavior is consistent with the low affinity of toluene for anionic fullerenes, which enhances ion-pairing. A graphical example of the TBA-${C_{60}}^{2-}$ ion-pairing captured by DESC is shown in Figure S9. Additionally, Figure S10 shows that toluene predominates as the solvent for the reduced states, a feature that DESC identifies automatically. Consequently, all explicit counterion calculations were performed with toluene as the solvent.

Table 3: Frontier molecular orbital energies (in eV) for the TBA-${C_{60}}^{-/2-}$ systems calculated using three methods: explicit cations, DESC, and ISM (COSMO).

	$\mathbf{{C_{60}}^{-}}$			$\mathbf{{C_{60}}^{2-}}$		
	Explicit	DESC	ISM	Explicit	DESC	ISM
HOMO	−5.086	−5.135	−3.173	−5.356	−5.477	−1.756
LUMO	−4.067	−4.113	−2.149	−4.412	−4.455	−0.731

The resulting frontier molecular orbital energies are shown in Table 3. DESC achieves excellent accuracy, with mean absolute errors of 47 meV for C_{60}^{-} and 83 meV for C_{60}^{2-}, in line to those reported for POMs. In contrast, COSMO yields very large errors (MAEs of 1920 meV and 3640 meV, respectively), highlighting its limitations in capturing counterion effects. Additionally, the reduction potential of C_{60}^{-} to C_{60}^{2-} calculated using DESC results in -1.39 V vs. $Fc^{+/0}$, in excellent agreement with the experimental value of -1.37 V vs. $Fc^{+/0}$. In comparison, the COSMO result significantly deviates with -3.26 V vs. $Fc^{+/0}$ (further details provided in the ESI). These results reinforce the effectiveness of DESC in accurately capturing the influence of counterions on the electronic properties of solutes, emphasizing its potential for broader applications.

5 Conclusion

The occurrence of molecular aggregation in solution, including ion pairing, is an important phenomenon affecting molecular properties and chemical processes. At the computational level, accounting for this particularity is not straightforward and requires an additional effort if the accurate electronic structure of the solute is the goal. We herein present a new computational strategy that accounts for complex solution environments, such as aggregation between a solute and the co-solutes, oriented to Quantum Mechanical calculations, called Dynamic Environment in Solution by Clustering (DESC). In such cases, this approach offers greater accuracy than conventional continuum implicit solvent models (ISMs), such as COSMO or PCM. DESC incorporates the specific effects of the co-solute in a statistically averaged, general and cost-effective manner. First, the dynamic nature of the bulk—comprising solvent, co-solutes (mostly counterions), and solute—is simulated using a regular classical Molecular Dynamics trajectory. DESC extracts the averaged number of aggregated counterions and their positions, and incorporates them into a standard QM calculation as effective point charges and bulky pseudo-atoms. By comparison to the explicit cations+ISM results

(the benchmark data), the DESC results largely outperform the standard ISM approach if aggregation occurs. The mean absolute errors in orbital energies obtained applying DESC mostly lie below 100 meV, whereas those for ISM are much larger, in many cases between 500 and 1500 meV. The other great advantage of DESC is the substantial reduction of computational time, if compared to equivalent QM calculations with explicit co-solutes, with speedup factors ranging 5-50 for the systems analyzed. Moreover, DESC skips the tedious hands-on work of preparing and averaging multiple QM calculations, as is the case for the explicit counterion approach that needs to account for the positional variability of counterions around the solute. This is an internal task performed by DESC in a few seconds.

Overall, the model presented here efficiently incorporates counterion mobility with a drastic reduction of computational effort and distinguishes the effects of different counterions, as evidenced by changes in the solute molecular orbital energies as the co-solute varies. This strategy was initially developed and tested with highly symmetrical polyoxometalate systems as target solutes but, in principle, there are no limitations in this aspect. Hence, any combination of solute, co-solute and solvent can be analyzed and processed by DESC. Tests conducted on different systems, such as non-symmetrical POMs and a fullerene, demonstrate that the methodology is robust also for solutes of other nature, clearly outperforming classical ISMs.

Further work on this model is ongoing in our group to generalize the current performance, especially when strong ion-pairing occurs, which might originate in important changes in the bulk solvent's relative permittivity. Classical electrolyte theories, such as Debye-Hückel theory or Born and Poisson-Boltzmann solvation models, are valid only in sufficiently dilute solutions,[68] where ion-ion correlations are negligible. Future improvements to the model could involve addressing this issue by refining the treatment of point charges, potentially employing a screened Coulomb potential, such as a Yukawa function or similar approach.[69] Yet another variation to the present version will include accounting correctly for the effects of solvent mixtures.

