COMPUTATIONAL METHODS FOR THE DESCRIPTION OF INTERMOLECULAR INTERACTIONS AND MOLECULAR MOTION IN CONFINING ENVIRONMENTS

EDITED BY: Heribert Reis, Piotr Żuchowski and Sonja Grubisic
PUBLISHED IN: Frontiers in Chemistry

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-410-5
DOI 10.3389/978-2-88976-410-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

COMPUTATIONAL METHODS FOR THE DESCRIPTION OF INTERMOLECULAR INTERACTIONS AND MOLECULAR MOTION IN CONFINING ENVIRONMENTS

Topic Editors:
Heribert Reis, National Hellenic Research Foundation, Greece
Piotr Żuchowski, Nicolaus Copernicus University in Toruń, Poland
Sonja Grubisic, University of Belgrade, Serbia

Citation: Reis, H., Żuchowski, P., Grubisic, S., eds. (2022). Computational Methods for the Description of Intermolecular Interactions and Molecular Motion in Confining Environments. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-410-5

Table of Contents

frontiers | Frontiers in Chemistry

EDITORIAL
published: 31 May 2022
doi: 10.3389/fchem.2022.941269

Editorial: Computational Methods for the Description of Intermolecular Interactions and Molecular Motion in Confining Environments

Heribert Reis[1]*, *Piotr Żuchowski*[2] and *Sonja Grubisic*[3]

[1]Institute of Chemical Biology, National Hellenic Research Foundation, Athens, Greece, [2]Institute of Physics, Nicolaus Copernicus University in Toruń, Toruń, Poland, [3]Institute of Chemistry, Technology, and Metallurgy, National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia

Keywords: molecular confinement, nuclear motion, intermolecular interactions, molecular simulation methods, carbon nanotubes, SAPT

Editorial on the Research Topic

Computational Methods for the Description of Intermolecular Interactions and Molecular Motion in Confining Environments

Methods of computational chemistry have become increasingly important in the last decades for the investigation of interaction and dynamics of small molecules enclosed in larger, micro- and mesoscopic structures, as witnessed by a large number of publications in a large number of diverse fields, such as design of pharmaceutical drugs Roy (2019), investigations on the mechanisms of drug delivery Garifalo et al. (2020), design of novel materials for bioimaging Bonačić-Koutecký and Antoine (2019); Combes et al. (2021), catalysis Li et al. (2021), gas storage Kundu et al. (2016), information and communication technologies Ghosh et al. (2021), but also on fundamental research, for example in astrophysics Puzzarini and Barone (2020); de Lara-Castells and Hauser (2020). Molecular confinement in single molecular structures may lead to new and unexpected phenomena, which are not easily predicted by classical theories of confinement in bulk materials. Due to the small to medium size of these confining structures, they can be treated in principle by modern computational methods with moderate to high accuracy, thus allaying the need for difficult experiments. However, new techniques and methodologies may be required for successful treatment of interactions and dynamics of molecules in confining structures. It is the purpose of this Research Topic of ten original research contributions to highlight current directions and present some applications of computational methods for molecular systems in confining environments.

Recent developments and an outlook on possible developments in the future on computational approaches for the treatment of small clusters comprised of He atoms or H_2 molecules in carbon nanotubes are reviewed in (de Lara-Castells and Mitrushchenkov). For an accurate description of these dispersion–dominated systems, difficult to treat by standard quantum-mechanical methods, a multi-scale approach using a DFT-based symmetry-adapted perturbation theory (SAPT) method is described. A second, wave-function based method, also allows the computation of shallow bound states in these systems.

Two contributions attest to the importance of research on the behaviour of small molecules in the confining environment of natural proteins in pharmaceutical and medicinal research. Zhang et al. use classical simulation techniques, i.e. molecular docking, molecular dynamics simulations and adaptive steered molecular dynamics simulations to investigate the interactions of two inhibitors drugs with the N-terminal Human Maltase Glucoamylase, in an effort to advance the search for effective drugs to treat diabetes 2.

Jiang et al. combine experiment and computational molecular simulations to study the interactions of two alternaria mycotoxins with a single-chain antibody fragment.

Edited and reviewed by:
Sam P. De Visser,
The University of Manchester,
United Kingdom

***Correspondence:**
Heribert Reis
hreis@eie.gr

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *11 May 2022*
Accepted: *16 May 2022*
Published: *31 May 2022*

Citation:
Reis H, Żuchowski P and Grubisic S (2022) Editorial: Computational Methods for the Description of Intermolecular Interactions and Molecular Motion in Confining Environments. Front. Chem. 10:941269. doi: 10.3389/fchem.2022.941269

Another study using classical molecular dynamics simulations to investigate a Research Topic of confinement of biological importance is presented by Vasiliu et al. They study the interactions between positively charged polyamines which act as counterions to negatively charged DNA fragments in cells using microsecond molecular dynamics simulations.

A contribution from the field of chemical engineering on the effect of confinement on solubilties of gases in ionic liquids enclosed by nanopores is reported by Sun et al., who improved the efficiency and applicability of electrolyte perturbed-chain statistical associating fluid theory–density functional theory (ePC-SAFT-DFT). The newly devised method was then applied to study the effect of several variables on the solubiltity of CO_2 in ionic liquids confined in silica nanopores, which led to new conclusions how to increase the solubility of CO_2 in ionic liquids.

Three studies are concerned with the effect and analysis of intermolecular interactions, clearly a dominant Research Topic for confined systems. These interactions may lead, for example, to hydrogen bonding, which are the focus of the first two papers, or other types of noncovalent bondings, such as halogen or chalcogen bonds.

In the first paper, Kaczmarek-Kedziera et al. present a study on the effect of intermolecular interactions, specifically hydrogen bonding, on two optical properties, namely electronic one- and two-photon absorption, applying quantum-mechanical methods on hydrogen bonded intermolecular complexes of two 4-substituted N,N′-diphenyl-urea or N,N′-diphenyl-thiourea with one central squaraine molecule. Using different substituents in position 4 of phenyl the acidity of the N-H protons can be varied.

In the second paper, the effect of a confining environment on the IR spectra of two hydrogen-bonded intermolecular complexes, namely HCN—HCN and HCN—HNC, is studied by Chołuj et al. The spatially confining environment is simulated by embedding the complexes in an external cylindrical harmonic oscillator potential, and the harmonic frequencies are computed with high-level *ab-initio* methods The simulated confinement leads to substantial changes of both transition intensities and vibrational frequencies.

In the third study Szczesniak and Chalasinski, report an interesting analysis of a number of complexes with a more general type of intermolecular interactions, i.e., weak electron-donor electron-acceptor complexes including examples of chalcogen and halogen bonds. With the help of Symmetry Adapted Perturbation Theory (SAPT), they challenge the prevailing interpretation of these interactions as of an electrostatic nature. Instead, they argue that exchange or Pauli repulsion is more relevant.

A question relevant for astrochemistry is studied by Upadhyay and Meuwly, namely the distribution of energy released due to exothermic reactions on amorphous solid water. Specifically, they investigate the energy transfer during energy relaxation of the product in the recombination reaction $CO(^1\Sigma^+)$ + O (^{1}D) → $CO_2(^1\Sigma_g^+)$ within and on the surface of amorphous solid water and the surrounding water molecules.

Finally, a novel embedding simulation method, called real-time time-dependent block orthogonalized Manby-Miller embedding approach is implemented by de Santis et al., and applied to study the X-ray absorption spectra of model systems comprised of fluoride and chloride ions confined by 8 water molecules ($[X@(H_2O)_8]^-$, X=Cl,F).

The field of computational techniques to treat confined systems in the extended definition used here is very large and extremely varied, and this Selection can only present a limited number of recent applications and developments in this Research Topic. Nevertheless, we believe that the selected papers are evidence of the vibrant activity of current research in this field. We expect that the role of computational treatments of confined systems will become more important in the future, due to the increasing search for systems with specific properties for applications on the microscopic and nanocscopic scale, as well as due to the steadily increase of computing power and the ongoing development of sophisticated modelling techniques.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

The authors thank Maria Pilar de Lara-Castells for her interest and encouragement during the compilation of this Research Topic.

REFERENCES

Bonačić-Koutecký, V., and Antoine, R. (2019). Enhanced Two-Photon Absorption of Ligated Silver and Gold Nanoclusters: Theoretical and Experimental Assessments. *Nanoscale* 11, 12436–12448. doi:10.1039/C9NR01826C

Combes, G. F., Vučković, A.-M., Perić Bakulić, M., Antoine, R., Bonačić-Koutecky, V., and Trajković, K. (2021). Nanotechnology in Tumor Biomarker Detection: The Potential of Liganded Nanoclusters as Nonlinear Optical Contrast Agents for Molecular Diagnostics of Cancer. *Cancers* 13, 4206. doi:10.3390/cancers13164206

de Lara-Castells, M. P., and Hauser, A. W. (2020). New Tools for the Astrochemist: Multi-Scale Computational Modelling and Helium Droplet-Based Spectroscopy. *Phys. Life Rev.* 32, 95–98. doi:10.1016/j.plrev.2019.08.001

Garofalo, M., Grazioso, G., Cavalli, A., and Sgrignani, J. (2020). How Computational Chemistry and Drug Delivery Techniques Can Support the Development of New Anticancer Drugs. *Molecules* 25, 1756. doi:10.3390/molecules25071756

Ghosh, D., Ivanov, S. A., and Tretiak, S. (2021). Structural Dynamics and Electronic Properties of Semiconductor Quantum Dots: Computational Insights. *Chem. Mat.* 33, 7848–7857. doi:10.1021/acs.chemmater.1c02514

Kundu, A., Piccini, G., Sillar, K., and Sauer, J. (2016). Ab Initio Prediction of Adsorption Isotherms for Small Molecules in Metal-Organic Frameworks. *J. Am. Chem. Soc.* 138, 14047–14056. doi:10.1021/jacs.6b08646

Li, S., Wang, Y., Qin, B., Zhou, Z., Zhou, S., Li, K., et al. (2021). "Mechanistic Studies toward the Rational Design of Oxide Catalysts for Carbon Dioxide Hydrogenation," in *Annual Reports in Computational Chemistry*. Editors D. A. Dixon and R. R. Chair (Amsterdam: Elsevier), Vol. 17, 211–270. doi:10.1016/bs.arcc.2021.09.001

Puzzarini, C., and Barone, V. (2020). A Never-Ending Story in the Sky: The Secrets of Chemical Evolution. *Phys. Life Rev.* 32, 59–94. doi:10.1016/j.plrev.2019.07.001

Roy K. (Editor) (2019). *In Silico Drug Design. Repurposing Techniques and Methodologies* (Cambridge, Massachusetts, USA: Acadenic Press).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 30 August 2021
doi: 10.3389/fchem.2021.711242

Targeting N-Terminal Human Maltase-Glucoamylase to Unravel Possible Inhibitors Using Molecular Docking, Molecular Dynamics Simulations, and Adaptive Steered Molecular Dynamics Simulations

Shitao Zhang[1], *Yi Wang*[1], *Lu Han*[1], *Xueqi Fu*[1], *Song Wang*[2], *Wannan Li*[1]* *and Weiwei Han*[1]*

[1]School of Life Science, Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, National Engineering Laboratory for AIDS Vaccine, Jilin University, Changchun, China, [2]Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun, China

OPEN ACCESS

Edited by:
Sonja Grubisic, University of Belgrade, Serbia

Reviewed by:
Dragan M. Popovic, University of Belgrade, Serbia
Francesca Mocci, University of Cagliari, Italy

***Correspondence:**
Wannan Li
liwannan@jlu.edu.cn
Weiwei Han
weiweihan@jlu.edu.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *18 May 2021*
Accepted: *26 July 2021*
Published: *30 August 2021*

Citation:
Zhang S, Wang Y, Han L, Fu X, Wang S, Li W and Han W (2021) Targeting N-Terminal Human Maltase-Glucoamylase to Unravel Possible Inhibitors Using Molecular Docking, Molecular Dynamics Simulations, and Adaptive Steered Molecular Dynamics Simulations. Front. Chem. 9:711242. doi: 10.3389/fchem.2021.711242

There are multiple drugs for the treatment of type 2 diabetes, including traditional sulfonylureas biguanides, glinides, thiazolidinediones, α-glucosidase inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists, dipeptidyl peptidase IV (DPP-4) inhibitors, and sodium-glucose cotransporter 2 (SGLT2) inhibitors. α-Glucosidase inhibitors have been used to control postprandial glucose levels caused by type 2 diabetes since 1990. α-Glucosidases are rather crucial in the human metabolic system and are principally found in families 13 and 31. Maltase-glucoamylase (MGAM) belongs to glycoside hydrolase family 31. The main function of MGAM is to digest terminal starch products left after the enzymatic action of α-amylase; hence, MGAM becomes an efficient drug target for insulin resistance. In order to explore the conformational changes in the active pocket and unbinding pathway for NtMGAM, molecular dynamics (MD) simulations and adaptive steered molecular dynamics (ASMD) simulations were performed for two NtMGAM-inhibitor [de-O-sulfonated kotalanol (DSK) and acarbose] complexes. MD simulations indicated that DSK bound to NtMGAM may influence two domains (inserted loop 1 and inserted loop 2) by interfering with the spiralization of residue 497–499. The flexibility of inserted loop 1 and inserted loop 2 can influence the volume of the active pocket of NtMGAM, which can affect the binding progress for DSK to NtMGAM. ASMD simulations showed that compared to acarbose, DSK escaped from NtMGAM easily with lower energy. Asp542 is an important residue on the bottleneck of the active pocket of NtMGAM and could generate hydrogen bonds with DSK continuously. Our theoretical results may provide some useful clues for designing new α-glucosidase inhibitors to treat type 2 diabetes.

Keywords: maltase-glucoamylase, inhibitors, molecular dynamics simulations, adaptive steered molecular dynamics simulations, conformational change

Abbreviations: MGAM, maltase-glucoamylase; NtMGAM, N-terminal human maltase-glucoamylase; CtMGAM, C-terminal luminal subunit; MD, molecular dynamics; DSK, de-O-sulfonated kotalanol; PCA, Principal Component Analysis; PCs, principal components; FEL, free energy landscape; ASMD, adaptive steered molecular dynamics simulations; PMF, potential of mean force; JA, Jarzynski average; RC, reaction coordinate.

INTRODUCTION

At present, there are multiple drugs for the treatment of type 2 diabetes, including traditional sulfonylureas (Stephen et al., 2018), biguanides (Schäfer, 1983), glinides (Chen et al., 2015), thiazolidinediones (Nanjan et al., 2018), α-glucosidase inhibitors (Kazufumi et al., 2014; Patel, 2015), glucagon-like peptide-1 (GLP-1) receptor agonists (Drucker, 2018), dipeptidyl peptidase IV (DPP-4) inhibitors, and sodium-glucose cotransporter 2 (SGLT2) inhibitors (Thornberry and Gallwitz, 2009; Kelly et al., 2019). These therapeutic drugs have been widely used in clinical trials because of their own characteristics in hypoglycemic control. For example, α-glucosidase inhibitors have been used to control postprandial glucose levels caused by type 2 diabetes since 1990 (Ríos et al., 2015; Flores-Bocanegra et al., 2017; Santos et al., 2018; Dhameja and Gupta, 2019; Usman et al., 2019; Mi et al., 2021; Tuyen et al., 2021). Acarbose (Chiasson et al., 2002) and miglitol (Satoru et al., 2015), which were clinically used for treating type 2 diabetes, may control blood glucose levels by targeting α-amylases and α-glucosidases (Lyann et al., 2010; Ren et al., 2011).

Glycoside hydrolases play significant roles in human metabolism, including digestion and decomposition of polysaccharides and biosynthesis of glycoprotein (Lovering et al., 2005). α-Glucosidases are rather crucial in the human metabolic system and are principally found in families 13 and 31 (Lovering et al., 2005). Maltase-glucoamylase (MGAM) (Sim et al., 2008) and sucrase-isomaltase (SI) (Sim et al., 2010) belong to glycoside hydrolase family 31. The main function of MGAM and SI is to digest terminal starch products left after the enzymatic action of α-amylase, which becomes an efficient drug target for insulin resistance (Van Beers et al., 1995). MGAM contains the following units: a small cytosolic domain of approximately 26 residues, a transmembrane domain (TMD) containing about 20 residues inserted into the intestinal epithelial cell membrane, an O-glycosylated linker, and two independent catalytic subunits: NtMGAM and C-terminal luminal subunit (CtMGAM) (Sim et al., 2008; Lyann et al., 2010) (**Supplementary Figure S1**). NtMGAM containing 864 residues (PDB ID: 3L4U) (Lyann et al., 2010) were used in this study. NtMGAM, known to be retaining α-glycosidases (Satoh et al., 2016), has received relatively little attention despite its importance and the number of different activities from a range of organisms, including mammals, plants, and microorganisms (Frandsen and Svensson, 1998; Yu et al., 1999; Lovering et al., 2005). The substrate specificities of MGAM vary and overlap from maltose (Quezada-Calvillo et al., 2008) to isomaltose (Elferink et al., 2020) and other small oligosaccharides.

Salacia reticulata (S. reticulata), a plant widely distributed in China and some other countries in Southeast Asia, is used as a traditional prescription for treating type 2 diabetes (Medagama, 2015). Sulfonium ion-containing compounds were isolated from aqueous extracts of *S. reticulata* by Jayakanthan and his colleagues (Jayakanthan et al., 2009), including de-O-sulfonated kotalanol (DSK). DSK comprises a 1,4-anhydro-4-

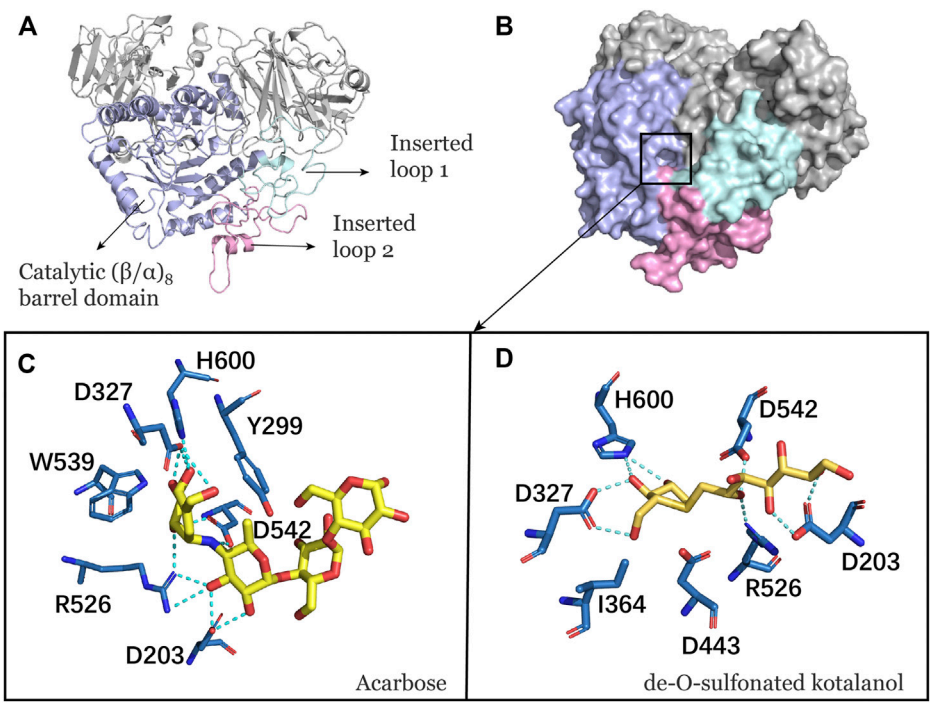


FIGURE 1 | (A) Catalytic domain of NtMGAM. The catalytic $(\beta/\alpha)_8$ barrel domain is tagged and colored differently. **(B)** Surface diagram of NtMGAM. **(C)** Binding of acarbose to NtMGAM. The active residues around acarbose binding to NtMGAM. **(D)** Binding of DSK to NtMGAM. The active residues around DSK binding to NtMGAM.

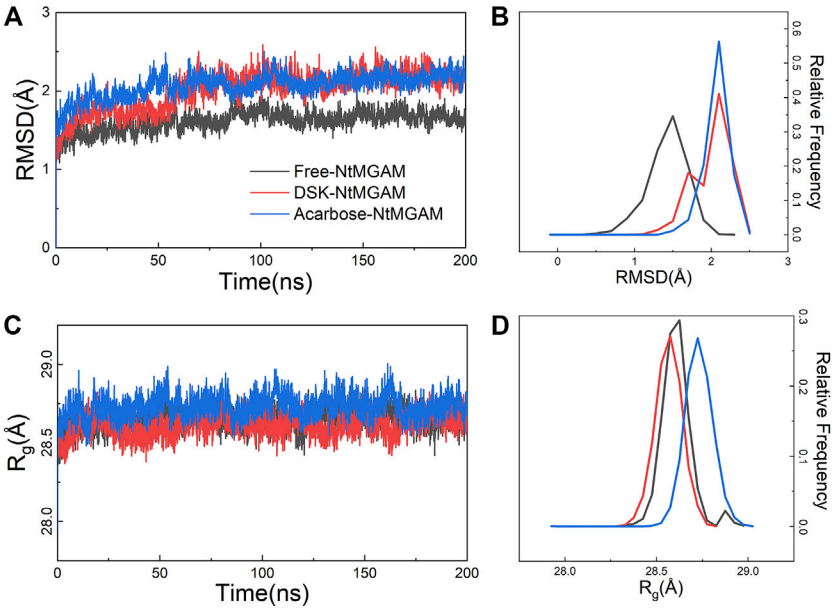


FIGURE 2 | Root-mean-square deviation (RMSD) and radius of gyration (R_g) analysis of NtMGAM with or without different ligands throughout 200 ns. **(A)** RMSD plot. **(B)** The relative frequency of the RMSD plot. **(C)** Gyration radius (R_g) plot. **(D)** The relative frequency of the R_g plot.

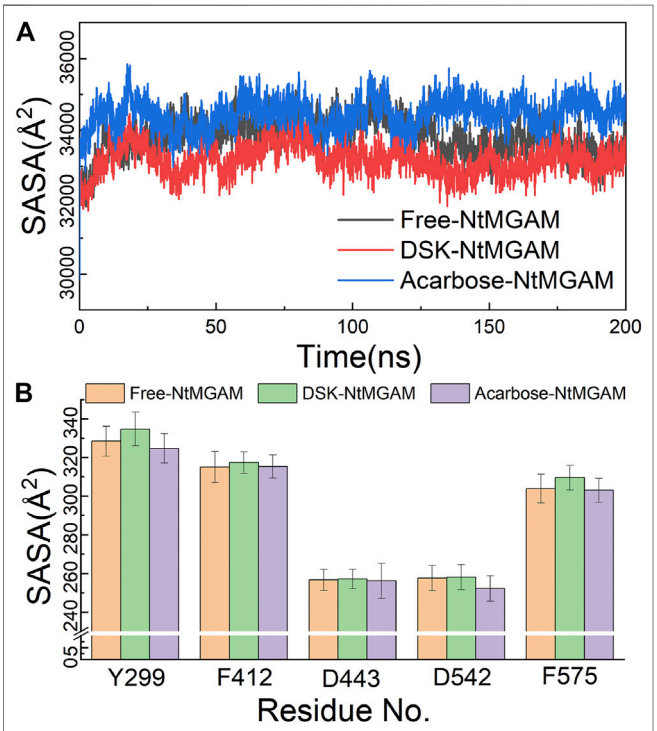


FIGURE 3 | (A) Solvent-accessible surface area (SASA) analysis of three systems over 200 ns MD. **(B)** SASA of active residues in different compound structures.

thio-d-arabinitol core and a polyhydroxylated acyclic chain (Lyann et al., 2010) and can act as an inhibitor of α-glucosidase in human bodies. A previous study reported that DSK could be an efficient inhibitor of NtMGAM with a K_i of 0.03 (±0.01) μM (Lyann et al., 2010).