The associated DESC Python code is open-access and can be freely installed.

Acknowledgement

The authors thank the economic support of the Spanish Ministry of Science and Innovation (grants PID2020-112762GB-I00 and PID2023-149905NB-I00, both funded by MCIN/AEI/ 10.13039/501100011033) and the Generalitat de Catalunya (grant 2021 SGR 00110). We also thank the colleagues at the Quantum Chemistry Group for fruitful discussions.

Supporting Information Available

Tables: DFT-optimized and DESC-generated interatomic distances for **AlK**; CPU times for DESC and explicit calculations. Figures: Solute-counterion RDFs; pre-DESC reorientation step for MD trajectories; molecular orbital energies obtained with DESC vs. the number of frames taken from the MD trajectory; comparisons of molecular orbital energies in DCM and MeCN solvents. Additional text: DESC-related files (main input, logfile, generated ADF input); redox information related to the C_{60} system.

References

(1) Born, M. Volumen und Hydratationswärme der Ionen. *Z. Phys.* **1920**, *1*, 45–48.

(2) Varghese, J. J.; Mushrif, S. H. Origins of complex solvent effects on chemical reactivity and computational tools to investigate them: a review. *React. Chem. Eng.* **2019**, *4*, 165–206.

(3) Bruessel, M.; Zahn, S.; Hey-Hawkins, E.; Kirchner, B. Theoretical Investigation of Solvent Effects and Complex Systems: Toward the calculations of bioinorganic systems

from ab initio molecular dynamics simulations and static quantum chemistry. *Adv. Inorg. Chem.* **2010**, *62*, 111–142.

(4) Cammi, R.; Mennucci, B.; Tomasi, J. Computational modelling of the solvent effects on molecular properties: an overview of the polarizable continuum model (PCM) approach. *Computational chemistry: reviews of current trends* **2003**, 1–79.

(5) Garay-Ruiz, D.; Bo, C. Rationalizing the mechanism of peroxyformate decomposition: computational insights to understand solvent influence. *Chem.–Eur. J.* **2021**, *27*, 11618–11626.

(6) Izzet, G.; Abécassis, B.; Brouri, D.; Piot, M.; Matt, B.; Serapian, S. A.; Bo, C.; Proust, A. Hierarchical self-assembly of polyoxometalate-based hybrids driven by metal coordination and electrostatic interactions: from discrete supramolecular species to dense monodisperse nanoparticles. *J. Am. Chem. Soc.* **2016**, *138*, 5093–5099.

(7) Skyner, R.; McDonagh, J.; Groom, C.; van Mourik, T. v.; Mitchell, J. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174–6191.

(8) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3094.

(9) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(10) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117–129.

(11) Miertuš, S.; Tomasi, J. Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chem. Phys.* **1982**, *65*, 239–245.

(12) Pascual-Ahuir, J.-L.; Silla, E.; Tuñón, I. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface. *J. Comput. Chem.* **1994**, *15*, 1127–1138.

(13) Toupalas, G.; Karlsson, J.; Black, F. A.; Masip-Sánchez, A.; López, X.; Ben M'Barek, Y.; Blanchard, S.; Proust, A.; Alves, S.; Chabera, P.; Clark, I. P.; Pullerits, T.; Poblet, J. M.; Gibson, E. A.; Izzet, G. Tuning Photoinduced Electron Transfer in POM-Bodipy Hybrids by Controlling the Environment: Experiment and Theory. *Angew. Chem. Int. Ed.* **2021**, *60*, 6518–6525.

(14) Malkani, A.; Li, J.; Oliveira, N.; He, M.; Chang, X.; Xu, B.; Lu, Q. Understanding the electric and nonelectric field components of the cation effect on the electrochemical CO reduction reaction. *Sci. Adv.* **2020**, *6*, eabd2569.

(15) Waegele, M. M.; Gunathunge, C. M.; Li, J.; Li, X. How cations affect the electric double layer and the rates and selectivity of electrocatalytic processes. *J. Chem. Phys.* **2019**, *151*.

(16) Shin, S.-J.; Choi, H.; Ringe, S.; Won, D. H.; Oh, H.-S.; Kim, D. H.; Lee, T.; Nam, D.-H.; Kim, H.; Choi, C. H. A unifying mechanism for cation effect modulating C1 and C2 productions from CO_2 electroreduction. *Nat. Commun.* **2022**, *13*, 5482.

(17) López, X.; Carbó, J. J.; Bo, C.; Poblet, J. M. Structure, properties and reactivity of polyoxometalates: a theoretical perspective. *Chem. Soc. Rev.* **2012**, *41*, 7537–7571.