In order to explore the conformational changes in the active pocket and unbinding pathway for NtMGAM, molecular dynamics (MD) simulations (Zhu et al., 2018; Zhu et al., 2019; Liu et al., 2020) and adaptive steered molecular dynamics (ASMD) simulations (Zhu et al., 2018) were performed between two inhibitors (DSK and acarbose) and NtMGAM (PDB ID: 3L4U) (Lyann et al., 2010) (workflow listed in **Supplementary Figure S2**). Our results may provide new ideas for the further design of α-glucosidase inhibitors.

MATERIALS AND METHODS

Preparation for the Structure of Protein Inhibitors

AutoDock 4.2 (Morris et al., 2009) was used for docking acarbose with NtMGAM using the Lamarckian genetic algorithm to identify a proper binding conformation with a grid box size of 66 Å × 58 Å × 66 Å points and a grid point spacing of 0.375 Å. The binding conformation with the lowest energy was chosen for simulations. The crystal structure of NtMGAM with DSK complex and the 3D structure of acarbose (PDB ID: 3JYR) (Vahedi-Faridi et al., 2010) was downloaded from Protein Data Bank (www.rcsb.org) for further studies.

MD Simulations

Simulations in our study were performed using the Amber16 package (D.A. Case et al., 2016) with the Amber ff99SB force

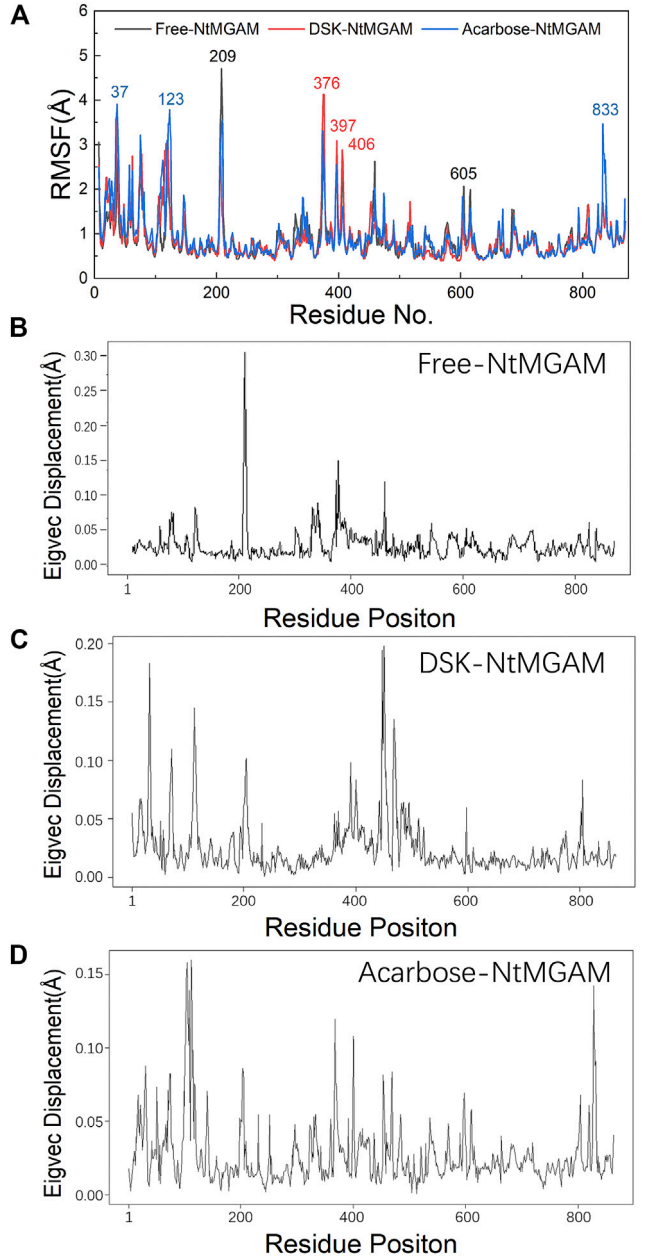


FIGURE 4 | (A) Average atom positional root-mean-square fluctuations (RMSF) of the backbone atoms per residue for the inhibitors bound NtMGAM. **(B–D)** The eigenvector components for atomic displacement along the first eigenvectors for MD-generated ensembles of free-NtMGAM, DSK-NtMGAM, and acarbose-NtMGAM, respectively.

field (Lindorff-Larsen et al., 2010). At the same time, the Gaff2 force field (Wang et al., 2004) was utilized to generate the parameterization of DSK and acarbose. It is well known that charged residues affect the environment of protein (Popović and Stuchebrukhov, 2004; Tashiro and Stuchebrukhov, 2005; Sugitani and Stuchebrukhov, 2009). H++ is an online tool that can automatically compute pKa values of ionizable groups in proteins (http://biophysics.cs.vt.edu/). We computed the ionizable groups of NtMGAM on H++ and then manually fixed the ionizable groups. All three complexes were analyzed using the MD simulations in a cubic periodic boundary box with the TIP3P water model (Bogunia and Makowski, 2020), which was prolonged to 12 Å from the protein atoms. Sodium ions were randomly added to the simulation systems for neutralization. To get the initial equilibrious structure, energy minimization was performed through the steepest descent method in 1,000 cycles. Subsequently, 50 ps of NVT (Berendsen temperature coupled with constant particle number, volume, and temperature) (Berendsen et al., 1984) and 50 ps of NPT (Parrinello–Rahman pressure coupled with constant particle number, pressure, and temperature) (Andersen, 1980) were performed to maintain the stability of the system (300 K, 1 bar). After stabilizing all thermodynamic properties, a 200 ns unconstrained MD simulation was performed with a time interval of 2 fs. The coordinates for all models were stored every 2 ps. During the simulation, the following options were specified: (I) bonds involving hydrogen are constrained and bond interactions involving H-atoms were omitted using the SHAKE algorithm (Miyamoto and Kollman, 1992), (II) the particle mesh Ewald summation algorithm (Essmann et al., 1995) was taken to calculate the long-range electrostatic interactions, (III) 1 atm constant pressure was maintained by the Langevin dynamics method (Rosenberg et al., 1986) (Guàrdia and Padró, 1985), and (IV) an optimum temperature (300 K) was maintained. MD simulations were performed three times for each system in this study (**Supplementary Figures S3, S4**). The root-mean-square deviation (RMSD), radius of gyration (R_g), root-mean-square fluctuation (RMSF), and solvent-accessible surface area (SASA) values were calculated using VMD (Humphrey et al., 1996).

MM-PBSA Calculations

MMPBSA.py (Miller et al., 2012) in the AmberTools17 package was employed to conduct free energy calculations for the two complexes. 200 conformations were extracted from each equilibrious trajectory (from 100 to 200 ns with an interval of 50 frames) for calculations. The binding free energies were calculated by subtracting the free energies of the receptor and the ligand DSK or acarbose from the free energy of the bound complex of two systems:

$$\Delta G_{binding,\ solvated} = \Delta G_{complex,\ solvated} - \left[\Delta G_{receptor,\ solvated} + \Delta G_{ligand,\ solvated}\right].$$

Then, the free energy change associated with each term of ΔG was calculated according to the following:

$$\Delta G_{solvated} = E_{gas} + \Delta G_{solvation} - TS_{solute},$$

where $\Delta G_{solvation}$ represents true free energy. To determine the relative stability, end-state method calculations were performed to estimate the energies, according to averages from the ensemble of these snapshots:

$$\Delta G_{solvated} \cong \langle E_{gas} \rangle + \langle \Delta G_{solvation} \rangle - T \langle S_{solute} \rangle = \frac{1}{N}\sum_{i=1}^{N} E_{i,\ gas} + \frac{1}{N}\sum_{i=1}^{N} \Delta G_{i,\ solvation} - \frac{T}{N}\sum_{i=1}^{N} S_{i,\ solute},$$

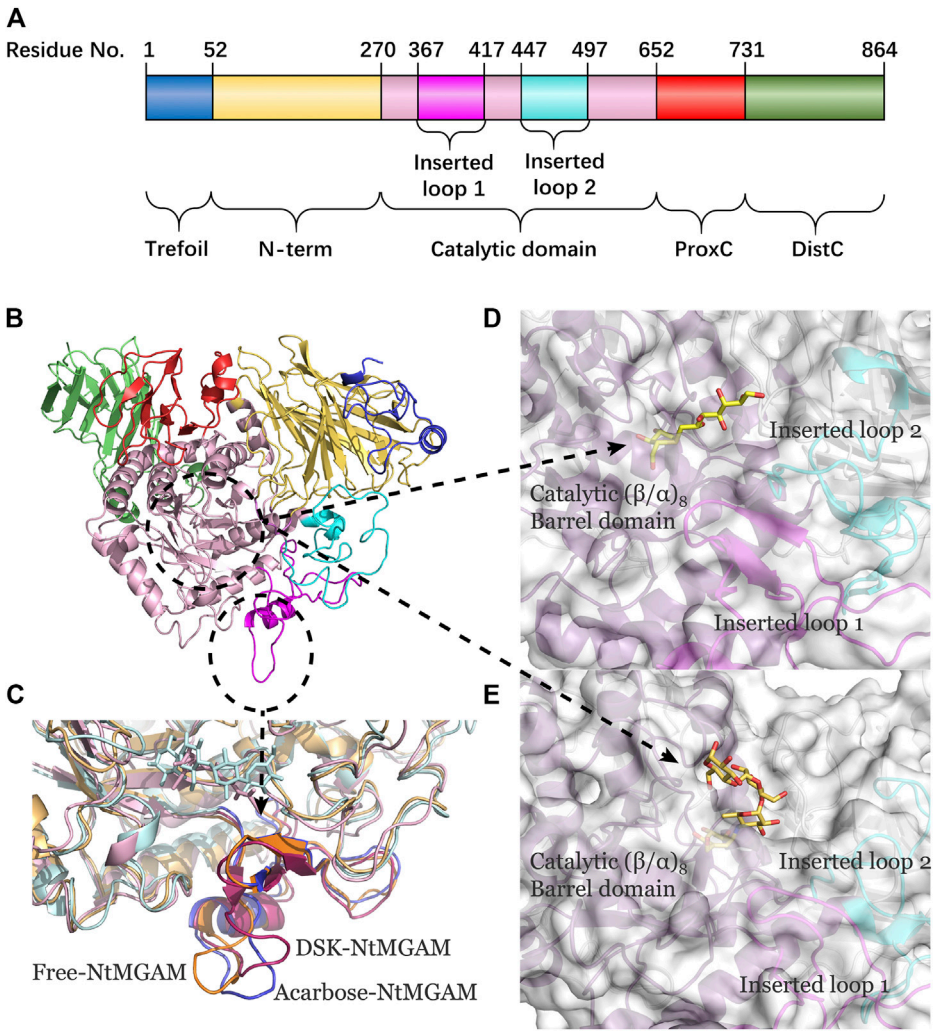


FIGURE 5 | (A) The NtMGAM domains with residue numbers labeled. **(B)** New cartoon diagram of NtMGAM with domains colored differently. **(C)** Inserted loop 1 in aligned compounds (the lowest energy conformations from FEL analysis). **(D,E)** NtMGAM active site pocket shown in surface representation occupied by DSK D and acarbose E. The structure of NtMGAM is shown in new cartoon and colored dissimilarly as follows: purple for catalytic domain, pink for inserted loop 1, cyan for inserted loop 2.

where *i* is the index of a particular frame and N is the total number of frames analyzed. The gas-phase energies (E_{gas}) can be computed from the quantum mechanical (QM) calculations and used further as a part of the force field parameterization; therefore, the E_{gas} energies can be abstracted from the molecular mechanical (MM) energies and the corresponding force field (Miller et al., 2012).

Principal Component Analysis and Free Energy Landscapes

PCA is a common statistical multivariate method, which can select the structure of each frame in an MD trajectory as a new set of variables, called principal components (PCs), with a minimal loss of information. In our study, we employed the bio3d package in R to refine structural superposition and examine the relationship between different conformers (Grant et al., 2006). The current protocol excludes the residues displaying the largest positional differences at each round and identifies only the core residues. Following the superposition of core residues, PCA was performed to examine the conformers based on their equivalent residues. The PCs collected during MD simulation are originally the eigenvector values collected from the covariance matrix, each corresponding to the change in protein trajectory (Al-Khafaji and Taskin Tok, 2020). In order to obtain a lower-dimensional representation of the structural dataset, we project the distribution onto the subspace defined by the largest principal components. In each dimension, the corresponding eigenvalue represents the percentage of the total mean square displacement (or variance) of atom positional fluctuations. In PCA, very few dimensions are generally enough to capture about 70% of the total variance in the structures to be studied (Grant et al., 2006). Therefore, the first few eigenvectors are sufficient to provide a useful description while still retaining most of the variance in

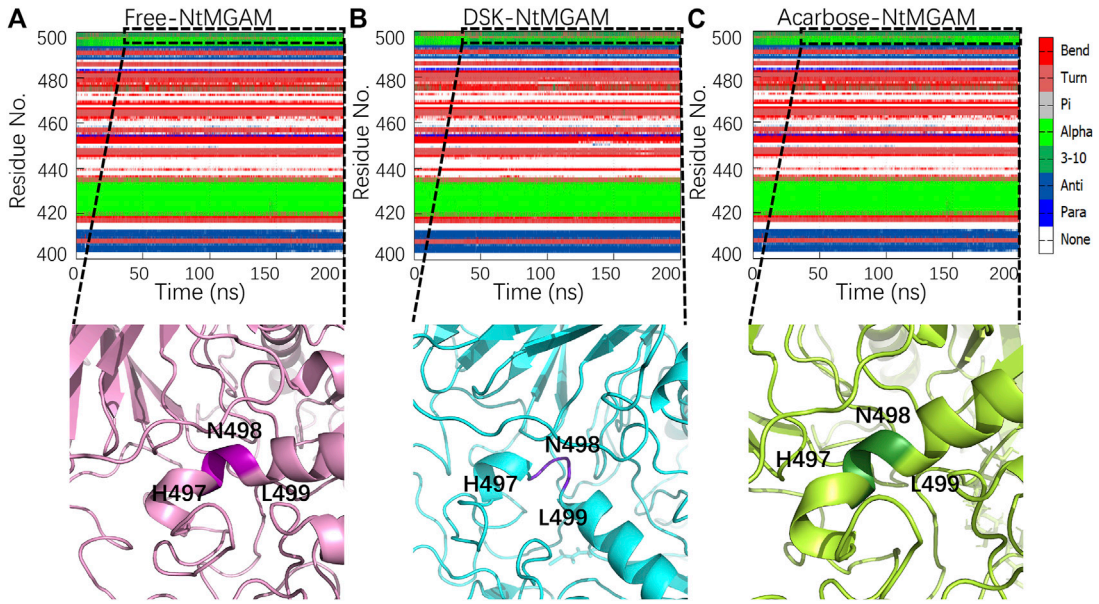


FIGURE 6 | Conformation changes in residues Asn491-Leu493 of the three systems: **(A)** free-NtMGAM, **(B)** DSK-NtMGAM, and **(C)** acarbose-NtMGAM.

TABLE 1 | The probability of secondary structures of residues H497 to L499.

Residue	Free-NtMGAM		DSK-NtMGAM		Acarbose-NtMGAM	
	α-Helix	Loop	α-Helix	Loop	α-Helix	Loop
H497	0.94	0.06	0.59	0.41	0.89	0.11
N498	0.94	0.06	0.59	0.41	0.92	0.08
L499	0.94	0.05	0.6	0.40	0.92	0.08

the original distribution. After PCA, the FEL (Frauenfelder et al., 1991) was obtained by calculating the joint probability distribution from the essential plane constructed from the top two eigenvectors (Singh et al., 2015).

Pathways Identified With CAVER

The software called CAVER has been widely used to analyze and visualize possible cavities and channels in protein structures. CAVER Analyst 2.0 (Jurcik et al., 2018) was employed to determine the channel position. The starting point was set at the position of the ligand for the channel computation. The minimum probe radius, clustering threshold, shell depth, and radius were set to be 0.9, 3.5, 4, and 3 Å, respectively. All the other parameters were used as default. 1,000 snapshots (we took one frame every ten frames in 10,000 from the 100–200 ns simulations) were analyzed utilizing CAVER Analyst 2.0.

Adaptive Steered Molecular Dynamics Simulations

ASMD simulations (Ozer et al., 2010; Ozer et al., 2012a; Ozer et al., 2012b; Ozer et al., 2014) were performed for NtMGAM with the two different ligands using the Amber 16 package. ASMD has been shown to alleviate the problem that many simulations must be run to converge the potential mean of force (PMF) in steered molecular dynamics (Izrailev et al., 1999) by dividing the predetermined reaction coordinate into numerous smaller stages. Each stage in the ASMD simulation contained multiple simulations that should be performed parallelly. In each stage, the trajectory with the work value closest to the Jarzynski average (JA) (Jarzynski, 1997) should be selected, and the coordinates at the end of that trajectory should be used as an initial coordinate for the next stage. Then, the JA structures were used for PMF calculation. Here, in our study, the distance between the ligand and NtMGAM in the initial conformation is 6 Å. The reaction coordinate in ASMD is predetermined to be set at 20 Å, at which the inhibitor is considered to escape from the enzyme. The stretching velocity was 10 Å/ns in this ASMD simulation, coupling a spring constant k of 40 kcal/(mol×Å^2). Each simulation was split every 2 Å into seven stages; each stage contains 14 simulations to reach the final reaction coordinate of 20 Å. As the distance between the two selected atoms reached 20 Å, there were no longer any interactions between the ligand and the receptor.

RESULTS AND DISCUSSION

Docking Pose and System Stabilize

It was reported that the structure of the NtMGAM substrate-binding site consisted of two sugar-binding sites, which in the acarbose-NtMGAM binding structure were occupied by the two nonreducing rings of acarbose (Lyann et al., 2010). To determine the docking pose, we chose to dock the DSK crystal structure to NtMGAM with AutoDock 4.2 (Morris et al., 2009). Comparing

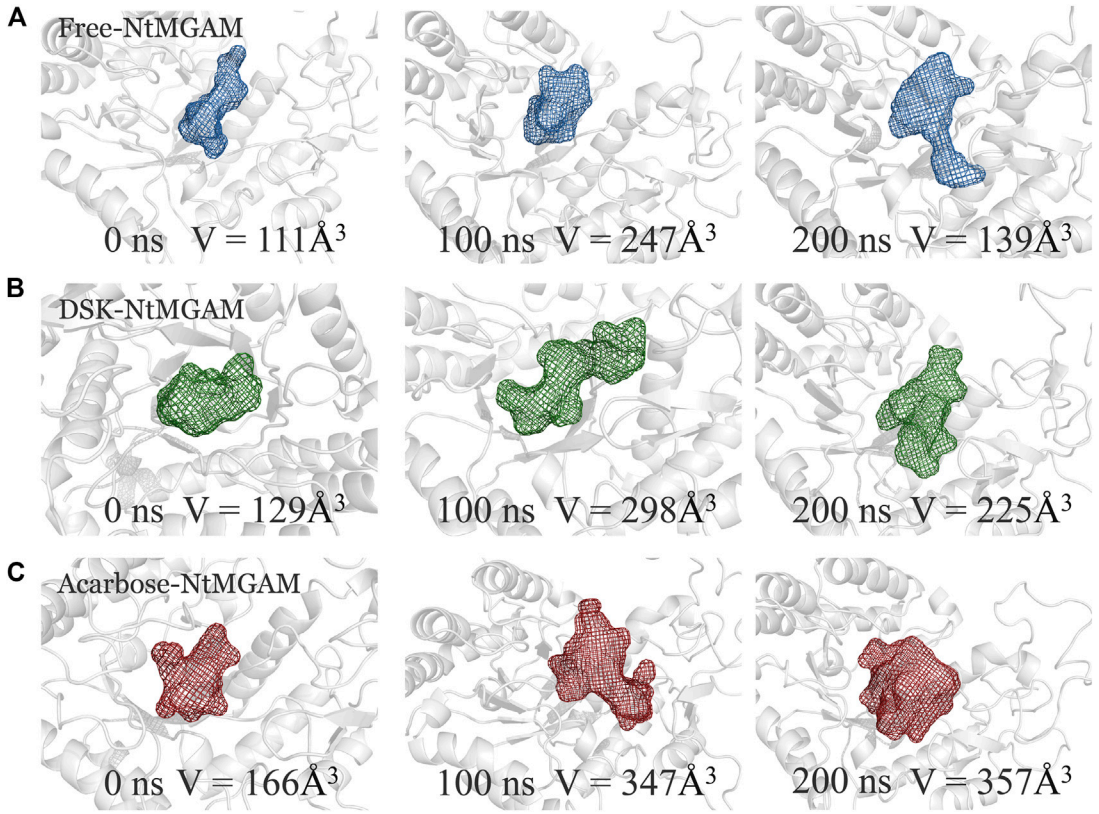


FIGURE 7 | The binding pocket conformations of free-NtMGAM **(A)**, DSK-NtMGAM **(B)**, and acarbose-NtMGAM **(C)** at 0, 100, and 200 ns with the volume of the pockets labeled.

the crystal structure of DSK-NtMGAM and the docked pose (**Supplementary Figure S5**), it can be concluded that they were similar (RMSD 0.50), which indicated that this system may use AutoDock 4.2 software to determine the binding pose. Acarbose was docked to NtMGAM using the same method.

NtMGAM has a typical catalytic $(\beta/\alpha)_8$ barrel domain (**Figure 1A**). DSK and acarbose were docked in an active pocket of NtMGAM (**Figure 1B**). **Figures 1C,D** shows that the active residues around acarbose and DSK were bound around NtMGAM. It can be seen that His600, Arg526, Asp542, Asp203, Trp406, and Asp327 were anchor residues for DSK binding in **Figure 1D**. In contrast, in the acarbose-NtMGAM, His594, Asp321, Arg520, Asp197, and Asp536 made hydrogen bonds with acarbose, indicating that they are important residues for acarbose binding to NtMGAM.

The MD simulations for three systems have been performed three times. The RMSD values of the Cα atom backbone of residues of three systems were calculated to evaluate the equilibrium of systems (**Supplementary Figures S3, S4** and **Supplementary Table S1**). It can be seen in **Supplementary Figure S3** that all MD simulations have got equilibrium. In **Supplementary Figure S4**, the RMSD of the three simulations of free-NtMGAM are slightly different from each other, with the average values of 1.67, 1.94, and 2.13 (from 100 to 200 ns), respectively. Although the RMSD values of each of the three repetitions seldom cross each other, the three simulations have all reached a state of relative equilibrium. The group with the most equilibrious and generally balanced RMSD was chosen for further study (**Supplementary Table S1**). The parameters of 200 ns MD simulations for three systems were listed in **Supplementary Table S2**. From **Figure 2A**, the RMSD values of NtMGAM, DSK-NtMGAM, and acarbose-NtMGAM are stabilized at about 1.67, 2.16, and 2.15 Å, respectively (**Figure 2B**), suggesting the structures of the three systems had reached a state of relative equilibrium. The R_g values of the three systems are shown in **Figures 2C,D** and are finally stabilized at 28.75 Å. In **Figure 3A**, the SASA values of acarbose-NtMGAM were stabilized at 34,000 $Å^2$; meanwhile, the other two systems were stabilized at 33,000 $Å^2$. These results indicate that three systems have attained stability and their states reached equilibrium. SASA values of single residue of the three systems were calculated for core residues. Tyr299, Phe412, Asn433, Asp542, and Phe575 had higher scores in the DSK-NtMGAM complex than the others. It was reported that Asp542 and Phe575 were important residues interacting with the hydroxyl groups of inhibitors binding to NtMGAM (Usman et al., 2019). Our results were consistent with the experimental data. In summary, all the systems were stabilized and can be used for further study.

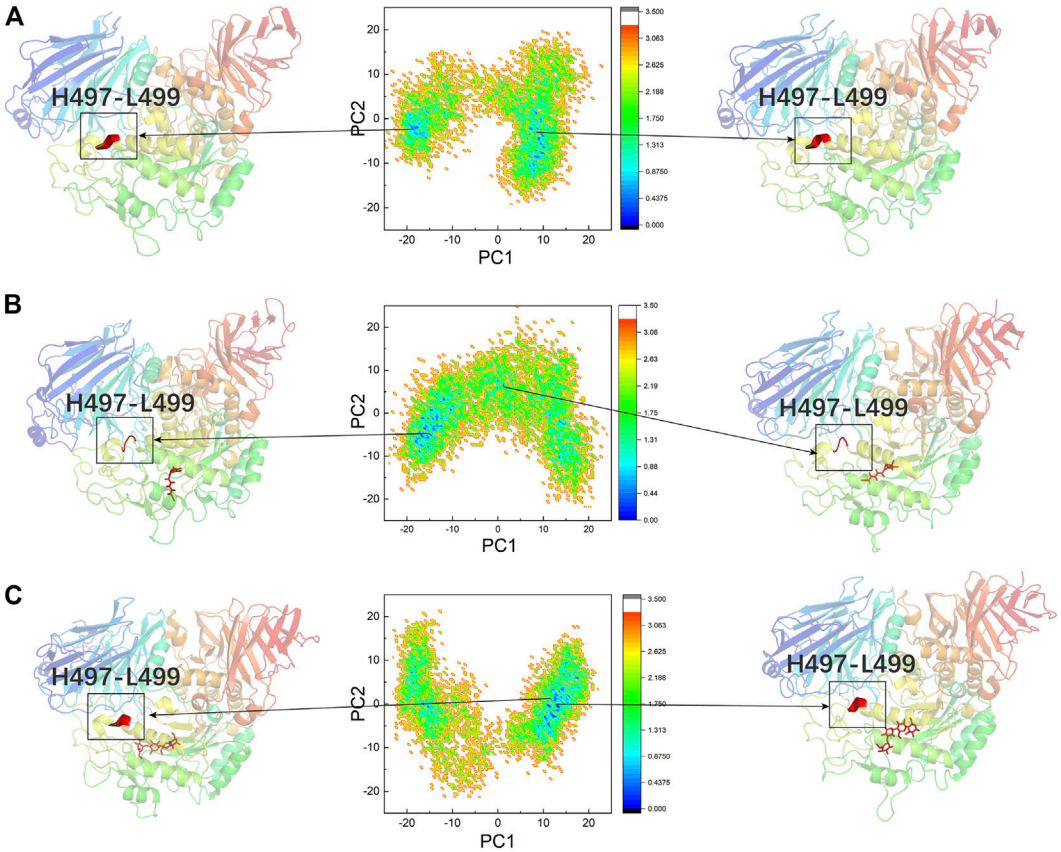


FIGURE 8 | PCA based FEL analysis of NtMGAM **(A)**. DSK-NtMGAM **(B)**. Acarbose-NtMGAM **(C)** as a function of projections of the MD trajectory onto the first (PC1) and second (PC2) eigenvectors. The structures of the two most stable conformations of the three systems are presented with the conformation of Asn491 to Leu493.

TABLE 2 | The probabilities of PC1 and PC2 of the three systems.

	PC 1 (%)	PC 2 (%)
Free-NtMGAM	16.46	6.28
DSK-NtMGAM	17.03	6.44
Acarbose-NtMGAM	19.61	5.46

TABLE 3 | MM-PBSA results.

	DSK	Acarbose
VDWAALS	−15.84 ± 0.39	−29.86 ± 1.01
EEL	−586.14 ± 2.32	−81.12 ± 4.95
EGB	565.10 ± 1.55	108.44 ± 4.87
ESURF	−5.02 ± 0.06	−6.22 ± 0.19
ΔG_{gas}	−601.98 ± 2.49	−110.99 ± 4.99
ΔG_{solv}	560.08 ± 1.50	102.21 ± 4.74
ΔTOTAL	−41.90 ± 1.14	−8.77 ± 1.08
K_i (μM)	0.03 ± 0.01	62 ± 13
ΔG_{exp}	−10.32	−5.76

Conformational Changes for Inhibitors Binding

To evaluate the deviation amount of displacement in three trajectories from MD simulations, atom positional RMSF values were calculated for the backbone atoms of three systems (**Figure 4A**). **Figures 4B–D** show that residue displacements correspond to the motions described by the first eigenvector for three complexes. These displacements represented the relative displacement of each residue caused by the motion described by a given eigenvector. It was reported that NtMGAM contains 868 residues, which can be divided into five structural domains: (I) a trefoil Type-P domain (residue No. 1–51); (II) N-terminal β-sandwich domain (residue No. 52–269); (III) catalytic $(\beta/\alpha)_8$ barrel domain (residue No. 270–651) with two inserted loops [inserted loop 1 (residue No. 367–416) and inserted loop 2 (residue No. 447–492)] protruding out between β3 and α3 and between β4 and α4, respectively; (IV) proximal C-terminal domain (residue No. 652–730); (V) distal C-terminal domain (residue No. 73–868), both with β-sandwich topologies (Sim et al., 2010) (see **Figure 5A**). In the DSK-NtMGAM complex, residues Ser376, Gly397, and Trp406, which are located at inserted loop 1 domain, exhibited distinct atom positional fluctuation amplitudes. This

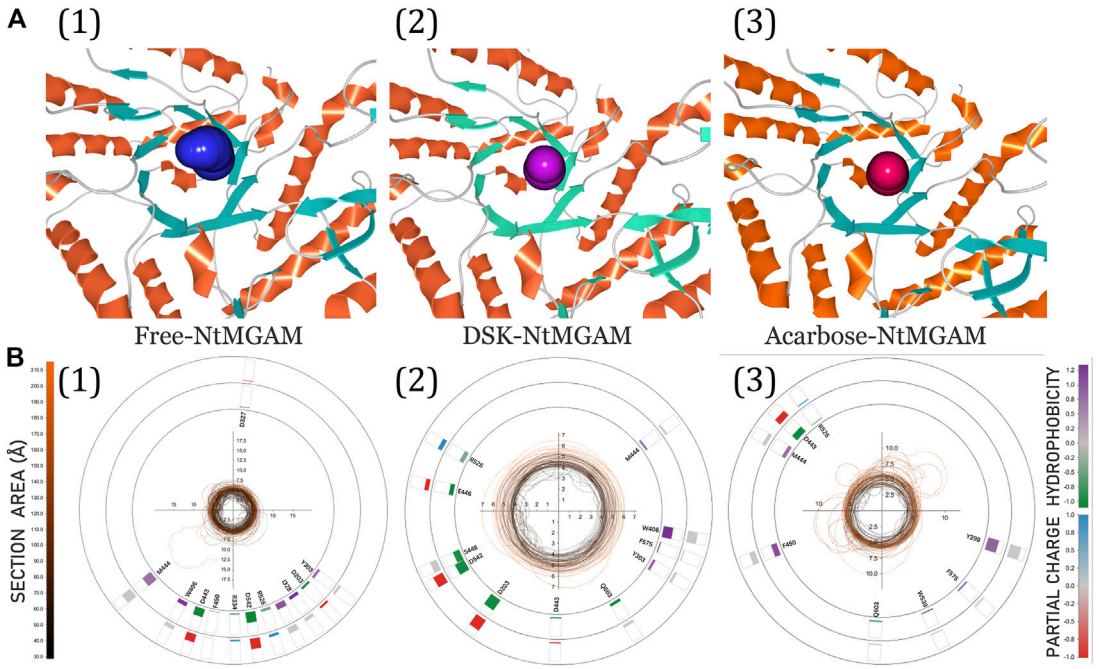


FIGURE 9 | (A) 3D visualization of candidate tunnels of (1) free-NtMGAM, (2) DSK-NtMGAM, and (3) acarbose-NtMGAM. **(B)** The details of the tunnel bottleneck contour over time of (1) free-NtMGAM, (2) DSK-NtMGAM, and (3) acarbose-NtMGAM.

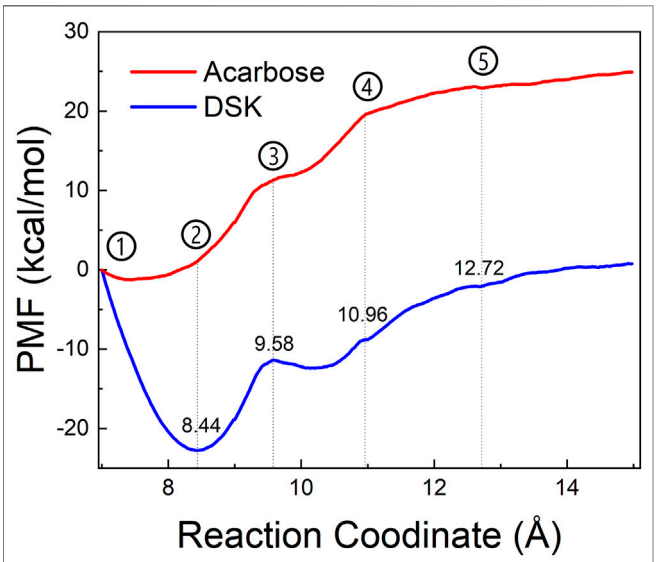


FIGURE 10 | PMF profiles along with the reaction coordinates of acarbose-NtMGAM (red) and DSK-NtMGAM (blue).

displacement widened the active site pocket, which would affect the inhibitors' binding.

Subsequently, secondary structure analysis was also performed, and the corresponding average secondary structure values for each residue are shown in **Figures 6A–C** and **Table 1**. It can be seen that the proportion of α-helix in Asn498 in free-NtMGAM and acarbose-NtMGAM was about 90%, whereas in DSK-NtMGAM, it was about 59% (**Table 1** and **Supplementary Table S3**). **Supplementary Table S3** shows that the proportion of α-helix in residue 497–499 in free-NtMGAM and acarbose-NtMGAM are almost above 90%. In contrast, in the DSK-NtMGAM complex, the odds are reduced to under 70%. The results revealed that the α-helix of DSK-NtMGAM during 200 ns MD simulations disappeared partly. Residues 497–499 are located near the inserted loop 2 domain, which is quite close to the opening of the $(\beta/\alpha)_8$ barrel. Despiralization of these residues can enlarge the domain of the inserted loop 2, therefore, contributing to the architecture of the inhibitor binding site.

POCASA 44 (http://altair.sci.hokudai.ac.jp/g6/service/pocasa/) (Yu et al., 2010) was utilized to predict the volume of the binding pocket. Parameters are listed as follows: the radius of probe sphere value was 1 Å, single point flag value was 10 Å, and protein depth flag value was 15Å. The active pocket conformations at 0, 100, and 200 ns were shown in **Figures 7A–C**. The volume of the pocket in free-NtMGAM was smaller than that of DSK-NtMGAM and acarbose-NtMGAM. Obviously, it could be considered that inhibitor of NtMGAM binding to the pocket with the nonreducing sugar ring in the −1 subsite and the reducing ring in the +1 subsite results in net retention of configuration at the anomeric center (Sim et al., 2010). Large active pockets will facilitate the inhibitor binding and entry. However, acarbose is so large that it binds to the NtMGAM active site primarily with its acarvosine unit (−1 and +1 subsite), few with its glycone rings (+2 and +3 subsite) (Sim et al., 2008).

In addition, the S group of DSK can generate charge interaction with residues (Trp400, Asp437, and Asp536)

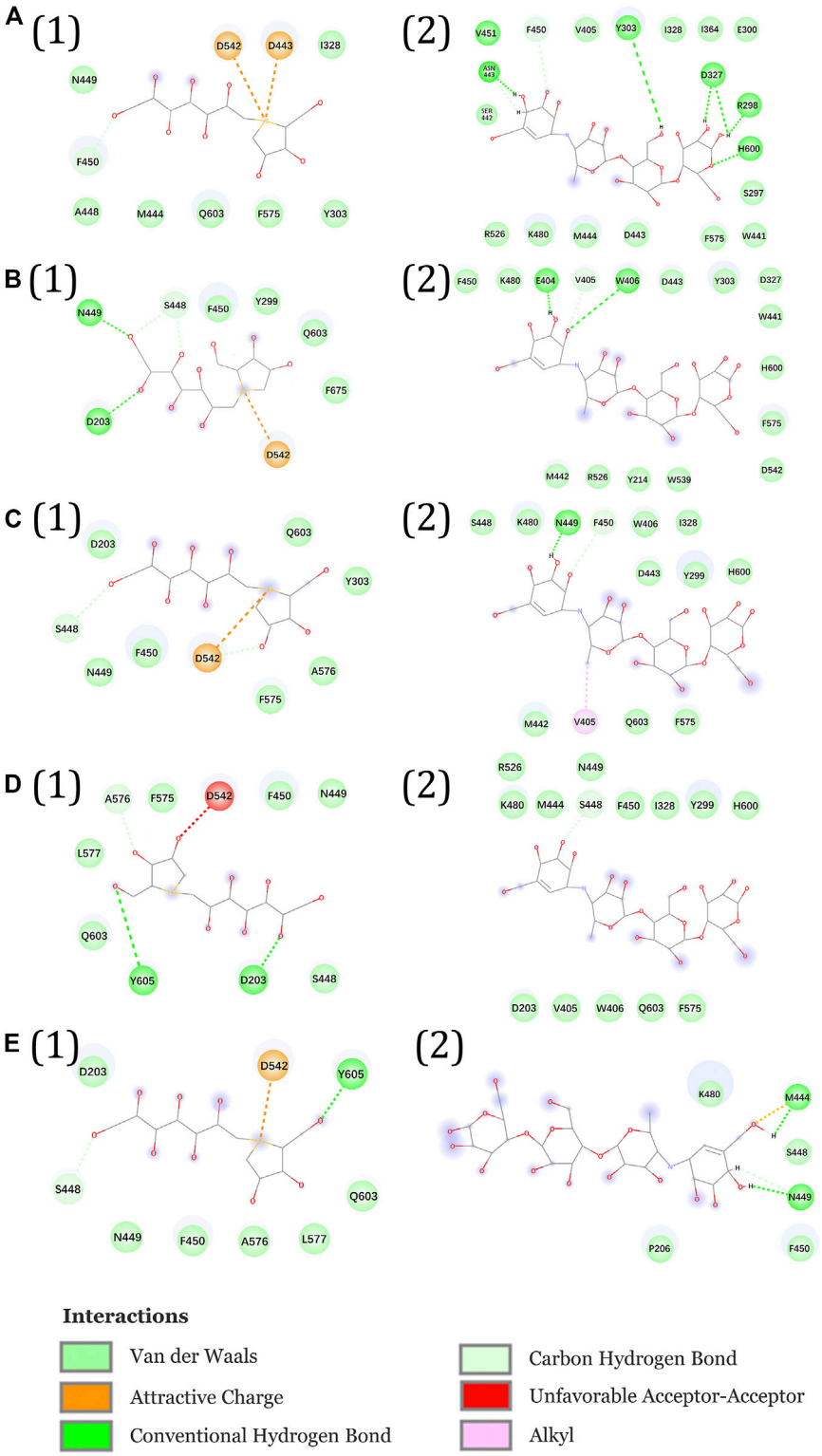


FIGURE 11 | Interactions during ASMD simulations between DSK (1) or acarbose (2) with NtMGAM. **(A)** The initial state of the two systems before the ASMD simulations. Reaction coordinates reached **(B)** 8.44 Å, **(C)** 9.58 Å, **(D)** 10.96 Å, and **(E)** 12.72 Å.

(**Supplementary Figure S5**), which may stabilize the DSK-NtMGAM complex.

PCA and FEL Analysis

PCA was performed to confirm whether the conformational changes of the three systems were continuous and stable, and the results are displayed through FEL (**Figures 8A–C**). **Table 2** lists the PC1 and PC2 probabilities of the three systems. The structures of the most stable conformations of the three systems revealed that the conformational changes in the residues 497–499 domain were complex in DSK-NtMGAM (α-helix disappeared partly). In summary, we can confirm that the conformational changes in the three systems were continuous and stable, and our analyses above are reliable.

MM-PBSA Calculations

We used the end-point method to calculate the free energy between NtMGAM and the two inhibitors. The results of the Generalized Born (GB) implicit solvent method with a SASA term calculation are shown in **Table 3**. The binding free energies were mainly contributed by electrostatic energy, which is calculated by the molecular mechanics force field and the electrostatic contribution to the solvation free energy calculated by GB. Meanwhile, the VDW interactions are approximately consistent. As shown in **Table 3**, the binding free energy of the DSK-NtMGAM complex (−41.90 ± 1.14 kcal/mol) is lower than that of the acarbose-NtMGAM complex (−8.77 ± 1.08 kcal/mol). The results have the same trend as the data calculated from the K_i value from Sim's works (Sim et al., 2008).

ASMD Simulations

The 3D visualization of the channel of the free-NtMGAM, DSK-NtMGAM, and acarbose-NtMGAM obtained by CAVER Analyst 2.0 is listed in **Figure 9A**. The detailed exploration of the channel bottleneck and surrounding residues is shown in **Figure 9B**. The bottleneck of the channel in DSK-NtMGAM was larger than that of acarbose-NtMGAM. The surrounding residues displayed around the contour demonstrated the frequency. It can be seen that there are more residues in the DSK-NtMGAM complex comparing to that in the acarbose-NtMGAM complex.

To explore the enzyme-inhibitor interactions and the affinity of the active sites of NtMGAM *via* inhibitors unbinding pathway, ASMD simulations were performed on the two complexes. In **Figure 10**, the PMF profile displayed the energy changes during the process of pulling the ligands out of the NtMGAM channel. In the DSK-NtMGAM complex, a lower energy barrier (13 kcal/mol) should be transferred to completely dissociate DSK from the channel of NtMGAM than that in the acarbose-NtMGAM complex (22 kcal/mol).

Figures 11A1–E1 show the interactional changes during the dissociation process of DSK-NtMGAM along the reaction coordinate (RC). First, for the initial coordinate (RC = 7 Å) [**Figure 11A1**], Asp542 and Asp443 formed a salt bridge with DSK. Then, at 8.44 Å [**Figure 11B1**], the free energy value dropped sharply due to the break of stronger hydrogen bonds between DSK and Asp203 and Asn449 residues. With the movement of DSK, the salt bridge between Asp443 and DSK disappeared and the salt bridge between Asp542 and DSK persisted. Nevertheless, at 9.58 Å [**Figure 11C1**], the residues that formed hydrogen bonds with the DSK were changed. Except for the salt bridge between Asp542 and DSK, the hydrogen bonds (salt bridge) were disappeared. Thereafter, the interconnections, including hydrogen bonds between channel residues and DSK, increased rapidly after 10.96 Å [**Figure 11D1**], giving rise to an increase in free energy value. Finally, at 12.72Å [**Figure 11D1**], DSK completely departed from the channel of NtMGAM and the curve of PMF tended to be flat. Asp542, as an important channel residue, could generate hydrogen bonds with DSK continuously, which was consistent with the results obtained in the channel analysis.

The acarbose dissociating from NtMGAM is shown in **Figures 11A2–E2**. At the beginning of the ASMD simulation [**Figure 11A2**], acarbose was tightly fixed due to the strong hydrogen bond interactions with Tyr303, Asp327, Arg298, and His600. Subsequently, at 8.44 Å [**Figure 11B2**], the free energy value increased slowly because of the stronger hydrogen bond interactions between acarbose and Glu404 and Trp406. At 9.58 Å [**Figure 11C2**], the acarbose made hydrogen bond interactions with Asn449, Val405, and Phe450, which were stronger than DSK, resulting in the increased free energy value (**Figure 11C**). Thereafter, the interconnections, including the hydrogen bonds between acarbose and channel residues, disappeared at about 10.96 Å besides the only hydrogen bond with Ser448 [**Figure 11D2**]. Finally, at 12.72 Å [**Figure 11E2**], acarbose was completely departed from the unbinding pathways of NtMGAM, and the curves of PMF tended to be flat.

To sum up, compared to acarbose, DSK escaped from NtMGAM easily with lower energy. Asp542 is an important residue on the bottleneck of the active pocket of NtMGAM, which could generate hydrogen bonds with DSK continuously. Our results may provide some useful clues for designing new medicine to relieve symptoms of postprandial hyperglycemia caused by type 2 diabetes. For example, we can modify the 3D structure of acarbose to get a new compound that is suitable for the active pocket of NtMGAM.

CONCLUSION

At present, there are multiple drugs for the treatment of type 2 diabetes on the market, including α-glucosidase inhibitors. MGAM has become an efficient drug target for insulin resistance. In order to explore the conformational changes in the active pocket and unbinding pathway for NtMGAM, MD simulations and ASMD simulations were performed between two inhibitors (DSK and acarbose) and NtMGAM. MD simulations indicated that DSK binding to NtMGAM may lead to an enlargement of the active pocket due to the flexibility of the two domains (inserted loop 1 and inserted loop 2), which would facilitate the binding of DSK to NtMGAM. ASMD simulation results indicated that Asp542 was an important channel residue, which could continuously generate hydrogen bonds with DSK. Our results may provide some interesting thoughts for designing new medicine for the treatment of type 2 diabetes based on the molecular structure

and specific intermolecular interactions between NtMGAM and DSK substrate in the binding pocket and the entrance channel.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

SZ wrote this paper and prepared the tables and figures. LH revised this paper. YW was responsible for the Supplementary Materials. XF and SW provided some revision advice. WL and WH provided the ideas and modified the papers.

FUNDING

This work was supported by the Overseas Cooperation Project of Jilin Province (20200801069GH).

ACKNOWLEDGMENTS

Amber 16 package was purchased and authorized to use by Jilin University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.711242/full#supplementary-material

REFERENCES

Al-Khafaji, K., and Taskin Tok, T. (2020). Molecular Dynamics Simulation, Free Energy Landscape and Binding Free Energy Computations in Exploration the Anti-invasive Activity of Amygdalin against Metastasis. *Comput. Methods Programs Biomed.* 195, 105660. doi:10.1016/j.cmpb.2020.105660

Andersen, H. C. (1980). Molecular Dynamics Simulations at Constant Pressure And/or Temperature. *J. Chem. Phys.* 72, 2384–2393. doi:10.1063/1.439486

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Dinola, A., and Haak, J. R. (1984). Molecular Dynamics with Coupling to an External bath. *J. Chem. Phys.* 81, 3684–3690. doi:10.1063/1.448118

Bogunia, M., and Makowski, M. (2020). Influence of Ionic Strength on Hydrophobic Interactions in Water: Dependence on Solute Size and Shape. *J. Phys. Chem. B* 124 (46), 10326–10336. doi:10.1021/acs.jpcb.0c06399

Case, D. A., Betz, R. M., Cerutti, D. S., Cheatham, T. E., III, Darden, T. A., Duke, R. E., et al. (2016). *Amber16*. San Francisco, CA: University of California, San Francisco.