(18) López, X.; Maestre, J. M.; Bo, C.; Poblet, J. M. Electronic properties of polyoxometalates: A DFT study of α/β-$[XM_{12}O_{40}]^{n-}$ relative stability (M = W, Mo and X a main group element). *J. Am. Chem. Soc.* **2001**, *123*, 9571–9576.

(19) López, X.; Poblet, J. M. DFT study on the five isomers of $[PW_{12}O_{40}]^{3-}$: Relative stabilization upon reduction. *Inorg. Chem.* **2004**, *43*, 6863–6865.

(20) López, X.; de Graaf, C.; Maestre, J. M.; Bénard, M.; Rohmer, M.-M.; Bo, C.; Poblet, J. M. Highly reduced polyoxometalates: ab initio and DFT study of $[PMo_8V_4O_{40}(VO)_4]^{5-}$. *J. Chem. Theory Comput.* **2005**, *1*, 856–861.

(21) Miro, P.; Poblet, J. M.; Ávalos, J. B.; Bo, C. Towards a computational treatment of polyoxometalates in solution using QM methods and explicit solvent molecules. *Can. J. Chem.* **2009**, *87*, 1296–1301.

(22) Solé-Daura, A.; Goovaerts, V.; Stroobants, K.; Absillis, G.; Jiménez-Lozano, P.; Poblet, J. M.; Hirst, J. D.; Parac-Vogt, T. N.; Carbó, J. J. Probing polyoxometalate–protein interactions using molecular dynamics simulations. *Chem.–Eur. J.* **2016**, *22*, 15280–15289.

(23) Segado, M.; Nyman, M.; Bo, C. Aggregation patterns in low-and high-charge anions define opposite solubility trends. *J. Phys. Chem. B* **2019**, *123*, 10505–10513.

(24) Maksimchuk, N. V.; Puiggalí-Jou, J.; Zalomaeva, O. V.; Larionov, K. P.; Evtushok, V. Y.; Soshnikov, I. E.; Solé-Daura, A.; Kholdeeva, O. A.; Poblet, J. M.; Carbó, J. J. Resolving the Mechanism for H_2O_2 Decomposition over Zr(IV)-Substituted Lindqvist Tungstate: Evidence of Singlet Oxygen Intermediacy. *ACS Catal.* **2023**, *13*, 10324–10339.

(25) Tzaguy, A.; Masip-Sánchez, A.; Avram, L.; Solé-Daura, A.; López, X.; Poblet, J. M.; Neumann, R. Electrocatalytic Reduction of Dinitrogen to Ammonia with Water as Proton and Electron Donor Catalyzed by a Combination of a Tri-ironoxotungstate and an Alkali Metal Cation. *J. Am. Chem. Soc.* **2023**, *145*, 19912–19924.

(26) Kremleva, A.; Rösch, N. Modeling the effect of the electrolyte on standard reduction potentials of polyoxometalates. *J. Phys. Chem. C* **2018**, *122*, 18545–18553.

(27) Matsui, T.; Kitagawa, Y.; Shigeta, Y.; Okumura, M. A density functional theory based protocol to compute the redox potential of transition metal complex with the correction of pseudo-counterion: General theory and applications. *J. Chem. Theory Comput.* **2013**, *9*, 2974–2980.

(28) Hruska, E.; Gale, A.; Liu, F. Bridging the experiment-calculation divide: machine learning corrections to redox potential calculations in implicit and explicit solvent models. *Journal of Chemical Theory and Computation* **2022**, *18*, 1096–1108.

(29) Roy, L. E.; Jakubikova, E.; Guthrie, M. G.; Batista, E. R. Calculation of one-electron redox potentials revisited. Is it possible to calculate accurate potentials with density functional methods? *The Journal of Physical Chemistry A* **2009**, *113*, 6745–6750.

(30) Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. Computational electrochemistry: prediction of liquid-phase reduction potentials. *Physical Chemistry Chemical Physics* **2014**, *16*, 15068–15106.

(31) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(32) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(33) Chaumont, A.; Wipff, G. Polyoxometalate Keggin Anions at Aqueous Interfaces with Organic Solvents, Ionic Liquids, and Graphite: a Molecular Dynamics Study. *J. Phys. Chem. C* **2009**, *113*, 18233–18243.

(34) López, X.; Nieto-Draghi, C.; Bo, C.; Avalos, J. B.; Poblet, J. M. Polyoxometalates in Solution: Molecular Dynamics Simulations on the α-$[PW_{12}O_{40}]^{3-}$ Keggin Anion in Aqueous Media. *J. Phys. Chem. A* **2005**, *109*, 1216–1222.