Chen, M., Hu, C., and Jia, W. (2015). Pharmacogenomics of Glinides. *Pharmacogenomics* 16 (1), 45–60. doi:10.2217/pgs.14.152

Chiasson, J.-L., Josse, R. G., Gomis, R., Hanefeld, M., Karasik, A., and Laakso, M. (2002). Acarbose for Prevention of Type 2 Diabetes Mellitus: the STOP-NIDDM Randomised Trial. *The Lancet* 359 (9323), 2072–2077. doi:10.1016/s0140-6736(02)08905-5

Dhameja, M., and Gupta, P. (2019). Synthetic Heterocyclic Candidates as Promising α-Glucosidase Inhibitors: An Overview. *Eur. J. Med. Chem.* 176, 343–377. doi:10.1016/j.ejmech.2019.04.025

Drucker, D. J. (2018). Mechanisms of Action and Therapeutic Application of Glucagon-like Peptide-1. *Cel Metab.* 27 (4), 740–756. doi:10.1016/j.cmet.2018.03.001

Elferink, H., Bruekers, J. P. J., Veeneman, G. H., and Boltje, T. J. (2020). A Comprehensive Overview of Substrate Specificity of Glycoside Hydrolases and Transporters in the Small Intestine. *Cell. Mol. Life Sci.* 77 (23), 4799–4826. doi:10.1007/s00018-020-03564-1

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 103, 8577–8593. doi:10.1063/1.470117

Flores-Bocanegra, L., González-Andrade, M., Bye, R., Linares, E., and Mata, R. (2017). α-Glucosidase Inhibitors from Salvia Circinata. *J. Nat. Prod.* 80 (5), 1584–1593. doi:10.1021/acs.jnatprod.7b00155

Frandsen, T. P., and Svensson, B. (1998). Plant Alpha-Glucosidases of the Glycoside Hydrolase Family 31. Molecular Properties, Substrate Specificity, Reaction Mechanism, and Comparison with Family Members of Different Origin. *Plant Mol. Biol.* 37 (1), 1–13. doi:10.1023/a:1005925819741

Frauenfelder, H., Sligar, S., and Wolynes, P. (1991). The Energy Landscapes and Motions of Proteins. *Science* 254 (5038), 1598–1603. doi:10.1126/science.1749933

Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006). Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* 22, 2695–2696. doi:10.1093/bioinformatics/btl461

Guàrdia, E., and Padró, J. A. (1985). Generalized Langevin Dynamics Simulation of Interacting Particles. *J. Chem. Phys.* 83, 1917–1920. doi:10.1063/1.449379

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual Molecular Dynamics. *J. Mol. Graphics* 14 (1), 33-38–27-38. doi:10.1016/0263-7855(96)00018-5

Izrailev, S., Stepaniants, S., Isralewitz, B., Kosztin, D., Lu, H., Molnár, F., et al. (1999). "Steered Molecular Dynamics," in *Computational Molecular Dynamics*. Berlin: Springer.

Jarzynski, C. (1997). Equilibrium Free-Energy Differences from Nonequilibrium Measurements: A Master-Equation Approach. *Phys. Rev. E* 56, 5018–5035. doi:10.1103/physreve.56.5018

Jayakanthan, K., Mohan, S., and Pinto, B. M. (2009). Structure Proof and Synthesis of Kotalanol and De-O-sulfonated Kotalanol, Glycosidase Inhibitors Isolated from an Herbal Remedy for the Treatment of Type-2 Diabetes. *J. Am. Chem. Soc.* 131 (15), 5621–5626. doi:10.1021/ja900867q

Jurcik, A., Bednar, D., Byska, J., Marques, S. M., Furmanova, K., Daniel, L., et al. (2018). CAVER Analyst 2.0: Analysis and Visualization of Channels and Tunnels in Protein Structures and Molecular Dynamics Trajectories. *Bioinformatics* 34 (20), 3586–3588. doi:10.1093/bioinformatics/bty386

Kazufumi, N., Hiroki, O., Hajime, K., Kenei, S., Shota, F., Kyoko, W., et al. (2014). DPP-4 Inhibitor and Alpha-Glucosidase Inhibitor Equally Improve Endothelial Function in Patients with Type 2 Diabetes: EDGE Study. *Cardiovasc. diabetology* 13, 110. doi:10.1186/s12933-014-0110-2

Kelly, M. S., Lewis, J., Huntsberry, A. M., Dea, L., and Portillo, I. (2019). Efficacy and Renal Outcomes of SGLT2 Inhibitors in Patients with Type 2 Diabetes and Chronic Kidney Disease. *Postgrad. Med.* 131 (1), 31–42. doi:10.1080/00325481.2019.1549459

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* 78 (8), 1950–1958. doi:10.1002/prot.22711

Liu, Y., Zhu, J., Guo, X., Huang, T., Han, J., Gao, J., et al. (2020). How Oncogenic Mutations Activate Human MAP Kinase 1 (MEK1): A Molecular Dynamics Simulation Study. *J. Biomol. Struct. Dyn.* 38 (13), 3942–3958. doi:10.1080/07391102.2019.1686065

Lovering, A. L., Lee, S. S., Kim, Y.-W., Withers, S. G., and Strynadka, N. C. J. (2005). Mechanistic and Structural Analysis of a Family 31 α-Glycosidase and its Glycosyl-Enzyme Intermediate. *J. Biol. Chem.* 280 (3), 2105–2115. doi:10.1074/jbc.M410468200

Lyann, S., Kumarasamy, J., Sankar, M., Ravindranath, N., D, J. B., Mario, P. B., et al. (2010). New Glucosidase Inhibitors from an Ayurvedic Herbal Treatment for Type 2 Diabetes: Structures and Inhibition of Human Intestinal Maltase-Glucoamylase with Compounds from Salacia Reticulata. *Biochemistry* 49 (3), 443–451. doi:10.1021/bi9016457

Medagama, A. B. (2015). Salacia Reticulata (Kothala Himbutu) Revisited; a Missed Opportunity to Treat Diabetes and Obesity? *Nutr. J.* 14, 21. doi:10.1186/s12937-015-0013-4

Mi, C.-N., Yuan, J.-Z., Zhu, M.-M., Yang, L., Wei, Y.-M., Wang, H., et al. (2021). 2-(2-Phenylethyl)chromone Derivatives: Promising α-glucosidase Inhibitors in Agarwood from Aquilaria Filaria. *Phytochemistry* 181, 112578. doi:10.1016/j.phytochem.2020.112578

Miller, B. R., 3rd, McGee, T. D., Jr., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012). MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theor. Comput.* 8 (9), 3314–3321. doi:10.1021/ct300418h

Miyamoto, S., and Kollman, P. (1992). Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* 13, 952–962. doi:10.1002/jcc.540130805

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30 (16), 2785–2791. doi:10.1002/jcc.21256

Nanjan, M. J., Mohammed, M., Kumar, B. R. P., and Chandrasekar, M. J. N. (2018). Thiazolidinediones as Antidiabetic Agents: A Critical Review. *Bioorg. Chem.* 77, 548–567. doi:10.1016/j.bioorg.2018.02.009

Ozer, G., Keyes, T., Quirk, S., and Hernandez, R. (2014). Multiple Branched Adaptive Steered Molecular Dynamics. *J. Chem. Phys.* 141 (6), 064101. doi:10.1063/1.4891807

Ozer, G., Quirk, S., and Hernandez, R. (2012a). Adaptive Steered Molecular Dynamics: Validation of the Selection Criterion and Benchmarking Energetics in Vacuum. *J. Chem. Phys.* 136 (21), 215104. doi:10.1063/1.4725183

Ozer, G., Quirk, S., and Hernandez, R. (2012b). Thermodynamics of Decaalanine Stretching in Water Obtained by Adaptive Steered Molecular Dynamics Simulations. *J. Chem. Theor. Comput.* 8, 4837–4844. doi:10.1021/ct300709u

Ozer, G., Valeev, E. F., Quirk, S., and Hernandez, R. (2010). Adaptive Steered Molecular Dynamics of the Long-Distance Unfolding of Neuropeptide Y. *J. Chem. Theor. Comput.* 6, 3026–3038. doi:10.1021/ct100320g

Patel, S. (2015). Cerebrovascular Complications of Diabetes: Alpha Glucosidase Inhibitor as Potential Therapy. *Horm. Metab. Res.* 48 (2), 83–91. doi:10.1055/s-0035-1565181

Popović, D. M., and Stuchebrukhov, A. A. (2004). Electrostatic Study of the Proton Pumping Mechanism in Bovine Heart Cytochrome C Oxidase. *J. Am. Chem. Soc.* 126 (6), 1858–1871. doi:10.1021/ja038267w

Quezada-Calvillo, R., Sim, L., Ao, Z., Hamaker, B. R., Quaroni, A., Brayer, G. D., et al. (2008). Luminal Starch Substrate "brake" on Maltase-Glucoamylase Activity Is Located within the Glucoamylase Subunit. *J. Nutr.* 138 (4), 685–692. doi:10.1093/jn/138.4.685

Ren, L., Qin, X., Cao, X., Wang, L., Bai, F., Bai, G., et al. (2011). Structural Insight into Substrate Specificity of Human Intestinal Maltase-Glucoamylase. *Protein & Cell* 2 (10), 827–836. doi:10.1007/s13238-011-1105-3

Ríos, J., Francini, F., and Schinella, G. (2015). Natural Products for the Treatment of Type 2 Diabetes Mellitus. *Planta Med.* 81 (12-13), 975–994. doi:10.1055/s-0035-1546131

Rosenberg, R., Boughaleb, Y., Nitzan, A., and Ratner, M. (1986). Effective Potentials from Langevin Dynamic Simulations of Framework Solid Electrolytes. *Solid State Ionics* 18-19, 127–135. doi:10.1016/0167-2738(86)90099-8

Santos, C. M. M., Freitas, M., and Fernandes, E. (2018). A Comprehensive Review on Xanthone Derivatives as α-glucosidase Inhibitors. *Eur. J. Med. Chem.* 157, 1460–1479. doi:10.1016/j.ejmech.2018.07.073

Satoh, T., Toshimori, T., Yan, G., Yamaguchi, T., and Kato, K. (2016). Structural Basis for Two-step Glucose Trimming by Glucosidase II Involved in ER Glycoprotein Quality Control. *Scientific Rep.* 6, 20575. doi:10.1038/srep20575

Satoru, S., Hisakazu, N., Kitaro, K., and Hajime, H. (2015). Review: Miglitol Has Potential as a Therapeutic Drug against Obesity. *Nutr. Metab.* 12, 51. doi:10.1186/s12986-015-0048-8

Schäfer, G. (1983). Biguanides. A Review of History, Pharmacodynamics and Therapy. *Diabete & metabolisme* 9 (2), 148–163.

Sim, L., Quezada-Calvillo, R., Sterchi, E. E., Nichols, B. L., and Rose, D. R. (2008). Human Intestinal Maltase-Glucoamylase: Crystal Structure of the N-Terminal Catalytic Subunit and Basis of Inhibition and Substrate Specificity. *J. Mol. Biol.* 375 (3), 782–792. doi:10.1016/j.jmb.2007.10.069

Sim, L., Willemsma, C., Mohan, S., Naim, H. Y., Pinto, B. M., and Rose, D. R. (2010). Structural Basis for Substrate Selectivity in Human Maltase-Glucoamylase and Sucrase-Isomaltase N-Terminal Domains. *J. Biol. Chem.* 285 (23), 17763–17770. doi:10.1074/jbc.M109.078980

Singh, B., Bulusu, G., and Mitra, A. (2015). Understanding the Thermostability and Activity of Bacillus subtilisLipase Mutants: Insights from Molecular Dynamics Simulations. *J. Phys. Chem. B* 119 (2), 392–409. doi:10.1021/jp5079554

Stephen, C., David, M., Leiter, A. L., Siew, P. C., Giorgio, S., and Michel, M. (2018). The Place of Gliclazide MR in the Evolving Type 2 Diabetes Landscape: A Comparison with Other Sulfonylureas and Newer Oral Antihyperglycemic Agents. *Diabetes Res. Clin. Pract.* 143, 1–14. doi:10.1016/j.diabres.2018.05.028

Sugitani, R., and Stuchebrukhov, A. A. (2009). Molecular Dynamics Simulation of Water in Cytochrome C Oxidase Reveals Two Water Exit Pathways and the Mechanism of Transport. *Biochim. Biophys. Acta (Bba) - Bioenerg.* 1787 (9), 1140–1150. doi:10.1016/j.bbabio.2009.04.004

Tashiro, M., and Stuchebrukhov, A. A. (2005). Thermodynamic Properties of Internal Water Molecules in the Hydrophobic Cavity Around the Catalytic Center of CytochromecOxidase. *J. Phys. Chem. B.* 109 (2), 1015–1022. doi:10.1021/jp0462456

Thornberry, N. A., and Gallwitz, B. (2009). Mechanism of Action of Inhibitors of Dipeptidyl-Peptidase-4 (DPP-4). *Best Pract. Res. Clin. Endocrinol. Metab.* 23 (4), 479–486. doi:10.1016/j.beem.2009.03.004

Tuyen, D. T., Yew, G. Y., Cuong, N. T., Hoang, L. T., Yen, H. T., Hong Thao, P. T., et al. (2021). Selection, Purification, and Evaluation of Acarbose–an α-glucosidase Inhibitor from Actinoplanes Sp. *Chemosphere* 265, 129167. doi:10.1016/j.chemosphere.2020.129167

Usman, B., Sharma, N., Satija, S., Mehta, M., Vyas, M., Khatik, G. L., et al. (2019). Recent Developments in Alpha-Glucosidase Inhibitors for Management of Type-2 Diabetes: An Update. *Curr. Pharm. Des.* 25 (23), 2510–2525. doi:10.2174/1381612825666190717104547

Vahedi-Faridi, A., Licht, A., Bulut, H., Scheffel, F., Keller, S., Wehmeier, U. F., et al. (2010). Crystal Structures of the Solute Receptor GacH of Streptomyces Glaucescens in Complex with Acarbose and an Acarbose Homolog: Comparison with the Acarbose-Loaded Maltose-Binding Protein of *Salmonella typhimurium*. *J. Mol. Biol.* 397 (3), 709–723. doi:10.1016/j.jmb.2010.01.054

Van Beers, E. H., Büller, H. A., Grand, R. J., Einerhand, A. W. C., and Dekker, J. (1995). Intestinal brush Border Glycohydrolases: Structure, Function, and Development. *Crit. Rev. Biochem. Mol. Biol.* 30 (3), 197–262. doi:10.3109/10409239509085143

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25 (9), 1157–1174. doi:10.1002/jcc.20035

Yu, J., Zhou, Y., Tanaka, I., and Yao, M. (2010). Roll: a New Algorithm for the Detection of Protein Pockets and Cavities with a Rolling Probe Sphere. *Bioinformatics* 26 (1), 46–52. doi:10.1093/bioinformatics/btp599

Yu, S., Bojsen, K., Svensson, B., and Marcussen, J. (1999). alpha-1,4-glucan Lyases Producing 1,5-Anhydro-D-Fructose from Starch and Glycogen Have Sequence Similarity to Alpha-Glucosidases. *Biochim. Biophys. Acta* 1433 (1-2), 1–15. doi:10.1016/s0167-4838(99)00152-1

Zhu, J., Li, Y., Wang, J., Yu, Z., Liu, Y., Tong, Y., et al. (2018). Adaptive Steered Molecular Dynamics Combined with Protein Structure Networks Revealing the Mechanism of Y68I/G109P Mutations that Enhance the Catalytic Activity of D-Psicose 3-Epimerase from Clostridium Bolteae. *Front. Chem.* 6, 437. doi:10.3389/fchem.2018.00437

Zhu, J., Qi, R., Liu, Y., Zhao, L., and Han, W. (2019). Mechanistic Insights into the Effect of Ligands on Structural Stability and Selectivity of Sulfotransferase 2A1 (SULT2A1). *ACS Omega* 4 (26), 22021–22034. doi:10.1021/acsomega.9b03136

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

MINI REVIEW
published: 08 December 2021
doi: 10.3389/fchem.2021.796890

Mini Review: Quantum Confinement of Atomic and Molecular Clusters in Carbon Nanotubes

María Pilar de Lara-Castells[1]* *and Alexander O. Mitrushchenkov*[2]*

[1]Instituto de Física Fundamental (AbinitSim Unit), IFF-CSIC, Madrid, Spain, [2]MSME, Univ Gustave Eiffel, CNRS UMR 8208, Univ Paris Est Creteil, Marne-la-Vallée, France

We overview our recent developments on a computational approach addressing quantum confinement of light atomic and molecular clusters (made of atomic helium and molecular hydrogen) in carbon nanotubes. We outline a multi-scale first-principles approach, based on density functional theory (DFT)-based symmetry-adapted perturbation theory, allowing an accurate characterization of the dispersion-dominated particle–nanotube interaction. Next, we describe a wave-function-based method, allowing rigorous fully coupled quantum calculations of the pseudo-nuclear bound states. The approach is illustrated by showing the transition from molecular aggregation to quasi-one-dimensional condensed matter systems of molecular deuterium and hydrogen as well as atomic ^{4}He, as case studies. Finally, we present a perspective on future-oriented mixed approaches combining, e.g., orbital-free helium density functional theory (He-DFT), machine-learning parameterizations, with wave-function-based descriptions.

Keywords: clusters of molecular hydrogen, clusters of atomic helium, carbon nanotubes, quantum confinement, ab initio intermolecular interaction theory, wave-function method for bound-state calculations, full quantum coupled characterizations

OPEN ACCESS

Edited by:
Heribert Reis, National Hellenic Research Foundation, Greece

Reviewed by:
Utpal Sarkar, Assam University, India

***Correspondence:**
María Pilar de Lara-Castells
Pilar.deLara.Castells@csic.es
Alexander O. Mitrushchenkov
Alexander.Mitrushchenkov@univ-eiffel.fr

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

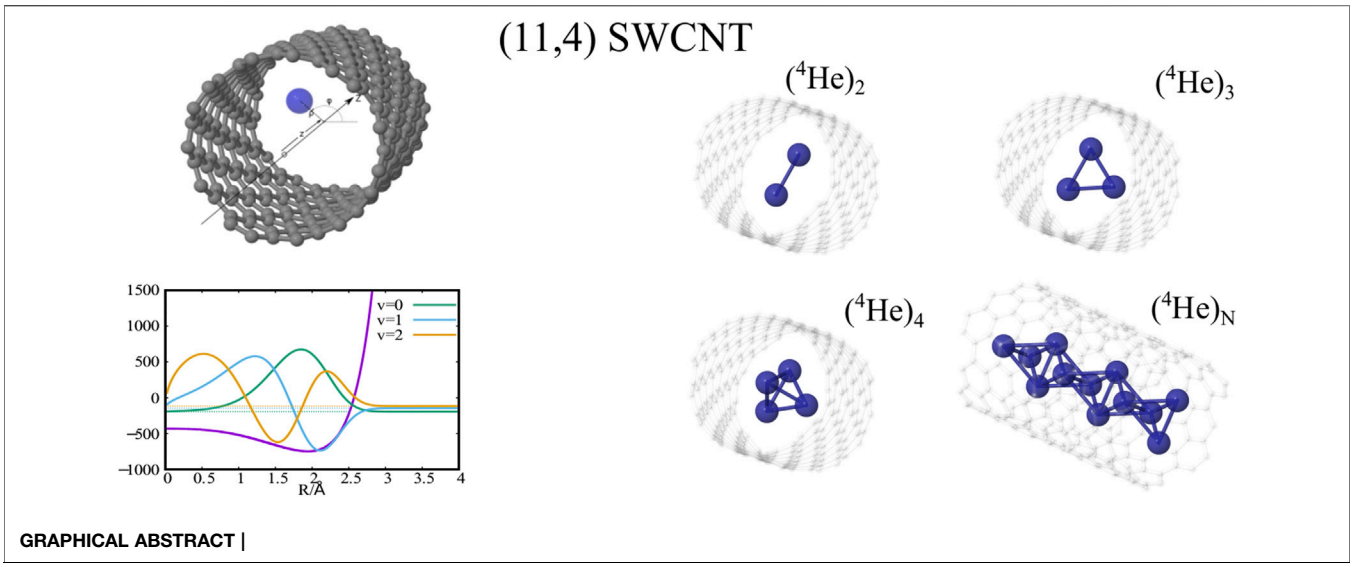
Received: *17 October 2021*
Accepted: *08 November 2021*
Published: *08 December 2021*

Citation:
de Lara-Castells MP and Mitrushchenkov AO (2021) Mini Review: Quantum Confinement of Atomic and Molecular Clusters in Carbon Nanotubes. Front. Chem. 9:796890. doi: 10.3389/fchem.2021.796890

1 INTRODUCTION

The cylindrical confinement provided by carbon nanotubes has offered the possibility of studying the pronounced quantum behaviour of ^{4}He atoms and H_2 molecules at reduced dimensionality. Recent measurements have demonstrated the formation of two-dimensional (2D) ^{4}He layers on the outer surface of single-walled carbon nanotubes (SWCNTs) (Noury et al., 2019). Also, an experimental study (Ohba, 2016) of gas adsorption at low (2–5 K) temperature revealed a quenched propagation of ^{4}He atoms through carbon nanopores with diameters below 7 Å despite of their small kinetic diameter. The application of orbital-free helium density functional theory (He-DFT) to carbon nanotubes immersed in a helium nanodroplet provided theoretical explication that the experimental observations stem from the exceptionally high zero-point energy of ^{4}He as well as its tendency to form two-dimensional (2D) layers upon adsorption at low temperatures (Hauser and de Lara-Castells, 2016). These conclusions were further confirmed by applying more accurate *ab initio* potential modelling along with a wave-function (WF)-based approach (Hauser et al., 2017). This study also showed that SWCNTs are filled by more molecules of N_2 than ^{4}He atoms, due to the higher zero-point energy of the latter. More generally, the interaction of small atomic and molecular systems with carbon-based nanostructures has attracted a lot of attention recently (see, e.g. Deb et al., 2016; Deb et al., 2019; Paul et al., 2019; Paul et al., 2020 and references therein).

For the case of molecular deuterium, our recent theoretical work has provided conclusive evidence for the transition from molecular aggregation to quantum solid-like packing in a SWCNT of 1 nm



GRAPHICAL ABSTRACT |

diameter (de Lara-Castells and Mitrushchenkov, 2020; de Lara-Castells and Mitrushchenkov, 2021), confirming a previous study using an embedding approach in a broader (ca. 1.4 nm) SWCNT (de Lara-Castells et al., 2017). Experimental evidence, using neutron scattering, on the formation of one-dimensional D_2 crystals under carbon nanotube confinement has been reported as well (Cabrillo et al., 2021). Altogether, these studies have confirmed the key role played by the quantum nature of the nuclear degrees of freedom in the confined atomic helium, molecular hydrogen, or molecular deuterium motion.

The quest for the understanding of the aggregation of molecular H_2 in carbon nanotubes is also application-oriented as it is being actively used as a clean energy source, substituting fossil fuels. In fact, its combustion produces only heat and water, and it can be efficiently combined with oxygen in a fuel cell to produce electricity. Yet, the usage of hydrogen as a profit fuel requires a substantial development of efficient storage materials (Schlapbach and Züttel, 2001; Züttel, 2003; Zheng et al., 2021) such as metal organic frameworks, Rosi et al. (2003) covalent organic frameworks (Zheng et al., 2021), and carbon-based nanoporous materials (Cheng et al., 2001; Liu et al., 2010). Hydrogen storage methods are aimed to pack hydrogen molecules as close as possible, existing direct experimental evidences for solid-like packing of hydrogen at temperatures of industrial importance (Ting et al., 2015). Actually, the question on the nature of the hydrogen packing in carbon nanotubes is truly fundamental. Thus, the possible existence of either a superfluid (Rossi and Ancilotto, 2016) or a crystalline phase (Del Maestro and Boninsegni, 2017; Ferré et al., 2017) for para-H_2 molecules inside carbon nanotubes at zero (Rossi and Ancilotto, 2016; Ferré et al., 2017) or ultra-low temperatures (Del Maestro and Boninsegni, 2017) have been addressed as well (see also (Cazorla and Boronat, 2017) for a recent review). Specially appealing in regards to the impact of quantum effects on the diffusion of H_2 and D_2 along carbon nanotubes is the experimental finding of reserved trends in their rates upon cooling (Nguyen et al., 2010), with the D_2 isotope becoming the faster inspite of its higher mass (Nguyen et al., 2010). The impact of quantum effects in confined H_2 and D_2 motion has been theoretically confirmed as well (Mondelo-Martell and Huarte-Larrañaga, 2016; Mondelo-Martell et al., 2017; Mondelo-Martell and Huarte-Larrañaga, 2021).

The high interest on the confinement of clusters of atomic ^{4}He is also related to the superfluid nature of ^{4}He droplets at a temperature close to absolute zero (0.4 K) (Mudrich and Stienkemeier, 2014) and, more specifically, to the onset of such fascinating property in doped clusters made of just a few ^{4}He atoms (Toennies and Vilesov, 2004). Then, the question is whether the confinement provided by SWCNTs can favour the emergence of such feature at reduced dimensionality. The effect of carbon nanotube confinement has been also shown to be remarkable for the ^{3}He isotope inside a SWCNT of 1 nm diameter. Thus, as opposed to the free dimer case, the confined ^{3}He dimer is predicted to be bound, mainly as a consequence of the strong localization along the radial ^{3}He–SWCNT direction (de Lara-Castells and Mitrushchenkov, 2021).

This overview article is organized as follows. The next section presents the computational approach addressing the modeling of intermolecular adsorbate–SWCNT interactions as well as fully coupled quantum calculations of the corresponding pseudo-nuclear wavefunctions. **Section 3** presents an illustrative application of the computational approach to the transition from molecular aggregation to quasi-one-dimensional (1D) matter systems composed by either D_2/H_2 molecules or ^{4}He atoms. Finally, **Section 4** closes with a summary of concluding remarks and a few future prospects.

2 COMPUTATIONAL APPROACH

2.1 *Ab Initio* Modelling of the Adsorbate-Nanotube Interaction

Achieving a correct description of the interaction of the lightest atomic and molecular clusters in nature (i.e., made of He or H_2) with its confining SWCNT environment is a challenge even for expensive *ab initio* methods since they are dominated by long-

range dispersion forces (van der Waals forces). Extended dispersion-corrected DFT methods are applicable including confinement effects in extended SWCNT structures but they bear a tendency to overshoot when applied to ultimate cases of dispersion-dominated (e.g., He-surface) interactions. One key idea has consisted in designing a functional [the so-called dlDF functional (Pernal et al., 2009)] which accounts for the dispersionless interaction energy only so that the dispersion corrections can be safely added later. Within a practical implementation of this idea, the incremental method (Stoll, 1992) is applied on non-periodic (small) cluster models of extended systems and combined with periodic dispersionless DFT calculations (de Lara-Castells et al., 2014b; de Lara-Castells et al., 2014a). This approach has been shown to be particularly successful when describing the interaction of atomic helium (de Lara-Castells et al., 2014a) as well as heavier noble gases (de Lara-Castells et al., 2015) with coronene/graphene/graphite surfaces. Very recently, it has been also demonstrated how modern *ab initio* intermolecular perturbation theory allows a cost-effective and accurate characterization of van der Waals-dominated He-SWCNT and H_2-SWCNT interactions using small clusters models of the SWCNTs. The use of small cluster models is justified as complicated, dispersionless contributions are mostly of short-range nature. Dispersion, on the other hand, is long-range, but the corresponding parameters show excellent transferability properties upon increasing the size of the surface cluster models (de Lara-Castells et al., 2014a; de Lara-Castells et al., 2015). Therefore, these parameters can be calculated at high level of *ab initio* theory on small clusters and then scaled to the actual SWCNT system. As previously emphasized (de Lara-Castells and Hauser, 2020), detailed energy decomposition schemes, an intrinsic feature of methods such as density functional theory (DFT)-based symmetry-adapted perturbation theory [SAPT(DFT)] (Hesselmann and Jansen, 2003; Misquitta et al., 2003; Heßelmann et al., 2005; Misquitta et al., 2005) has been shown to be particularly useful in this respect.

In our most recent works (de Lara-Castells and Mitrushchenkov, 2020; de Lara-Castells and Mitrushchenkov, 2021), the SAPT(DFT) method has been used to derive by fitting the dispersionless and dispersion SAPT(DFT) energy contributions to an additive pairwise potential model (PPM) (de Lara-Castells et al., 2016; Hauser and de Lara-Castells, 2017; de Lara-Castells et al., 2017; Hauser et al., 2017), which is a modified version of that proposed by Carlos and Cole (Carlos and Cole, 1980). For the illustrative cases presented in this work, these terms were calculated for the interaction between a single H_2 molecule (He atom) and a hydrogen-saturated (unsaturated) nanotube SWCNT(5, 5) tube made of 62 (40) atoms, considering two transverse sections of the tube. The SAPT(DFT) MOLPRO (Heßelmann et al., 2005; Werner et al., 2012) implementation was applied, using the Perdew-Burke-Ernzerhof (PBE) density functional (Perdew et al., 1996) and the augmented polarized correlation-consistent triple-zeta basis (Woon et al., 1994) for all atoms but the hydrogen atoms saturating the dangling bonds, for which the correlation-consistent double-zeta basis was used instead.

The PPM functional form (de Lara-Castells et al., 2016; Hauser and de Lara-Castells, 2017; de Lara-Castells et al., 2017; Hauser et al., 2017) for the dispersionless energy contribution accounts for the typical exponential growth of the dominant dispersionless term, the exchange-repulsion, but also including a Gaussian-type "cushion" to describe weakly attractive tails stemming from other dispersionless terms

$$E_{\rm int}^{\rm disp-less}(\{\mathbf{R}_{\rm A-C}\}) = \sum_{\rm C}[1+\gamma_R(1-\frac{6}{5}\cos^2\theta_{\rm C})] \times A\, e^{(-\alpha\, R_{\rm A-C}-\beta\, R_{\rm A-C}^2)}, \quad R_{\rm A-C} < R_c, \qquad (1)$$

where R_c is a cut-off distance, $R_{\rm A\text{-}C}$ stands for the distance between the adsorbate center-of-mass and one carbon atom of the SWCNT, and $\theta_{\rm C}$ is the angle between the radial vector going from the SWCNT center to one carbon atom and the vector $\mathbf{R}_{\rm A\text{-}C}$ pointing from the adsorbate center-of-mass to the same C atom. The dimensionless factor $\gamma_{\rm R}$ in the first term accounts for the anisotropy of the C – C bonds. The sum in **Eq. (1)** runs over all carbon atoms of the nanotube. For the dispersion part, we apply the typical C_6/C_8 expansion with the damping functions of Tang and Toennies f_n (n = 6, 8) (Tang and Toennies, 1984)

$$E_{\rm int}^{\rm disp}(\{\mathbf{R}_{\rm A-C}\}) = -\sum_{\rm C}[1+\gamma_A(1-\frac{3}{2}\cos^2\theta_{\rm C})] \times \sum_{n=6,8}\frac{\sqrt{C_n^{\rm A}C_n^{\rm C}}}{R_{\rm A-C}^n} f_n\Big(\sqrt{\beta_{\rm A}\beta_{\rm C}}\,R_{\rm AC}\Big), \qquad (2)$$

where $\gamma_{\rm A}$ is also a dimensionless anisotropy parameter. The inclusion of γ_A and γ_R anisotropy terms has been found to be important when modelling potential corrugation effects on both curved carbon and metallic surfaces (de Lara-Castells et al., 2016; Hauser and de Lara-Castells, 2017; Hauser et al., 2017).

2.2 The Pseudo-Nuclear Wave-Function Problem

Our method has been designed to allow bound-state calculations for N identical atoms or molecules (referred to as "pseudo-nuclei", PNs) inside a SWCNT. Using cylindrical coordinates ($x_i, y_i, z_i \rightarrow \rho_i, \phi_i, z_i$) for each PN (see left panel of **Figure 1**), together with a Jacobian transformation of the volume to keep the Hamiltonian "explicitly Hermitian", $\Psi \rightarrow \sqrt{\rho_1\rho_2\cdots\rho_N}\,\Psi$, the Hamiltonian describing a motion of N PN's takes the form

$$\hat{H}_N = \sum_{i=1}^{N}\left\{-\frac{1}{2M}\left[\frac{\partial^2}{\partial\rho_i^2}+\frac{\partial^2}{\partial z_i^2}+\frac{1}{\rho_i^2}\left(\frac{\partial^2}{\partial\phi_i^2}+\frac{1}{4}\right)\right]+V_1(\rho_i)\right\} + \sum_{i<j}V_2(r_{ij}) \qquad (3)$$

where the distance r_{ij} is explicitly given by

$$r_{ij}^2 = \rho_i^2+\rho_i^2+(z_i-z_j)^2-2\rho_i\rho_j\cos(\phi_i-\phi_j). \qquad (4)$$

We note that, in order to preserve the cylindrical symmetry of the system, the very small corrugation appearing along the azimuthal degree of freedom ϕ is not accounted for.

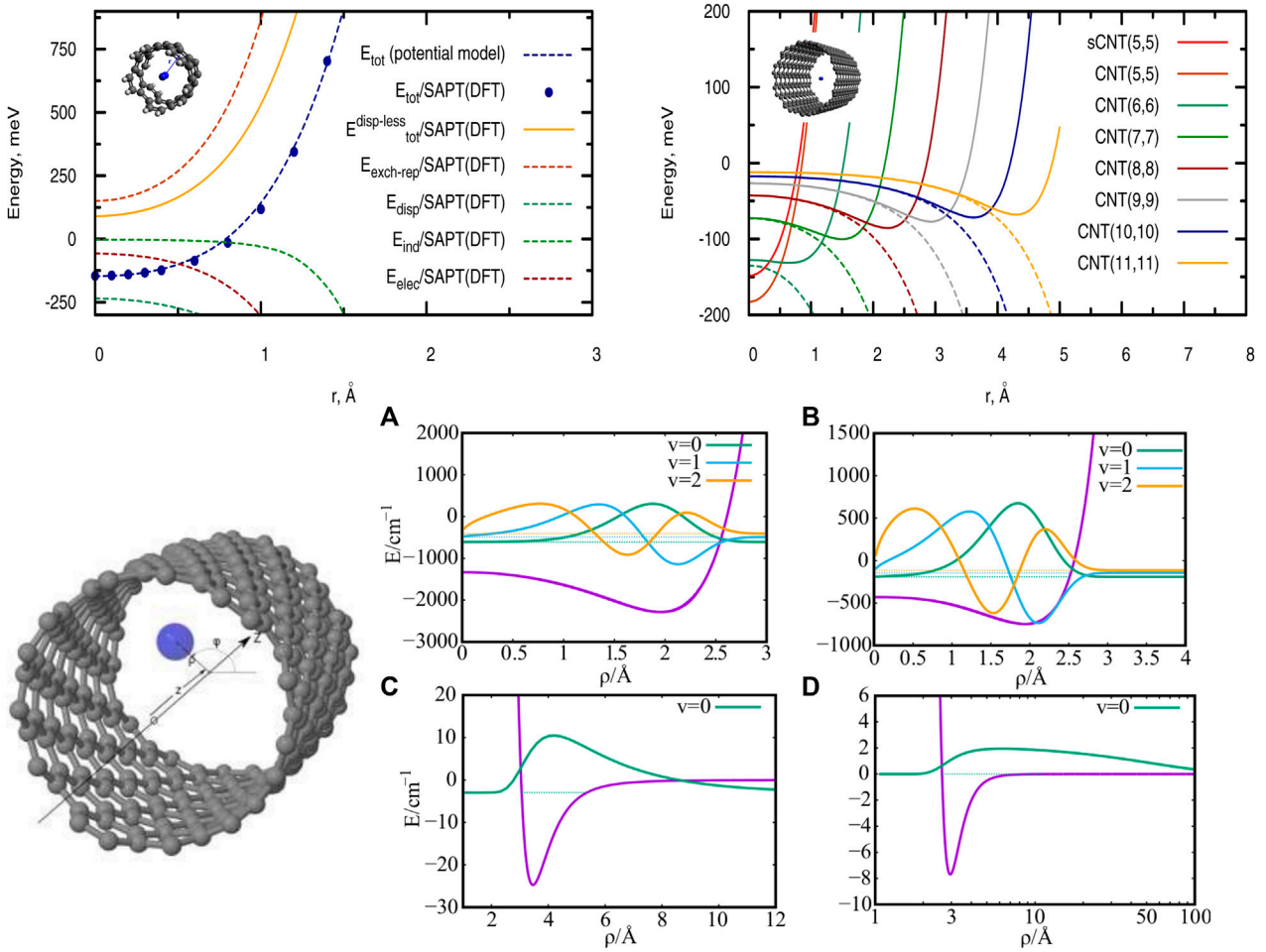


FIGURE 1 | Top panel. Left-hand side: Radial scan of the interaction energies between a single H_2 molecule and a short (single-walled) carbon nanotube (sCNT) of helicity index (5, 5). The pairwise potential model is compared with reference *ab initio* calculations using the SAPT(DFT) approach. Right-hand side: Radial scan of the total interaction energies (full lines) and dispersion contributions (dashed lines) between a single H_2 molecule and carbon nanotubes with helicity indexes of increasing value. Source: (de Lara-Castells et al., 2017) Bottom panel. Left-hand side: Cylindrical coordinates describing a single particle in a tube. Right-hand side: interaction potentials PN-CNT **(A, B)** and PN-PN **(C, D)** together with the corresponding low-lying bound states of a single confined PN. **(A, C)**: PN = H_2, **(B, D)**: PN = ^{4}He. Source: (de Lara-Castells and Mitrushchenkov, 2021).

Considering long nanotubes, it is thus assumed that the interaction potential depends only on ρ. The atomic structure of the SWCNT, however, is implicitly considered through the SAPT(DFT) calculations used to fit the parameters of our PPM (see **Section 3.1**). These potentials are used within the Discrete Variable Representation (DVR) approach (Bačič and Light, 1989) upon fitting them to polynomials of ρ:

$$V_1(\rho) = \sum_{l=0}^{l_{\max}} p_l \rho^l \tag{5}$$

To separate the overall Z-translation and overall rotational motions, we introduce relative coordinates as $t_i = z_i - z_N$ and $\chi_i = \phi_i - \phi_N$, for $i = 1 \ldots N-1$. The overall Z-translation and overall rotation coordinates are conveniently defined as $Z = Z_{\mathrm{COM}} = \frac{1}{N}\sum_{i=1}^{N} z_i$ and $\Phi = \phi_N$. The overall rotation quantum number reads $\Lambda = m_1 + m_2 + \cdots + m_N$, where m_i is the integer standing for the projection of the angular momentum of the i-th particle onto the nanotube axis. Since both PN–SWCNT and PN–PN interactions do not depend on Φ, the coordinate Φ can be separated via the factor $\exp(i\Lambda\Phi)\Psi(\chi_i)/\sqrt{2\pi}$. Finally, the full kinetic energy is written as $K_N = -\frac{1}{2M}\sum_{i=1}^{N}\Delta(\mathbf{r}_i) = K_N(z) + K_N(\rho) + K_N(\phi)$ where

$$K_N(z) = -\frac{1}{M}\sum_{i=1}^{N-1}\frac{\partial^2}{\partial t_i^2} - \frac{1}{M}\sum_{i<j=1}^{N-1}\frac{\partial}{\partial t_i}\frac{\partial}{\partial t_j} \tag{6}$$

$$K_N(\rho) = -\frac{1}{2M}\sum_{i=1}^{N}\frac{\partial^2}{\partial \rho_i^2} \tag{7}$$

and

$$K_N(\phi) = -\frac{1}{2M}\sum_{i=1}^{N-1}\frac{1}{\rho_i^2}\left[\frac{\partial^2}{\partial \chi_i^2} + \frac{1}{4}\right] - \frac{1}{2M}\frac{1}{\rho_N^2}\left[\left(i\Lambda - \sum_{i=1}^{N-1}\frac{\partial}{\partial \chi_i}\right)^2 + \frac{1}{4}\right] \tag{8}$$

We note that the internal Hamiltonian does not account explicitly for the bosonic permutation symmetry of all N particles. The symmetry is automatically included for the particles labelled as $1 \ldots N-1$ since it is equivalent to a simple exchange of the corresponding coordinates t_i, ρ_i, χ_i. The exchange of the particles labelled as i and N, however, results in linear transformations of the coordinates t_i and χ_i. The symmetry with respect to the $i \leftrightarrow N$ exchange is analyzed *a posteriori* by defining a "bosonic symmetry factor" Q as the matrix element:

$$Q = \langle \Psi_i | (1 \leftrightarrow N) | \Psi_j \rangle \tag{9}$$

The Q factor is unity for true bosonic solutions. This test is used as a criteria to evaluate the wavefunction accuracy.

To calculate the eigenvalues of the internal Hamiltonian, we use the DVR approach (Bačič and Light, 1989; de Lara-Castells and Mitrushchenkov, 2020; de Lara-Castells and Mitrushchenkov, 2021). This approach allows an easy and quick evaluation of the pair interaction potential V_2, with the basis set being obtained as a direct product of functions for the different coordinates. Sinc-DVR functions are conveniently employed for the t and χ coordinates. When dealing with the polar radii ρ, however, it is necessary to explicitly treat the singular kinetic energy term $-1/8M\rho_i^2$, which stems from the Jacobian transformation to cylindrical coordinates. For this purpose, we use the DVR basis obtained for the finite basis set representation (FBR) built from the radial functions of two-dimensional (2D) Harmonic oscillator functions. However, this basis renders the grid size too large for clusters with, e.g., 4 PNs since the internal Hamiltonian becomes 10-dimensional (10 D). This problem is efficiently solved by using potential-optimized DVR (PO-DVRs) functions (Echave and Clary, 1992), allowing for a very fast energy convergence as well. The diagonalization of the resulting Hamiltonian matrix is carried out using the Jacobi-Davidson algorithm (Sleijpen and van der Vorst, 1996). This technique has been very successful even for characterizing ill-behaved interactions such as, e.g., the "hard-core" interaction problem of doped helium clusters (de Lara-Castells et al., 2009a; de Lara-Castells et al., 2009b).

3 ILLUSTRATIVE APPLICATIONS

In the first subsection, we will illustrate how the interaction of molecular hydrogen with various SWCNTs can be modelled at *ab initio* level. In the second subsection, we will show how our computational approach has allowed to reveal the transition from van-der-Waals-type molecular aggregation to quasi-1D condensed matter systems for atomic ^{4}He and molecular H_2 and D_2 inside a SWCNT of 1 nm diameter.

3.1 H_2-Nanotube Interaction Potentials for SWCNTs of Increasing Diameter

As an illustrative case, the top (left-hand) panel of **Figure 1** shows the total H_2/CNT interaction potential as a function of the radial distance r between the molecule center-of-mass and the SWCNT(5,5) center, along with the electrostatic E_{elec}, exchange-repulsion $E_{exch\text{-}rep}$, induction E_{ind}, and dispersion contributions E_{disp}. Upon applying the pairwise PPM presented in **Section 3.1**, the radial scans of interaction energies shown in the upper (right-hand) panel of **Figure 1** are obtained for SWCNTs of increasing diameter.

As can be observed in the top panel of **Figure 1** (left-hand side), the H_2/CNT attractive interaction is dispersion-dominated, with the exchange-repulsion dominating the whole dispersionless term. This repulsive energy contribution grows exponentially as the molecule-surface distance decreases although such behaviour is somewhat smoothed out by the attractive electrostatic contribution. The induction term contributes very little at the potential minimum but considerably at the SWCNT cage. The potential minimum is located at the center of the narrowest nanotube (diameter of 6.74 Å), where the adsorbates benefit from the dispersion interaction with carbon atoms at both sides of the carbon cage. However, upon increasing the nanotube diameter (see right-hand panel at the top of **Figure 1**), the dispersion becomes very small at the nanotube center and the potential minima shift towards the carbon cage. For a nanotube of diameter of 27.21 Å, the value of the well-depth (−56 meV) is already close to that obtained for a graphene sheet and measured for the H_2 molecule adsorbed onto graphite (−51.7 meV from Ref. 29). From the top (left-hand side) panel of **Figure 1** it can also be observed that the pairwise potential model closely reproduces the results obtained with the SAPT(DFT) method. Similar conclusions hold for the case of the He/SWCNT interaction (Hauser et al., 2017), also emphasizing the usefulness of the SAPT(DFT)-derived PPM (see **Section 3.1**).

3.2 From van-der-Waals Aggregation to Quasi-One-Dimensional Chains

Considering a SWCNT of 1 nm diameter and helicity index (11, 4), the He-SWCNT and H_2-SWCNT one-particle potentials $V_1(\rho)$ are shown in the bottom (right-side) panels of **Figure 1** along with a few supported bound states of a single ^{4}He atom or H_2 molecule. **Figure 2** presents the plots of the ground-state two-dimensional (2D) densities of PN_4 clusters (PN = ^{4}He, para-H_2, and ortho-D_2).