(35) Masip-Sánchez, A. topoMOx: Topologies for Metal Oxides. 2024; `https://github.com/qcgurv/topoMOx`.

(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian˜16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

(37) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(38) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8*, 61–74.

(39) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(40) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(41) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.

(42) Pye, C. C.; Ziegler, T. An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package. *Theor. Chem. Acc.* **1999**, *101*, 396–408.

(43) Errington, R. J.; Petkar, S. S.; Middleton, P. S.; McFarlane, W.; Clegg, W.; Coxall, R. A.; Harrington, R. W. Synthesis and Reactivity of the Methoxozirconium Pentatungstate (nBu$_4$N)$_6$[{(μ–MeO)ZrW$_5$O$_{18}$}$_2$]: Insights into Proton-Transfer Reactions, Solution Dynamics, and Assembly of $\{ZrW_5O_{18}\}^{2-}$ Building Blocks. *J. Am. Chem. Soc.* **2007**, *129*, 12181–12196.

(44) Priyadarshini, M.; Shanmugan, S.; Kirubakaran, K. P.; Thomas, A.; Prakash, M.; Senthil, C.; Lee, C. W.; Vediappan, K. High energy storage of Li-ions on keggin-type polyoxometalate as electrodes for rechargeable lithium batteries. *J. Phys. Chem. Solids* **2020**, *142*, 109468.

(45) Solé-Daura, A.; Notario-Estévez, A.; Carbó, J. J.; Poblet, J. M.; de Graaf, C.; Monakhov, K. Y.; López, X. How Does the Redox State of Polyoxovanadates Influence the Collective Behavior in Solution? A Case Study with $[I@V_{18}O_{42}]^{q-}$ (q = 3, 5, 7, 11, and 13). *Inorg. Chem.* **2019**, *58*, 3881–3894.

(46) Chaumont, A.; Wipff, G. Ion aggregation in concentrated aqueous and methanol solutions of polyoxometallates Keggin anions: the effect of counterions investigated by molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6940–6953.

(47) Kremleva, A.; Aparicio, P. A.; Genest, A.; Rösch, N. Quantum chemical modeling of

tri-Mn-substituted W-based Keggin polyoxoanions. *Electrochim. Acta* **2017**, *231*, 659–669.

(48) Tzaguy, A.; Masip-Sánchez, A.; Avram, L.; Solé-Daura, A.; López, X.; Poblet, J. M.; Neumann, R. Electrocatalytic Reduction of Dinitrogen to Ammonia with Water as Proton and Electron Donor Catalyzed by a Combination of a Tri-ironoxotungstate and an Alkali Metal Cation. *J. Am. Chem. Soc.* **2023**, *145*, 19912–19924.

(49) Masip-Sánchez, A.; López, X. DESC: Dynamic Environment in Solution by Clustering. 2024; `https://github.com/qcgurv/DESC`.

(50) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.

(51) Richard J. Gowers; Max Linke; Jonathan Barnoud; Tyler J. E. Reddy; Manuel N. Melo; Sean L. Seyler; Jan Domåski; David L. Dotson; Sébastien Buchoux; Ian M. Kenney; Oliver Beckstein MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the 15th Python in Science Conference. 2016; pp 98 – 105.

(52) Masip-Sánchez, A. TCL Script for Aligning Trajectories in VMD. 2024; `https://github.com/qcgurv/DESC/tree/main/scripts`.

(53) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contrib-

utors SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272.

(54) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(55) Hubert, L.; Arabie, P. Comparing partitions. *J. Classif.* **1985**, *2*, 193–218.

(56) Allinger, N. L.; Zhou, X.; Bergsma, J. Molecular mechanics parameters. *J. Mol. Struct.: THEOCHEM* **1994**, *312*, 69–83.

(57) Chaur, M. N.; Melin, F.; Ortiz, A. L.; Echegoyen, L. Chemical, electrochemical, and structural properties of endohedral metallofullerenes. *Angew. Chem. Int. Ed.* **2009**, *48*, 7514–7538.

(58) Puente Santiago, A. R.; Fernandez-Delgado, O.; Gomez, A.; Ahsan, M. A.; Echegoyen, L. Fullerenes as key components for low-dimensional (photo) electrocatalytic nanohybrid materials. *Angew. Chem. Int. Ed.* **2021**, *60*, 122–141.

(59) Balch, A. L.; Winkler, K. Electrochemistry of fullerene/transition metal complexes: Three decades of progress. *Coord. Chem. Rev.* **2021**, *438*, 213623.