As previously reported (de Lara-Castells and Mitrushchenkov, 2020; de Lara-Castells and Mitrushchenkov, 2021) the structuring of PN_N clusters, with $N = 1 - 3$, is characterized by a high pseudo-nuclear delocalization. Thus, the triangular-like structure exhibited by confined $(^4He)_3$ clusters (see graphical abstract) freely rotates inside the nanotube. However, as can be observed in **Figure 2**, whatever the bosonic particle be, the most probable structure is pyramidal-like. This confinement effect is specially remarkable for the case of ^{4}He, due to the high delocalization of the bare PN–PN dimer (see bottom panel (d) of **Figure 1**). Yet, it features an increasing spatial delocalization when going from molecular ortho-D_2 through molecular para-H_2 to atomic ^{4}He. Inspite of para-H_2 molecules being twice as light as ^{4}He atoms, the three times deeper attractive well of the pair potential (see bottom panels (a) and (b) of **Figure 1**)

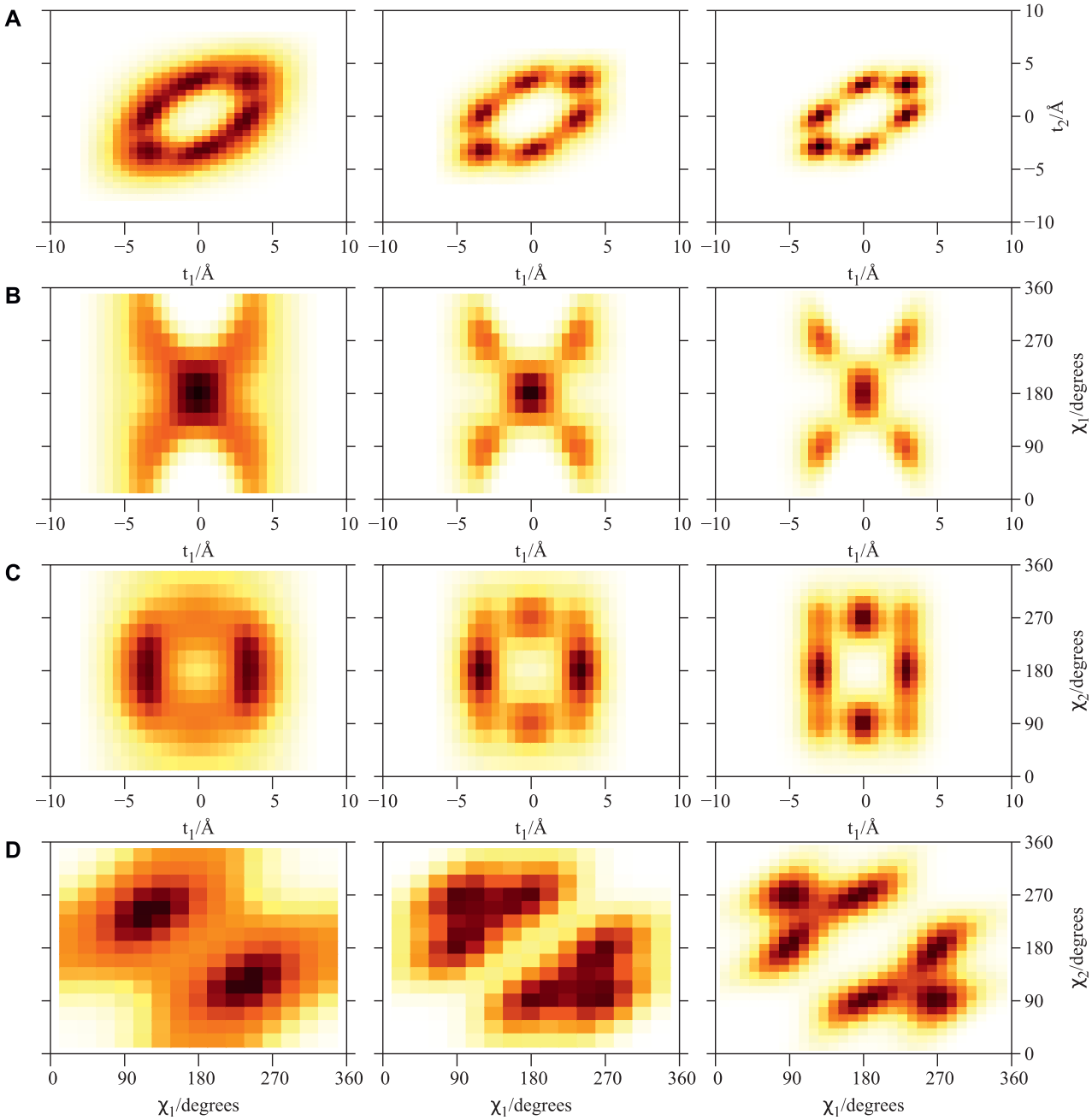


FIGURE 2 | Plot of 2D densities for the ground state of PN_4 c SWCNT(11, 4) complexes. Left to right: PN = ^{4}He, para-H_2, ortho-D_2. **(A)** t_1 and t_2 coordinates; **(B)** t_1 and χ_1 coordinates; **(C)** t_1 and χ_2 coordinates; **(D)** χ_1 and χ_2 coordinates. Source: (de Lara-Castells and Mitrushchenkov, 2021).

causes a much more compact structuring of the (para-H_2)$_4$ cluster when compared with the ^{4}He counterpart. Particularly, we note that the extended profile of the 2D densities in the χ coordinates reflects a quasi-independent relative rotational motion of two (^{4}He)$_2$ dimers lying orthogonal to the tube axis. Since ortho-D_2 molecules are twice as heavy as para-H_2 molecules, the solid-like nature of the (ortho-D_2)$_4$ structure (a regular tetrahedron) becomes even more apparent than for the (para-H_2)$_4$ counterpart. The formation of (vibrationally averaged) tetrahedral structures of (para-H_2)$_4$ clusters is not a special effect from the SWCNT confinement. In fact, the same feature has also been revealed in clathrate hydrate cages (Sebastianelli et al., 2008; Witt et al., 2010). These similarities indicate that the aggregation of multiple light PNs are driven by quantum nuclear effects and very weak dispersive PN–PN interaction forces rather than the much more stronger PN–SWCNT (or PNT-cage) energy contribution (see bottom panels (a) and (b) of **Figure 1**). As discussed in

detail in our previous work (de Lara-Castells and Mitrushchenkov, 2020), the onset of solid-like packing for molecular deuterium is explained by analyzing the potential minima landscape, allowing also to predict the formation of a one-dimensional chain of tetrahedral structures along the tube axis when the number of D_2 molecules increase. Similar conclusions holds for the cases of ^{4}He atoms and H_2 molecules (de Lara-Castells and Mitrushchenkov, 2021). Moreover, these special structures feature stabilization of collective rotational motion resembling the behaviour of quantum rings exhibiting persistent current (charged particles) or persistent flow (neutral particles). This characteristic has been connected with the onset of superfluid motion and persistent flow in quantum rings made of ^{4}He atoms by Bloch (Bloch, 1973). It is remarkable that such feature occurs already in clusters made of four ^{4}He atoms at the reduced dimensionality offered by the confining medium.

4 CONCLUDING REMARKS AND FUTURE DIRECTIONS

Summarizing, this mini review shows how a challenging case of confined molecular system can be accurately characterized using multi-scale first-principles modelling. In order to emphasize the key role of quantum effects, we have chosen the lightest atomic and molecular species in nature as the confined object and SWCNTs as the confining medium. In the first part, we have overviewed how an *ab initio*-derived pairwise potential model can accurately characterize the intermolecular interaction between the confining medium and the confined object. In the second part, we have illustrated how a computational approach, allowing for rigorous fully coupled quantum calculations of the pseudo-nuclear bound states, has been capable of revealing the transition from van der Waals-type molecular aggregation to quasi-1D condensed matter systems under cylindrical carbon nanotube confinement. Remarkably, it has been also shown that the structuring is driven by purely dispersive PN-PN interactions together with quantum pseudo-nuclear effects and not the much more stronger PN-tube interaction forces. In this way, it can be also understood why, inspite of its weak cohesive PN-PN interaction, neutron scattering experimental measurements have very recently shown the formation of quasi-one-dimensional structures of molecular D_2 corresponding to cylindrical sections of the hexagonal close-packed bulk crystal (Cabrillo et al., 2021).

As a future perspective, we are aimed to adapt our Full Configuration Interaction Nuclear Orbital (FCI-NO) approach (de Lara-Castells et al., 2009b), originally developed for doped helium clusters, to the present case in which the carbon nanotube acts as the "dopant" species. The essential advantage of this *ansatz* is that the wave function is expanded using products of nuclear orbitals, allowing the employment of second quantization techniques, an automatic inclusion of Fermi–Dirac and Bose–Einstein nuclear spin statistical effects, and, then, a characterization of ^{3}He and ^{4}He isotopes on an equal footing. The second future prospect is the creation of embedding approaches linking our wave-function-based description with orbital free He-DFT (Dalfovo et al., 1995; Ancilotto et al., 2017), following previous efforts (Hauser et al., 2017). In this sense, the highly accurate method presented in our latest work (de Lara-Castells and Mitrushchenkov, 2021) is expected to provide benchmark results guiding new, machine-learning driven, parameterizations in He-DFT, as already illustrated in electronic structure theory (Meyer et al., 2020).

AUTHOR CONTRIBUTIONS

MPdLC and AOM have equally contributed to the conceptualization of this work, and the writing and edition of this manuscript.

FUNDING

This work has been partly supported by the Spanish Agencia Estatal de Investigación (AEI) and the Fondo Europeo de Desarrollo Regional (FEDER, UE) under Grant Nos. PID 2020-117605GB-I00 and MAT 2016-75354-P.

ACKNOWLEDGMENTS

We thank Andreas W. Hauser Graz University of Technology and Ricardo Fernández-Perea IEM-CSIC, Spain for their previous and very fruitful collaboration on the topic of this mini review.

REFERENCES

Ancilotto, F., Barranco, M., Coppens, F., Eloranta, J., Halberstadt, N., Hernando, A., et al. (2017). Density Functional Theory of Doped Superfluid Liquid Helium and Nanodroplets. *Int. Rev. Phys. Chem.* 36, 621–707. doi:10.1080/0144235X.2017.1351672

Bačić, Z., and Light, J. C. (1989). Theoretical Methods for Rovibrational States of Floppy Molecules. *Annu. Rev. Phys. Chem.* 40, 469–498. doi:10.1146/annurev.pc.40.100189.002345

Bloch, F. (1973). Superfluidity in a Ring. *Phys. Rev. A.* 7, 2187–2191. doi:10.1103/PhysRevA.7.2187

Cabrillo, C., Fernández-Perea, R., Bermejo, F. J., Chico, L., Mondelli, C., González, M. A., et al. (2021). Formation of One-Dimensional Quantum Crystals of Molecular Deuterium inside Carbon Nanotubes. *Carbon* 175, 141–154. doi:10.1016/j.carbon.2020.12.067

Carlos, W. E., and Cole, M. W. (1980). Interaction between a He Atom and a Graphite Surface. *Surf. Sci.* 91, 339–357. doi:10.1016/0039-6028(80)90090-4

Cazorla, C., and Boronat, J. (2017). Simulation and Understanding of Atomic and Molecular Quantum Crystals. *Rev. Mod. Phys.* 89, 035003. doi:10.1103/RevModPhys.89.035003

Cheng, H.-M., Yang, Q.-H., and Liu, C. (2001). Hydrogen Storage in Carbon Nanotubes. *Carbon* 39, 1447–1454. doi:10.1016/s0008-6223(00)00306-7

Dalfovo, F., Lastri, A., Pricaupenko, L., Stringari, S., and Treiner, J. (1995). Structural and Dynamical Properties of Superfluid Helium: A Density-Functional Approach. *Phys. Rev. B* 52, 1193–1209. doi:10.1103/PhysRevB.52.1193

Deb, J., Bhattacharya, B., Paul, D., and Sarkar, U. (2016). Interaction of Nitrogen Molecule with Pristine and Doped Graphyne Nanotube. *Phys. E* 84, 330–339. doi:10.1016/j.physe.2016.08.006

Deb, J., Paul, D., and Sarkar, U. (2019). Donor-acceptor Decorated Graphyne: A Promising Candidate for Nonlinear Optical Application. *Dae Solid State Phys. Symp.* doi:10.1063/1.5113008

de Lara-Castells, M. P., Bartolomei, M., Mitrushchenkov, A. O., and Stoll, H. (2015). Transferability and Accuracy by Combining Dispersionless Density Functional and Incremental post-Hartree-Fock Theories: Noble Gases Adsorption on Coronene/Graphene/Graphite Surfaces. *J. Chem. Phys.* 143, 194701. doi:10.1063/1.4935511

de Lara-Castells, M. P., Fernández-Perea, R., Madzharova, F., and Voloshina, E. (2016). Post-Hartree-Fock Studies of the He/Mg(0001) Interaction: Anti-corrugation, Screening, and Pairwise Additivity. *J. Chem. Phys.* 144, 244707. doi:10.1063/1.4954772

de Lara-Castells, M. P., Hauser, A. W., Mitrushchenkov, A. O., and Fernández-Perea, R. (2017). Quantum Confinement of Molecular Deuterium Clusters in Carbon Nanotubes: Ab Initio Evidence for Hexagonal Close Packing. *Phys. Chem. Chem. Phys.* 19, 28621–28629. doi:10.1039/C7CP05869A

de Lara-Castells, M. P., and Hauser, A. W. (2020). New Tools for the Astrochemist: Multi-Scale Computational Modelling and Helium Droplet-Based Spectroscopy. *Phys. Life Rev.* 32, 95–98. doi:10.1016/j.plrev.2019.08.001

de Lara-Castells, M. P., and Mitrushchenkov, A. O. (2021). A Nuclear Spin and Spatial Symmetry-Adapted Full Quantum Method for Light Particles inside Carbon Nanotubes: Clusters of ^{3}He, ^{4}He, and Para-H_2. *Phys. Chem. Chem. Phys.* 23, 7908–7918. doi:10.1039/d0cp05332e

de Lara-Castells, M. P., Mitrushchenkov, A. O., Delgado-Barrio, G., and Villarreal, P. (2009a). Using a Jacobi-Davidson "Nuclear Orbital" Method for Small Doped ^{3}He Clusters. *Few-body Syst.* 45, 233–236. doi:10.1007/s00601-009-0035-6

de Lara-Castells, M. P., and Mitrushchenkov, A. O. (2020). From Molecular Aggregation to a One-Dimensional Quantum crystal of Deuterium inside a Carbon Nanotube of 1 nm Diameter. *J. Phys. Chem. Lett.* 11, 5081–5086. doi:10.1021/acs.jpclett.0c01432

de Lara-Castells, M. P., Stoll, H., Civalleri, B., Causà, M., Voloshina, E., Mitrushchenkov, A. O., et al. (2014a). Communication: A Combined Periodic Density Functional and Incremental Wave-Function-Based Approach for the Dispersion-Accounting Time-Resolved Dynamics of ^{4}He Nanodroplets on Surfaces: ^{4}He/Graphene. *J. Chem. Phys.* 141, 151102. doi:10.1063/1.4898430

de Lara-Castells, M. P., Stoll, H., and Mitrushchenkov, A. O. (2014b). Assessing the Performance of Dispersionless and Dispersion-Accounting Methods: Helium Interaction with Cluster Models of the TiO_2(110) Surface. *J. Phys. Chem. A.* 118, 6367–6384. doi:10.1021/jp412765t

de Lara-Castells, M. P., Villarreal, P., Delgado-Barrio, G., and Mitrushchenkov, A. O. (2009b). An Optimized Full-Configuration-Interaction Nuclear Orbital Approach to a "Hard-Core" Interaction Problem: Application to (^{3}He$_N$)-Cl_2(B) Clusters (N≤4). *J. Chem. Phys.* 131, 194101. doi:10.1063/1.3263016

Del Maestro, A., and Boninsegni, M. (2017). Absence of Superfluidity in a Quasi-One-Dimensional Parahydrogen Fluid Adsorbed inside Carbon Nanotubes. *Phys. Rev. B* 95, 054517. doi:10.1103/PhysRevB.95.054517

Echave, J., and Clary, D. C. (1992). Potential Optimized Discrete Variable Representation. *Chem. Phys. Lett.* 190, 225–230. doi:10.1016/0009-2614(92)85330-D

Ferré, G., Gordillo, M. C., and Boronat, J. (2017). Luttinger parameter of quasi-one-dimensional para–H_2. *Phys. Rev. B* 95, 064502. doi:10.1103/PhysRevB.95.064502

Sleijpen, G. L., and van der Vorst, H. A. (1996). A Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems. *SIAM J. Matrix Anal. Appl.* 17, 401–425. doi:10.1137/s0895479894270427

Hauser, A. W., and de Lara-Castells, M. P. (2016). Carbon Nanotubes Immersed in Superfluid Helium: The Impact of Quantum Confinement on Wetting and Capillary Action. *J. Phys. Chem. Lett.* 7, 4929–4935. doi:10.1021/acs.jpclett.6b02414

Hauser, A. W., and de Lara-Castells, M. P. (2017). Spatial Quenching of a Molecular Charge-Transfer Process in a Quantum Fluid: the C^{sx}-C_{60} Reaction in Superfluid Helium Nanodroplets. *Phys. Chem. Chem. Phys.* 19, 1342–1351. doi:10.1039/C6CP06858H

Hauser, A. W., Mitrushchenkov, A. O., and de Lara-Castells, M. P. (2017). Quantum Nuclear Motion of Helium and Molecular Nitrogen Clusters in Carbon Nanotubes. *J. Phys. Chem. C* 121, 3807–3821. doi:10.1021/acs.jpcc.6b12959

Hesselmann, A., and Jansen, G. (2003). Intermolecular Dispersion Energies from Time-dependent Density Functional Theory. *Chem. Phys. Lett.* 367, 778–784. doi:10.1016/S0009-2614(02)01796-7

Heßelmann, A., Jansen, G., and Schütz, M. (2005). Density-functional Theory-Symmetry-Adapted Intermolecular Perturbation Theory with Density Fitting: A New Efficient Method to Study Intermolecular Interaction Energies. *J. Chem. Phys.* 122, 014103. doi:10.1063/1.1824898

Liu, C., Chen, Y., Wu, C.-Z., Xu, S.-T., and Cheng, H.-M. (2010). Hydrogen Storage in Carbon Nanotubes Revisited. *Carbon* 48, 452–455. doi:10.1016/j.carbon.2009.09.060

Mattera, L., Rosatelli, F., Salvo, C., Tommasini, F., Valbusa, U., and Vidali, G. (1980). Selective Adsorption of $^1H_2$ and $^2H_2$ on the (0001) Graphite Surface. *Surf. Sci.* 93, 515–525. doi:10.1016/0039-6028(80)90279-4

Meyer, R., Weichselbaum, M., and Hauser, A. W. (2020). Machine Learning Approaches toward Orbital-free Density Functional Theory: Simultaneous Training on the Kinetic Energy Density Functional and its Functional Derivative. *J. Chem. Theor. Comput.* 16, 5685–5694. doi:10.1021/acs.jctc.0c00580

Misquitta, A. J., Jeziorski, B., and Szalewicz, K. (2003). Dispersion Energy from Density-Functional Theory Description of Monomers. *Phys. Rev. Lett.* 91, 033201. doi:10.1103/PhysRevLett.91.033201

Misquitta, A. J., Podeszwa, R., Jeziorski, B., and Szalewicz, K. (2005). Intermolecular Potentials Based on Symmetry-Adapted Perturbation Theory with Dispersion Energies from Time-dependent Density-Functional Calculations. *J. Chem. Phys.* 123, 214103. doi:10.1063/1.2135288

Mondelo-Martell, M., and Huarte-Larrañaga, F. (2021). Competition of Quantum Effects in H2/D2 Sieving in Narrow Single-Wall Carbon Nanotubes. *Mol. Phys.* 119, e1942277. doi:10.1080/00268976.2021.1942277

Mondelo-Martell, M., and Huarte-Larrañaga, F. (2016). Diffusion of H_2 and D_2 Confined in Single-Walled Carbon Nanotubes: Quantum Dynamics and Confinement Effects. *J. Phys. Chem. A.* 120, 6501–6512. doi:10.1021/acs.jpca.6b00467

Mondelo-Martell, M., Huarte-Larrañaga, F., and Manthe, U. (2017). Quantum Dynamics of H_2 in a Carbon Nanotube: Separation of Time Scales and Resonance Enhanced Tunneling. *J. Chem. Phys.* 147, 084103. doi:10.1063/1.4995550

Mudrich, M., and Stienkemeier, F. (2014). Photoionisaton of Pure and Doped Helium Nanodroplets. *Int. Rev. Phys. Chem.* 33, 301–339. doi:10.1080/0144235X.2014.937188

Nguyen, T. X., Jobic, H., and Bhatia, S. K. (2010). Microscopic Observation of Kinetic Molecular Sieving of Hydrogen Isotopes in a Nanoporous Material. *Phys. Rev. Lett.* 105, 085901. doi:10.1103/PhysRevLett.105.085901

Noury, A., Vergara-Cruz, J., Morfin, P., Plaçais, B., Gordillo, M. C., Boronat, J., et al. (2019). Layering Transition in Superfluid Helium Adsorbed on a Carbon Nanotube Mechanical Resonator. *Phys. Rev. Lett.* 122. doi:10.1103/physrevlett.122.165301

Ohba, T. (2016). Limited Quantum Helium Transportation through Nano-Channels by Quantum Fluctuation. *Sci. Rep.* 6. doi:10.1038/srep28992

Paul, D., Deb, J., and Sarkar, U. (2019). Influence of Noble Gas Atoms on B12N12 Fullerene: A DFT Study. *Dae Solid State Phys. Symp.* doi:10.1063/1.5113010

Paul, D., Dua, H., and Sarkar, U. (2020). Confinement Effects of a Noble Gas Dimer inside a Fullerene Cage: Can it Be Used as an Acceptor in a DSSC?. *Front. Chem.* 8. doi:10.3389/fchem.2020.00621

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* 77, 3865–3868. doi:10.1103/physrevlett.77.3865

Pernal, K., Podeszwa, R., Patkowski, K., and Szalewicz, K. (2009). Dispersionless Density Functional Theory. *Phys. Rev. Lett.* 103, 263201. doi:10.1103/physrevlett.103.263201

Rosi, N. L., Eckert, J., Eddaoudi, M., Vodak, D. T., Kim, J., O'Keeffe, M., et al. (2003). Hydrogen Storage in Microporous Metal-Organic Frameworks. *Science* 300, 1127–1129. doi:10.1126/science.1083440

Rossi, M., and Ancilotto, F. (2016). Superfluid Behavior of Quasi-One-dimensionalp–H_2 inside a Carbon Nanotube. *Phys. Rev. B* 94, 100502. doi:10.1103/PhysRevB.94.100502

Schlapbach, L., and Züttel, A. (2001). Hydrogen-Storage Materials for Mobile Applications. *Nature* 414, 353–358. doi:10.1038/35104634

Sebastianelli, F., Xu, M., and Bačić, Z. (2008). Quantum Dynamics of Small H_2 and D_2 Clusters in the Large Cage of Structure II Clathrate Hydrate: Energetics, Occupancy, and Vibrationally Averaged Cluster Structures. *J. Chem. Phys.* 129, 244706. doi:10.1063/1.3049781

Stoll, H. (1992). On the Correlation Energy of Graphite. *J. Chem. Phys.* 97, 8449–8454. doi:10.1063/1.463415

Tang, K. T., and Toennies, J. P. (1984). An Improved Simple Model for the Van Der Waals Potential Based on Universal Damping Functions for the Dispersion Coefficients. *J. Chem. Phys.* 80, 3726–3741. doi:10.1063/1.447150

Ting, V. P., Ramirez-Cuesta, A. J., Bimbo, N., Sharpe, J. E., Noguera-Diaz, A., Presser, V., et al. (2015). Direct Evidence for Solid-like Hydrogen in a Nanoporous Carbon Hydrogen Storage Material at Supercritical Temperatures. *ACS Nano* 9, 8249–8254. doi:10.1021/acsnano.5b02623.PMID:26171656

Toennies, J. P., and Vilesov, A. F. (2004). Superfluid Helium Droplets: A Uniquely Cold Nanomatrix for Molecules and Molecular Complexes. *Angew. Chem. Int. Ed.* 43, 2622–2648. doi:10.1002/anie.200300611

Werner, H. J., Knowles, P. J., Knizia, G., Manby, F. R., Schütz, M., Celani, P., et al. (2012). Molpro, Version 2012.1, a Package of *Ab Initio* Programs, See Available at: http://www.molpro.net [Dataset].

Witt, A., Sebastianelli, F., Tuckerman, M. E., and Bačić, Z. (2010). Path Integral Molecular Dynamics Study of Small H_2 Clusters in the Large Cage of Structure II Clathrate Hydrate: Temperature Dependence of Quantum Spatial Distributions. *J. Phys. Chem. C* 114, 20775–20782. doi:10.1021/jp107021t

Woon, D. E., DunningJr., and Dunning, T. H. (1994). Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of Static Electrical Response Properties. *J. Chem. Phys.* 100, 2975–2988. doi:10.1063/1.466439

Zheng, J., Wang, C.-G., Zhou, H., Ye, E., Xu, J., Li, Z., et al. (2021). Current Research Trends and Perspectives on Solid-State Nanomaterials in Hydrogen Storage. *Research*. doi:10.34133/2021/3750689

Züttel, A. (2003). Materials for Hydrogen Storage. *Mater. Today* 6, 24–33. doi:10.1016/S1369-7021(03)00922-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 05 January 2022
doi: 10.3389/fchem.2021.800541

Supramolecular Approach to Tuning the Photophysical Properties of Quadrupolar Squaraines

Anna Kaczmarek-Kędziera[1]*, *Borys Ośmiałowski*[1], *Piotr S. Żuchowski*[2] and *Dariusz Kędziera*[1]

[1]Faculty of Chemistry, Nicolaus Copernicus University in Toruń, Toruń, Poland, [2]Institute of Physics, Nicolaus Copernicus University in Toruń, Toruń, Poland

In the present study, the influence of the hydrogen bonding for the one- and two-photon absorption of the prototypical squaraine dye is investigated with quantum chemistry tools. The central squaraine unit is bound by strong hydrogen bonds with 4-substituted N,N′-diphenylurea and, alternatively, N,N′-diphenylthiourea molecules, which affects to a high extend the properties of the squaraine electron accepting moiety, thus shifting its maximum absorption wavelength and enhancing the TPA cross section. The replacement of oxygen by sulfur atoms in the squaraine central ring, known to affect its photophysical behavior, is considered here as the way of modifying the strength and nature of the intermolecular contacts. Additionally, the influence of the oxygen-by-sulfur replacement is also considered in the N,N′-diphenylurea moiety, as the factor affecting the acidity of the N–H protons. The introduction of the sequence of the substituents of varying electron-donating or electron-withdrawing characters in the position 4 of N,N′-diphenyl(thio)urea subsystems allows to finely tune the hydrogen bonding with the central squaraine unit by further modification of the N–H bond characteristics. All of these structural modifications lead to the controlled adjustment of the electron density distribution, and thus, the properties affected such as transition moments and absorption intensity. *Ab initio* calculations provide strong support for this way of tailoring of one- or two-photon absorption due to the obtained strong hypsochromic shift of the maximum one-photon absorption wavelength observed particularly for thiosquaraine complexes and an increase in the TPA wavelength together with the increase in the TPA cross section. Moreover, the source of the strong modification of the thiosquaraine OPA in contrast to the pristine oxosquaraine upon N,N′-diphenyl(thio)urea substitution is determined. Furthermore, for the first time, the linear dependence of the non-additivity in the interaction energy on the Hammett substituent constant is reported. The stronger the electron-donating character of the substituent, the larger the three-body non-additive components and the larger their percentage to the total interaction energy.

Keywords: squaraine dye, photophysical properties, DFT calculations, interaction energy, SAPT energy decomposition, one-photon absorption, hydrogen bonding, non-additivity

OPEN ACCESS

Edited by:
Nino Russo,
University of Calabria, Italy

Reviewed by:
Anton Stasyuk,
Universitat de Girona, Spain
Gloria Cardenas-Jiron,
University of Santiago, Chile

***Correspondence:**
Anna Kaczmarek-Kędziera
teoadk@chem.umk.pl

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *23 October 2021*
Accepted: *03 December 2021*
Published: *05 January 2022*

Citation:
Kaczmarek-Kędziera A, Ośmiałowski B, Żuchowski PS and Kędziera D (2022) Supramolecular Approach to Tuning the Photophysical Properties of Quadrupolar Squaraines. Front. Chem. 9:800541. doi: 10.3389/fchem.2021.800541

1 INTRODUCTION

The growing interest in the applications of chromophores and fluorophores in the fields of biomedical techniques such as bioimaging or photodynamic therapy or in photovoltaic devices requires the rational design of the photoactive systems with respect to both their photophysical characteristics and solubility, stability, non-toxicity, and availability. Despite the numerous classes of molecules exhibiting the desired features at least in one of these fields (for instance, azobenzenes, merocyanines, coumarins, phthalocyanines, or dipyrromethene derivatives, just to mention a few), the continuous need of the precise and controlled tuning of their properties increases the number of contributions devoted to their modifications.

Squaraine dyes belong to the class of quadrupolar molecules of high interests in material chemistry, bioimaging, non-linear optics, or photonics. Their peculiar photooptical properties arise from their unique structure: an electron-deficient four-membered squaric acid ring (denoted further by A as acceptor) is placed in between two electron-rich donating groups (denoted by D as donors). This D–A–D structure results in a specific strong and sharp one-photon absorption (OPA) in the visible or near-IR region, exceptional brightness, and unique non-linear properties. However, squaraines are rarely used as probes or in biomedical applications due to their reported low solubility and lack of stability in biological media Ros-Lis et al. (2002); Ros-Lis et al. (2004); Karpenko et al. (2015). Additionally, the extended π-electron scaffold of squaraines promotes the stacking intermolecular interactions in polar solvents, causing the fluorescence quenching Arunkumar et al. (2007).

Numerous structural modifications of squaraine dyes have been investigated with respect to the particular features required for given applications. It is well-known that the proper introduction of the electron-withdrawing (EW) and electron-donating (ED) moieties to the molecular framework allows to finely tune both the maximum absorption wavelength and the TPA cross section, σ_{TPA}. It has been demonstrated only recently that the terminal electron–withdrawing substituents inserted symmetrically in the indolenine squaraine dyes affect the transition dipole moments, the difference between static ground and excited state dipole moment, and absorption wavelength stronger than the electron-donating groups placed alike in the study by Barcenas et al. (2021). However, the maximum absorption wavelength in both cases (EW and ED) is shifted bathochromically and other properties considered by Barcenas et al. are also modified in the same direction for electrons being transferred by the substituent from or to the central squaric ring moiety. This can be perceived as a severe limitation of the tailoring strategies for the squaraine dyes in context of their desired features vital for further usage.

The applications of squaraines as photosensitizers in photodynamic therapy or in photon upconversion demands the efficient generation of stable triplet states. The enhancement of the intersystem crossing can be achieved by the introduction of heavy atoms. However, even only the oxygen-to-sulfur replacement in the squaric acid ring has been shown to increase the quantum yield for the triplet excited state, and thus it make squaraines the attractive triplet state photosensitizers with relatively weak modification of other photophysical properties, such as two-photon absorption Webster et al. (2010); Avirah et al. (2012); Peceli et al. (2013).

The presence of the D–A–D motif may lead to the efficient two-photon absorption (TPA) of squaraine systems. The simple symmetrical dibutylaniline squaraine dye (SQ) has been shown to exhibit the strong two-photon absorption with three bands: vibronic coupling band of the TPA cross section equal to 200 GM, second band at 850 nm with σ^{TPA} = 2000 GM corresponding to the S_2 excited state, and the third one at 700 nm with the cross section of 15,000 GM is governed by the excitation to the S_4 state (Webster et al., 2010; Ferrer et al. (2019)). Squaraines of the extremely large values of σ_{TPA} have been obtained (for instance, 27 ,000 GM for heterocyclic pyrrole-substituted squaric acid ring) (Pawlicki et al., 2009). Among the popular techniques of optimization of the TPA properties of chromophores one should mention the extension of the π-electron–conjugated chain of the system, modification of the electron-donating and electron-accepting properties of the central and side moieties, or the conformational flexibility of the parts of the system. However the development of the whole molecule often aggravates the practical applications of the obtained system, for instance, significantly decreasing the solubility. For that reason, it is desired to design small molecules exhibiting the enhanced non-linear response.

It has been shown that the hydrogen bonding (HB)–governed aggregation can lead to the enhanced TPA (K. Liu and Y. Wang and Y. Tu and H. Å gren and Y. Liu (2007), Liu (2008)). Therefore, one could exploit the supramolecular architecture in order to improve the squaraine two-photon optical properties, at the same time aiming at the more flexible yet controllable variations of the dipole moments upon excitation or photophysical parameters. Among the systems widely investigated in the field of hydrogen-bonded supramolecular aggregation, the N,N′-diphenylurea (DPU) and their derivatives have occurred particularly interesting due to their easy crystallization and a well-defined crystal structure. This allows us to study the influence of the substitution effects on the structural and electronic properties of the aggregates. Since the DPU moiety possesses two strong hydrogen bond donors, it is expected to form stable and regular structures with a matching hydrogen bond acceptor. One can thus construct supramolecular systems built of the squaraine central moiety caught by the two urea derivatives forming the bifurcated hydrogen bonds to each of the oxygen atoms of the central four-membered squaraine ring, as presented schematically in **Figure 1**. The application of the N,N′-diphenylurea allows to finely tune the acidity of the N–H protons by the appropriate choice of the substituents in position 4 of the phenyl rings. Additionally, a similar sequence of the systems could be obtained by the utilization of the thiourea despite the urea system, since they are known to exhibit higher acidity of the N-H protons and are widely applied in organocatalysis.

The hydrogen bonding properties of numerous molecules can be influenced by substitution of the oxygen atom by sulfur.

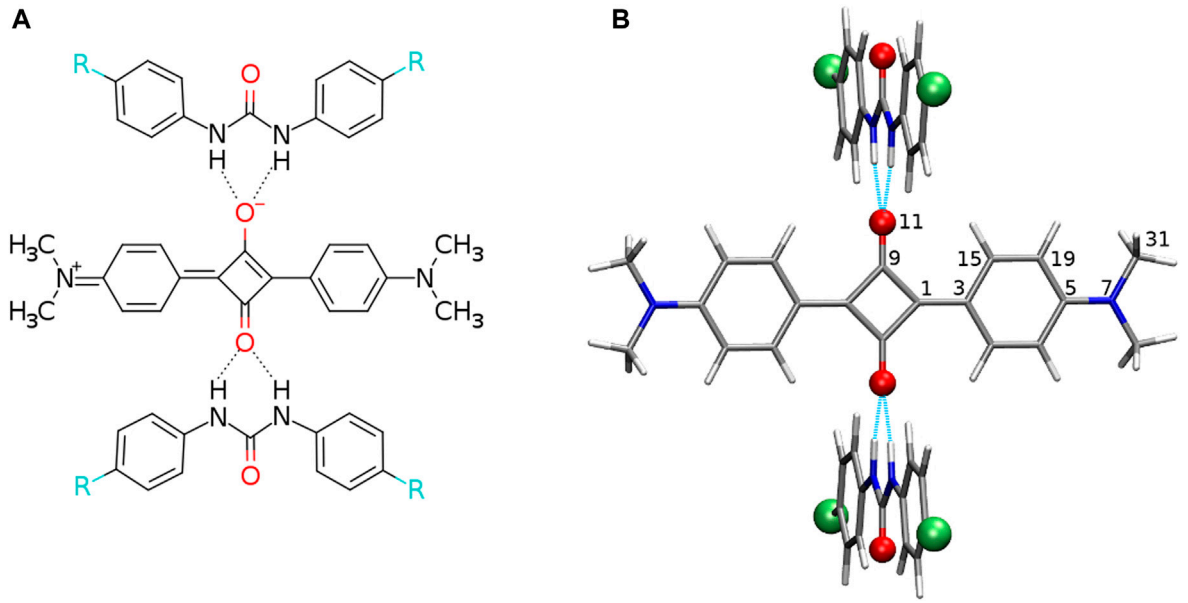


FIGURE 1 | (A) Structural formula of the analyzed hydrogen-bonded squaraine complex with 2 N,N′-diphenylurea molecules (R denotes electron-donating or electron-accepting functional groups) and **(B)** exemplary optimized structure with the urea plane perpendicular to the squaraine plane (red balls placed for oxygen or sulfur atoms, green balls depict the protons substituted with electron-donating or electron-accepting functional groups, and intermolecular hydrogen bonds are shown in cyan dashed line). Atom numbering for the symmetrical and unique atoms is given.

This was exemplified by numerous carbonyl and thiocarbonyl compounds (Hinchliffe, 1984; Platts et al., 1996; Allen et al., 1997; Rablen et al., 1998; Steiner, 2000; Krepps et al., 2001; Wennmohs et al., 2003; Valdés-Martínez et al., 2004; Jabłoński et al., 2006; Corpinot et al., 2017). The sulfur atom is known to be significantly weaker hydrogen bond acceptor than oxygen but also to exhibit the preference for larger deviation from 180° than in oxygen moieties (Platts et al., 1996). This arises from the fact, that opposite to the charge–charge interaction dominating in the H . . . O contacts, the hydrogen–sulfur attraction is stabilized by the charge–quadrupole interaction. Therefore, from the point of view of the photophysical properties of the systems investigated in the present study, the application of the sulfur derivatives of the squaraine dyes can be also vital. This modification is expected to change the order of the excited states of the squaraine moiety by introducing the low-lying $n \rightarrow \pi^*$ transition exploiting the electron pair of sulfur. Hence, the reversed order of the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ with respect to the original oxygen-containing squaraines allow for the stronger intersystem crossing according to the El Sayed rule (El-Sayed (1963)) and so further on, for instance, the enhancement of the singlet oxygen generation quantum yield necessary for the photodynamic therapy applications Peceli et al. (2013); Chetti et al. (2021). It has been shown that simple oxygen-by-sulfur substitution in both places in squaric ring in squaraines drastically modifies neither the one-photon nor the two-photon absorption of the dye. However, it occurs to be a convenient tool for the subtle tuning of the photooptical properties of squaraines and has got even more attention recently in the design of triplet–triplet annihilation systems for photon upconversion Peceli et al. (2013); Pristash et al. (2020); Chetti et al. (2021). The importance of this strategy for the rational design of fluorophores exhibiting bathochromic shift of the absorption and emission bands has been confirmed earlier also for different class of systems (Jedrzejewska et al., 2016).

Therefore, in the present study, the photophysical properties of the sequence of the supramolecular aggregates composed of the squaraine or thiosquaraine central moiety hydrogen bonded by the 2 N,N′-diphenylurea or 2 N,N′-diphenylthiourea (DPTU) derivatives are analyzed with the computational chemistry tools. The aim is to determine the structural modifications which allow tuning the one- or two-photon absorption and singlet–triplet energy gap in oxo- and thiosquaraines. The sequence of the electron-donating and electron-accepting substituents in the 4-position of the phenyl ring of the DPU and DPTU systems has been selected according to the Hammett electronic effects (Hansch et al., 1991). The systematic change of the substituent can be used to tune the intermolecular interactions but also may be used as a research tool to study the given system systematically (Ośmiałowski et al., 2013).

One need to underline that the prototypical dimethylaniline squaraine dye is challenging from the theoretical point of view due to the interplay of several competing effects: the charge transfer between the electron-accepting central squaric acid ring and the electron-donating side groups, the mild biradicaloid character, and the significant double excitations affecting the two-photon theoretical description (Prabhakar et al., 2005b; Prabhakar et al., 2010; Prabhakar et al., 2005a; Yesudas et al., 2006; Srinivas et al., 2007; Ferrer et al., 2019).

TABLE 1 | Hammett constant values σ_p for the sequence of the analyzed substituents (Hansch et al., 1991).

	–	–	–	–	–	–	–	–
Substituent	$-NMe_2$	$-NH_2$	$-NHNH_2$	–OH	–NHOH	$-CMe_3$	–Me	–H
σ_p	–0.83	–0.66	–0.55	–0.37	–0.34	–0.20	–0.17	0.00
Substituent	–Cl	$-CONH_2$	–CHO	$-CCl_3$	$-CF_3$	–COCl	–CN	$-NO_2$
σ_p	0.23	0.36	0.42	0.46	0.54	0.61	0.66	0.78

2 METHODS

The complexes of squaraine (OSQ) and thiosquaraine (SSQ) with 2 N,N′-diphenylurea and N,N′-diphenylthiourea molecules have been investigated for the geometry frozen at the D_{2h} symmetry point group, as presented in **Figure 1**, in order to separate electronic effect from structural deformation. The modification of the acidity of the N–H protons of urea by the proper site substitution or O-by-S replacement is expected to affect both the N–H ... O/N–H ... S distances, electron density distribution ρ, and several other properties such as quadrupole moment, thus influencing intermolecular interaction energy and photophysics of the investigated chromophore. Therefore, the 4-substituted derivatives of urea and thiourea are analyzed. The sequence of the substituents is selected in order to ensure the wide range of electron-donating to electron-accepting properties: $-NMe_2$, $-NH_2$, $-NHNH_2$, –OH, –NHOH, $-CMe_3$, –Me, –H, –Cl, $-CONH_2$, –CHO, $-CCl_3$, $-CF_3$, –COCl, –CN, and $-NO_2$ (Hansch et al., 1991). The corresponding values of σ_p for these substituents are presented in **Table 1**. In order to establish the influence of the hydrogen bonds on the charge transfer inside the squaraine moiety, the artificially symmetrized perpendicular structures have been investigated (**Figure 1**). For these systems, $\mathcal{D}_{2h}$ symmetry has been enforced and partial optimization has been performed in order to allow for the restricted geometry relaxation only. This arises from the fact that the extended π-electron aromatic scaffolds of squaraines prefer the dispersion interaction with the phenyl rings of the urea, when optimized freely. However, in order to limit the considerations to the hydrogen-bonded complexes, which could be observed in crystals, among other interactions, and affect the properties of the analyzed systems significantly, the free relaxation to the stacked architecture has been forbidden in the present study. The ωB97X-D functional has been chosen for its wide applicability and good performance for the molecular and aggregate structures (Chai and Head-Gordon, 2008a,b). The choice of the 6-31+G(d) basis set has been imposed by the presence of polarization and diffuse functions for heavy atoms accompanied by the moderate size of the basis set.

Mutual interactions between the squaraine and urea derivatives have been investigated within the supermolecular approach. The counterpoise-corrected interaction energy $\Delta E_{SM}^{CP}(ABC)$ has been determined for the three-body system according to the site–site counterpoise procedure (Wells and Wilson, 1983; Richard et al., 2018) as follows:

$$\Delta E_{SM}^{CP}(ABC) = E_{ABC}(ABC) - E_{ABC}(A) - E_{ABC}(B) - E_{ABC}(C), \quad (1)$$

where A, B, and C denote one of the subsystems, respectively, while $E_{ABC}(X)$ stands for the energy of the subsystem X = A,B,C calculated in the basis set of the whole complex (basis set indicated in the subscript). The geometry of the subsystems was considered frozen at the complex geometry and no relaxation was allowed.

The non-additivity arising from the three-body effects is estimated as the difference between the total interaction energy in the trimer, $\Delta E_{SM}^{CP}(ABC)$, and the sum of the interaction energies in dimers as follows (Chałasinski et al., 1991; Elrod and Saykally, 1994; Góra et al., 2011):

$$\Delta_{non-add} = \Delta E_{SM}^{CP}(ABC) - \left[\Delta E_{SM}^{CP}(AB) + \Delta E_{SM}^{CP}(AC) + \Delta E_{SM}^{CP}(BC)\right]. \quad (2)$$

All the contributions here are calculated within the full trimer basis set (Řezáč et al., 2015) with ωB97X-D, MP2, and DLPNO-CCSD(T) approaches.

The energy of the hydrogen bonds in the investigated complexes is estimated on the basis of the theory of atoms in molecules according to Espinosa et al. (1998). The analysis of the partial charge distribution has been carried out with the natural population analysis (Foster and Weinhold, 1980; Reed et al., 1988) because of its small dependence on the basis set size, and the quality of the obtained data has been verified with Hirshfeld charges (Hirshfeld, 1977), presented in **Supplementary Material** for OSQ-DPU and SSQ-DPU complexes. The modification of the charge distribution upon excitation has been investigated with the Le Bahers indexes (Bahers et al., 2011) and $\Delta\sigma$ defined as the difference for the root-mean-square deviation of distribution for particle and hole is given in **Supplementary Material**.

Since the bis(N,N-dimethylamine) squaraine molecule represents the challenge for the computational methods due to the double excitation effects balanced with the charge transfer and moderate biradicaloid character, the literature recommendations have been applied for the computational methodology for the electronic excitations (Quartarolo et al., 2009; Alberto et al., 2014; Bassal et al., 2017; Ferrer et al., 2019). It has been proven that the conventional long-range–corrected functionals correctly predict the one-photon absorption spectrum of bis(N,N-dimethylamine) squaraine and despite strong overestimation of the energy of the states involved in two-photon transitions allow the correct assignment of the experimentally observed states. Therefore, the one-photon and two-photon absorption has been described within the Coulomb-attenuated B3LYP functional (Beerepoot et al., 2018). For comparison, OPA parameters were also determined with M06-2X and PBE0 functionals and are provided in **Supplementary Material** (Zhao and Truhlar,

2008; Jacquemin et al., 2012; Leang et al., 2012; Isegawa et al., 2012; Charaf-Eddin et al., 2013; Houari et al., 2014; Azarias et al., 2016).

According to the selection rules for OPA, the transitions in the centrosymmetric systems are possible only with the change of the symmetry of the state (from gerade to ungerade or from ungerade to gerade), in opposition to the TPA-allowed electric dipole transitions, which require the preservation of the symmetry of the state. The oscillator strength f_{if} corresponding to the probability of the one-photon absorption in the electric dipole approximation depends on the energy difference between the involved initial and final states E_f and E_i and the transition dipole moment as follows:

$$f_{if} = \frac{2}{3}\frac{m_e}{\hbar^2}\left(E_f - E_i\right)\sum_{\alpha=x,y,z}|\langle\psi_i|R_\alpha|\psi_f\rangle|^2, \quad (3)$$

where m_e denotes the electron mass, $\hbar$ is the reduced Planck constant, and the sum runs over the Cartesian coordinates.

The intersystem crossing rate depends on the spin–orbit coupling matrix elements and is inversely proportional to the singlet–triplet energy gap E_T–E_S for the involved states. Due to the artificial geometry of the investigated complexes, only the singlet–triplet energy gap is calculated in the present study for the estimation of the influence of the O-by-S replacement and substituent effect in hydrogen-bonded systems on the intersystem crossing.

The two-photon absorption cross section σ^{TPA} is defined in GCS units for the photon energy $\hbar\omega$ as follows:

$$\sigma^{TPA} = \frac{4\pi^3\alpha a_0^5\omega^2}{c}\langle\delta^{TPA}\rangle g(2\omega, \omega_0, \Gamma), \quad (4)$$

where δ^{TPA} denotes the TPA strength, calculated for the linearly polarized light as follows:

$$\langle\delta^{TPA}\rangle = \frac{1}{15}\sum_{ab}\left(2S_{ab}\bar{S}_{ab} + S_{aa}\bar{S}_{bb}\right), \quad (5)$$

with S being the TPA transition moment defined in Beerepoot et al. (2015), and a and b denoting the Cartesian components. $g(2\omega, \omega_0, \Gamma)$ stands for the lineshape function responsible for inclusion of spectral broadening effects, α is the fine structure constant, and a_0, the Bohr radius.

Geometry optimization, supermolecular interaction energy, and one-photon absorption calculations have been performed with the Gaussian16 package (Frisch et al., 2016). Two-photon absorption cross sections have been determined with the Dalton2018 program package (Aidas et al., 2018; Aidas et al., 2014). The non-covalent interaction analysis has been performed within NCIPlot software (Boto et al., 2020; Johnson et al., 2010; Contreras-Garcia et al., 2011) and AIMAll (Keith, 2019; Bader, 1990; Popelier, 2000), and charge transfer indexes have been calculated with Multiwfn3.8 (Lu and Chen, 2021. LPNO-CCSD(T) counterpoise-corrected interaction energies have been obtained within the ORCA 4.2.1 package (Neese et al., 2020a; Neese et al., 2020b; Neese, 2012; Neese, 2018; Altun et al., 2021).

In all the graphics along the article, blue lines and points denote the oxosquaraine systems, while the green ones cover the thiosquaraines. The darkest shade is applied for the isolated (thio) squaraine, the middle shade for the DPU complexes, while the lightest shade for the DPTU complexes.