(60) Rudolf, M.; Kirner, S.; Guldi, D. A multicomponent molecular approach to artificial photosynthesis–the role of fullerenes and endohedral metallofullerenes. *Chem. Soc. Rev.* **2016**, *45*, 612–630.

(61) El-Khouly, M. E.; Ito, O.; Smith, P. M.; D'Souza, F. Intermolecular and supramolecular photoinduced electron transfer processes of fullerene–porphyrin/phthalocyanine systems. *J. Photoch. Photobio. C* **2004**, *5*, 79–104.

(62) Iagatti, A.; Cupellini, L.; Biagiotti, G.; Caprasecca, S.; Fedeli, S.; Lapini, A.; Ussano, E.; Cicchi, S.; Foggi, P.; Marcaccio, M.; Mennucci, B.; Di Donato, M. Efficient photoinduced charge separation in a BODIPY–C_{60} dyad. *J. Phys. Chem. C* **2016**, *120*, 16526–16536.

(63) Campanera, J. M.; Bo, C.; Olmstead, M. M.; Balch, A. L.; Poblet, J. M. Bonding within the Endohedral Fullerenes $Sc_3N@C_{78}$ and $Sc_3N@C_{80}$ as Determined by Density Functional Calculations and Reexamination of the Crystal Structure of $Sc_3N@C_{78}\cdot Co(OEP)\cdot 1.5C_6H_6\cdot 0.3CHCl_3$. *J. Phys. Chem. A* **2002**, *106*, 12356–12364.

(64) Valencia, R.; Rodríguez-Fortea, A.; Clotet, A.; de Graaf, C.; Chaur, M. N.; Echegoyen, L.; Poblet, J. M. Electronic structure and redox properties of metal nitride endohedral fullerenes $M_3N@C_{2n}$ (M = Sc, Y, La, and Gd; 2n = 80, 84, 88, 92, 96). *Chem.–Eur J* **2009**, *15*, 10997–11009.

(65) Puente Santiago, A. R.; Sanad, M. F.; Moreno-Vicente, A.; Ahsan, M. A.; Cerón, M. R.; Yao, Y.-R.; Sreenivasan, S. T.; Rodriguez-Fortea, A.; Poblet, J. M.; Echegoyen, L. A new class of molecular electrocatalysts for hydrogen evolution: Catalytic activity of $M_3N@C_{2n}$ (2n = 68, 78, and 80) fullerenes. *J. Am. Chem. Soc.* **2021**, *143*, 6037–6042.

(66) Qiu, J.; Abella, L.; Du, X.; Cao, Z.; He, Z.; Meng, Q.; Yan, Y.; Poblet, J. M.; Sun, L.; Rodríguez-Fortea, A.; Chen, N. $CaY@C_{2n}$: Exploring Molecular Qubits with Ca–Y Metal–Metal Bonds. *J. Am. Chem. Soc.* **2024**, *146*, 24310–24319.

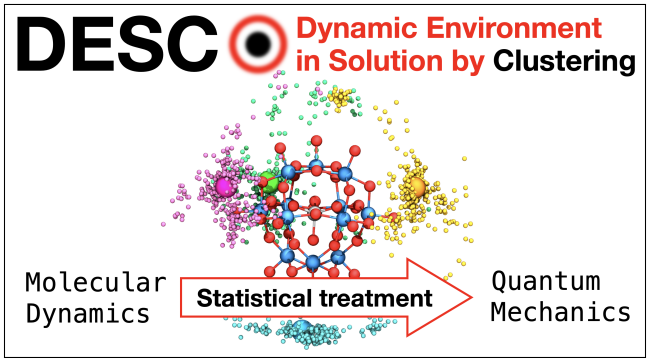
(67) Xie, Q.; Perez-Cordero, E.; Echegoyen, L. Electrochemical detection of $C_{60}^{6-}$and C_{70}^{6-}: Enhanced stability of fullerides in solution. *J. Am. Chem. Soc.* **1992**, *114*, 3978–3980.

(68) Ramirez, R.; Kjellander, R. Effective multipoles and Yukawa electrostatics in dressed molecule theory. *J. Chem. Phys.* **2006**, *125*, 144110.

(69) Forsberg, B.; Ulander, J.; Kjellander, R. Dressed ion theory of size-asymmetric elec-

trolytes: Effective ionic charges and the decay length of screened Coulomb potential and pair correlations. *J. Chem. Phys.* **2005**, *122*, 064502.

TOC Graphic


DESC
Dynamic Environment
in Solution by Clustering
Molecular
Dynamics
Statistical treatment
Quantum
Mechanics