3 RESULTS

3.1 Structural Properties

Intermolecular interactions of squaraine moiety with the N,N′-diphenyl(thio)urea derivatives affect both the structure of the squaraine scaffold itself and the geometrical parameters of the D–H … A contacts. **Figures 2**, **3** present the corresponding modification of the bond length in the (thio)squaraine. It can be noticed that the most pronounced modification along the squaraine symmetry axis occurs for the C5–N7 bond (compare **Figure 1** for atom numbering), connecting the squaraine phenylene moiety with the -NMe_2 terminal substituent (**Figure 2A**). The sequence of the applied urea substituents causes the C5–N7 bond shortening with the growing electron-withdrawing character of the substituent by about 0.004 Å. The corresponding modification in the case of C15–C19 (and equivalent) bonds in the phenylene rings does not exceed 0.002 Å, and for C1–C3 bond, this adjustment remains of the order of 0.003 Å only. These tendencies are independent on the sulfur present either in squaraine itself or in urea. Although small, these differences indicate the stronger bond localization in the systems containing strong EW groups in urea in opposition to the delocalization upon the substitution with strong ED, when the bond length alternation becomes smaller.

On the other hand, the double bond between the C9 carbon in the squaric ring and the O11 oxygen or S11 sulfur atom in OSQ and SSQ, respectively, undergoes modest elongation upon the introduction of the terminal EW group in urea (not exceeding 0.004 Å). Again this elongation is almost not dependent on the sulfur presence either in the urea or in the squaraine. This is the first geometrical indication that the urea substitution with EW groups promotes the O/S … H–N hydrogen bonds shortening and thus strengthening. This observation remains in agreement with the O11/S11 … H (urea) and N (urea)–H (urea) distance modification upon urea substitution, as shown in **Figure 4**.

Since the size of the sulfur atom is significantly larger than the oxygen one, the scale on both plots is shifted by 0.6 Å for convenience. The O-by-S replacement in squaraine (for urea complexes compare dark blue and dark green lines in upper panels of **Figure 4**) directly affects the hydrogen bond formed between the subsystems. The sulfur presence in the hydrogen bond makes it more prone to the electronic effects arising from the modification of the substitution pattern, and the hydrogen bond donor … proton distance decreases strongly with the growth of the Hammett substituent constant compared with that in the case of the O … H–N interactions in oxygen-bearing squaraine. This tendency is similar for the influence of the O-by-S replacement in squaraine on the hydrogen bond distance in the thiourea complexes, which can be deduced from the comparable difference of the slopes (see light blue and light green lines in upper panels of **Figure 4**). The strong electron-withdrawing substituent present in the urea units may

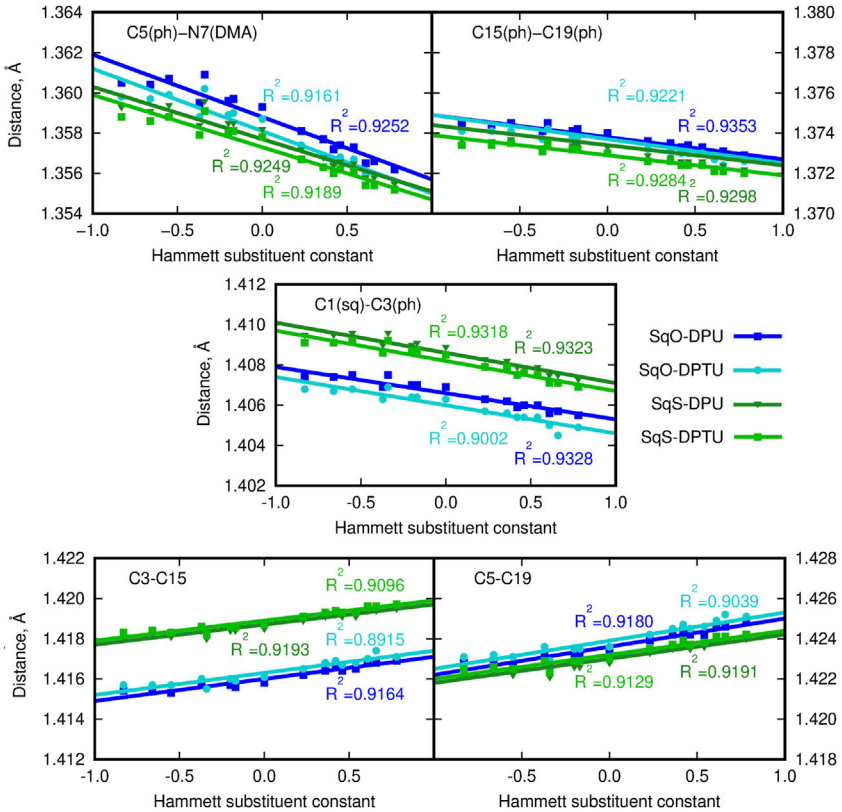


FIGURE 2 | Geometrical parameters for squaraine derivatives (atom numbering in **Figure 1B**; here and in the article; blue lines/points denote the oxygen-containing squaraine complexes and green lines/points, its thioderivative; while darker lines are used for urea and lighter for thiourea). (DMA) denotes the terminal dimethylamino group and (ph) stands for the phenyl ring in (thio)squaraine.

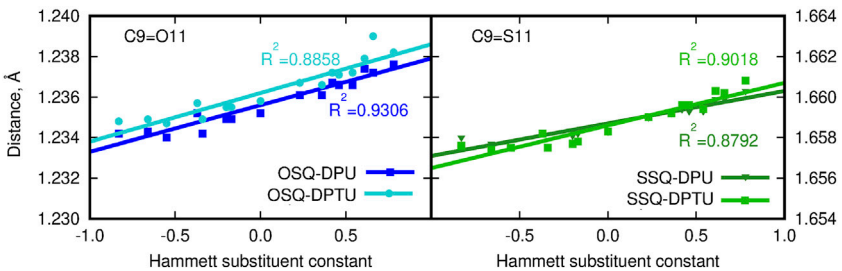


FIGURE 3 | C9-O/S11 bond length for squaraine derivatives (atom numbering in **Figure 1B**; blue lines/points denote the oxygen-containing squaraine complexes and green lines/points, its thioderivative, while darker lines are used for urea and lighter for thiourea).

cause the shortening of the O . . . H distance with respect to the most electron-donating group by as much as 0.05 Å in the OSQ complexes and 0.09 Å in SSQ complexes. The effects observed here can be diminished by the considerable size of the squaraine moiety, which prevents the effective substantial intramolecular charge transfer upon the urea substitution. The corresponding modifications for the O . . . H distance in the formaldehyde–urea and thioformaldehyde–urea complexes are significantly larger, reaching 0.157 9 Å (about 5.5% of the starting value) for the thio derivative of formaldehyde. Therefore, this simple geometrical analysis indicates that the proper design of the supramolecular complexes bearing the chromophore unit allows for the subtle tuning of the electron charge density distribution and thus all of the resulting features.

3.2 Charge Distribution in Ground State

The presence of the ED or EW substituents in the urea derivatives is expected to modify the charge distribution in the urea molecules itself. However it could be also assumed that such an adjustment of the electron density in the hydrogen

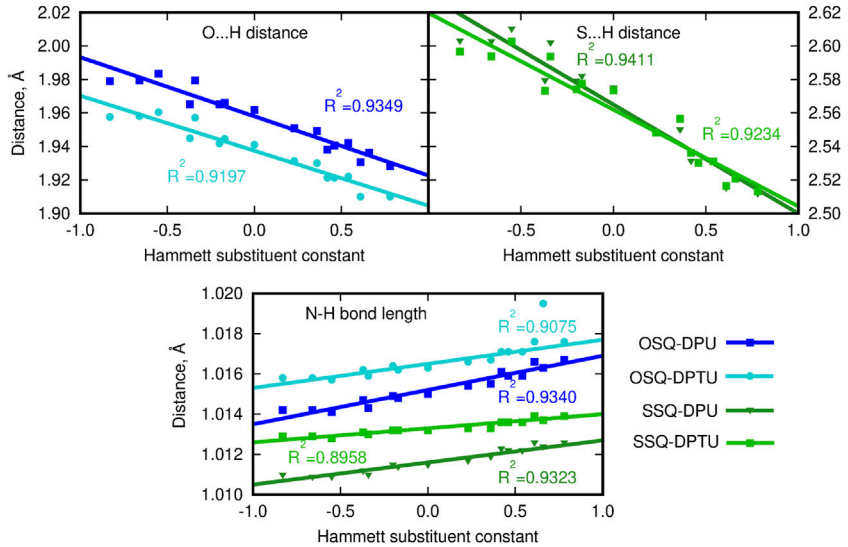


FIGURE 4 | Geometrical parameters for hydrogen bonds (A . . . H and D–A distances, where A denotes the hydrogen bond acceptor and D, hydrogen bond donor).

bond–donating species is strong enough to affect the hydrogen bond, and thus hydrogen bond acceptor and its properties as well. Therefore the charge distribution in the squaraine moiety is presented with respect to the Hammett constant in the substituted N,N′-diphenyl(thio)urea (**Figure 5**). The sum of NBO charges for the (thio)squaraine, presented in panel (D) of **Figure 5**, indicates that the interaction in the considered hydrogen-bonded complex can be perceived as a way of drawing the electrons from the squaraine unit. This effect is almost twice as strong in the oxygen-bearing squaraines as for their thio-counterparts. Additionally, the stronger the electron-accepting character of the substituents in N,N′-diphenyl(thio)urea, the more positive the partial charge in squaraine that remains in agreement with chemical intuition. The O-by-S exchange in urea on the other hand affects partial charge in (thio)squaraine only marginally.

The central squaraine system is divided into three parts [panels (A–C) of **Figure 5**, respectively], according to the character of these parts, namely, the terminal N,N–dimethylamino group serves as the strong electron donor and the squaric ring in the center is an electron-withdrawing moiety, while phenyl rings provide the scaffold enabling communication between those moieties. Introduction of the electron-donating substituents in the N,N′-diphenyl(thio)urea causes the decrease in the negative charge cumulated in the central squaric ring, while the electron-accepting groups present in the N,N′-diphenyl(thio)urea leads to the increase of the negative charge in this part of the system. The opposite trend can be noticed for the terminal N,N-dimethylamino group, which gather the negative charge in the case of the ED group introduced in the N,N′-diphenyl(thio)urea and give it back to the squaric ring in the case of EW substituent. Moreover, at the N,N-dimethylamino groups the charge redistribution in squaraine moiety seems to be most spectacular, since the charge increases by about 50% with the growth of the Hammett constant, depending on the system [compare panel (B) in **Figure 5**]. The O-by-S substitution both in squaraine and in urea units does not affect the general tendencies for the charge distribution particularly for the terminal N,N-dimethylamino groups. In the case of the central squaric ring, the sulfur presence in the thiosquaraine strongly increases the negative charge gathered by the squaric ring in comparison to the pristine oxosquaraine, while the O-by-S exchange in urea is noticed also to increase the negative charge accumulated in the central part of the complex, but only tiny effect is observed. On the other hand, for the thiosquaraine, the presence of the sulfur in the thiourea moiety does not introduce any further meaningful modifications. Additionally, the Le Bahers charge transfer indexes show that due to the centosymmetric architecture of the investigated systems, the general shift of the charge distribution upon excitation is equal to zero. Moreover, the $\Delta\sigma$ parameter indicates that the effect is tiny and smaller for the thiosquaraine complexes than for the oxosquaraine ones.

3.3 Substituent Effects in the Interaction Energy

The dependence of the supermolecular interaction energy on the Hammett constant of the substituent in urea derivative, estimated within the DFT, MP2, and DLPNO-CCSD(T) approaches, is presented in **Figure 7**. The blue lines denote the OSQ complexes, while the green ones, the SSQ systems. One can see that the general tendencies are independent on the methodology applied and hold for all of the employed methods. First of all, the interaction is twice as large for the pristine, oxygen-containing squaraines as for their thio-analogs, and this difference is larger than 15 kcal/mol for all analyzed cases. This can be expected due to the known larger strength of the N–H . . . O hydrogen bonds than the N–H . . . S ones. This

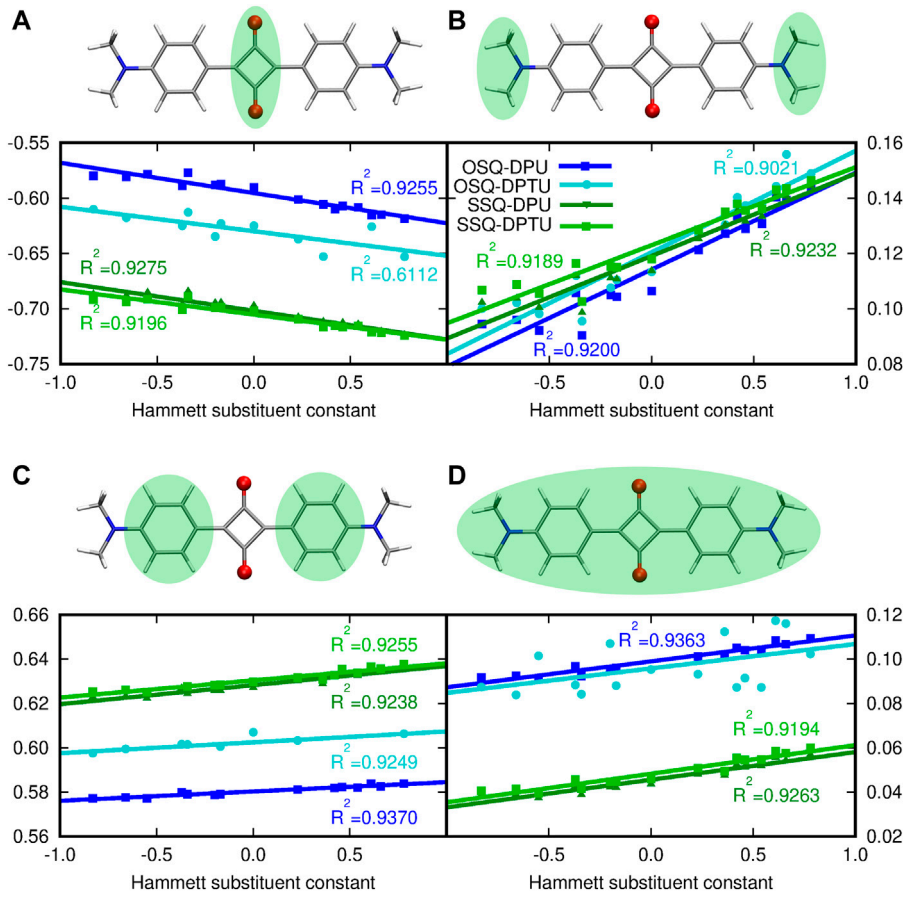


FIGURE 5 | Modification of the NPA partial charge distribution in the central (thio)squaraine unit upon substitution in N,N′-diphenyl(thio)urea fragments: **(A)** in the central squaric ring, **(B)** in the terminal N,N-dimethylamino groups, **(C)** in phenyl rings, and **(D)** in the whole (thio)squaraine moiety. Sum of partial charges from the considered fragment, presented in the green background, is given as an ordinate.

trend is confirmed by the hydrogen bond strength, E_{HB}, estimated according to Espinosa et al. (1998) as the negative of the half of the potential energy density (virial field) within AIM (Bader, 1990; Popelier, 2000) (see **Figure 6** and **Supplementary Material**). E_{HB} is roughly three times larger for the oxygen-containing squaraines than for the thiosquaraines (lower panel in **Figure 6**). On the other hand, the O-by-S replacement in the diphenylurea species leads only to the tiny enhancement of the attraction. The introduction of sulfur atoms to the urea derivatives amplifies the H-bonds up to 10% that arises from the larger acidity of the N–H protons in thiourea than in urea. This makes the difference of the order of 0.5 kcal/mol for oxosquaraine complexes and less than 0.1 kcal/mol for thiosquaraine ones. All of the lines plotted in panels of **Figure 7** remain almost parallel, which means that the substituent effect is the same in both series. However, the stronger the electron-accepting character of the substituent is, the interaction energy becomes more attractive, since the urea withdraws the electrons from the central squaraine moiety, thus making the hydrogen bond stronger. These effects are slightly more pronounced for oxygen-bearing squaraines (attraction enhancement by about 10 kcal/mol) than for their thio-analogs (7–8 kcal/mol).

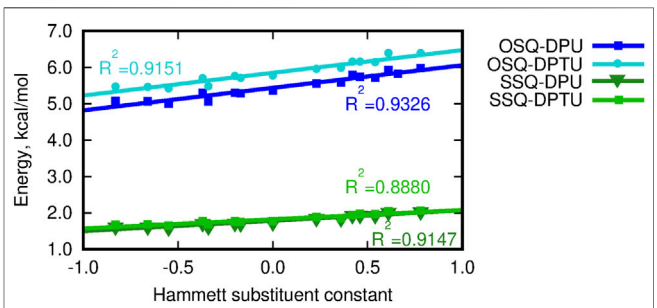


FIGURE 6 | Hydrogen bond energy estimated according to Espinosa et al. (1998) versus the Hammett substituent constant.

It is interesting to investigate to what extent the interactions in OSQ and SSQ complexes with DPU originate from extra interaction, beyond pairwise molecule–molecule interactions. In other words, it is of importance if there is any cooperative

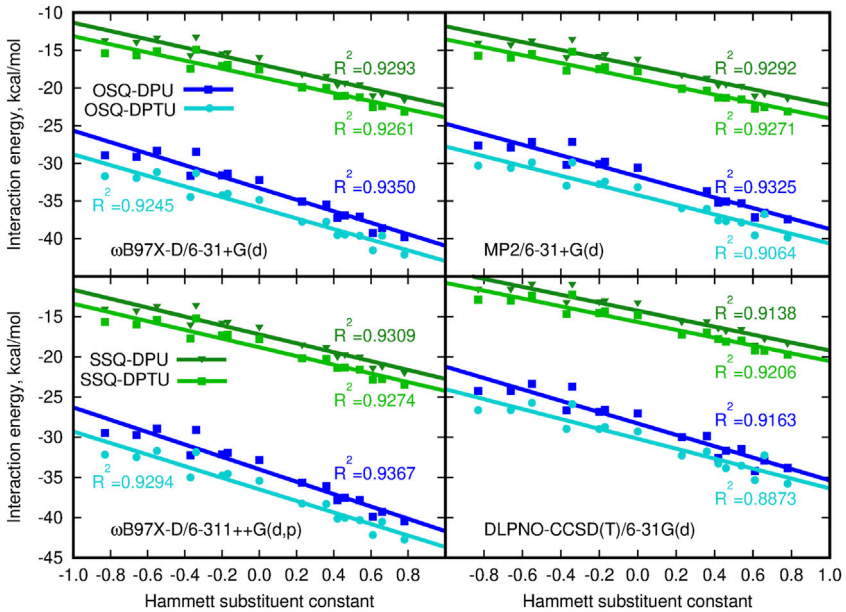


FIGURE 7 | Counterpoise-corrected supermolecular interaction energy for the three-component system calculated according to the site–site counterpoise scheme versus the Hammett substituent constant.

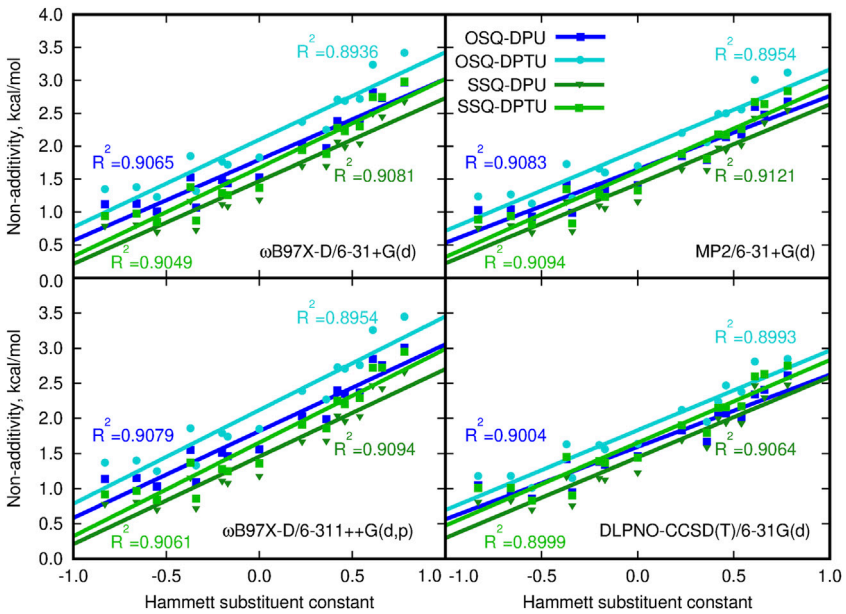


FIGURE 8 | Non-additivity in the supermolecular interaction energy versus the Hammett substituent constant.

effect in the complex which cannot be reproduced merely by adding SQ molecules and DPU units. The significance of the three-body effect for the total interaction is presented in **Figure 8**. The observed effects are destabilizing: inclusion of the third subsystem in the complex decreases the mutual attraction observed when only two-body terms are taken into account. The non-additivity is of similar order of magnitude for all of the investigated series of systems and clearly exhibits the dependence on the substituent character. These effects equal to about 1 kcal/mol for the most electron-donating N,N-dimethylamino group and increasing to about 2.5–3.5 kcal/mol for the most electron-withdrawing nitro group and are slightly higher for the oxosquaraine than for its thio-counterpart. It should be also noticed that beside the enhancement of the absolute value of the three-body effects, its percentage into the total supermolecular interaction energy is doubled within the

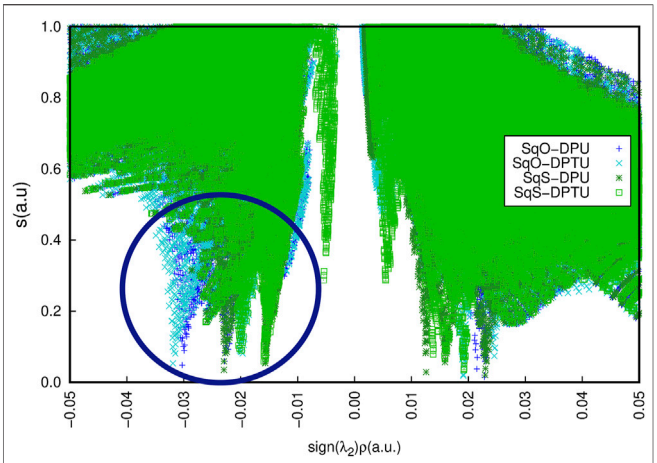


FIGURE 9 | Non-covalent interactions for the unsubstituted systems according to the NCIPlot analysis for OSQ-DPU (blue), OSQ-DPTU (turquoise), SSQ-DPU (dark green), and SSQ-DPTU (light green).

analyzed substituent series for all of the oxo- and thio-complexes. However here, due to the significantly weaker interaction in thiosquaraine systems than in oxosquaraine ones, the overall percentage of the three-body effects is larger for squaraine-containing sulfur atoms and reaches from 5.5% for -NMe_2 substituent to 12.5% for -NO_2. Notwithstanding, the magnitude of three-body effects in these complexes is relatively small, which allows for simplified treatment of bigger analogs of these molecules.

Additionally, the tiny difference for OSQ and SSQ complexes in the character of the interactions can be noticed by the NCI analysis (**Figure 9**). The difference observed in the range of −0.035 to −0.02 a. u. (marked by the navy oval) concerns the character of the contacts responsible for the hydrogen bonds.

3.4 Substituent Effect on the Orbital Energy Levels

Since the introduction of the sulfur into the system is known to insert additional energy levels arising from the lone pairs of sulfur, it is expected to provide a desired tool for tailoring of the chromophores properties. Thus, the sulfur presence in the squaric ring leads to the inversion of the lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ excitations in this manner decreasing the singlet–triplet energy gap (Webster et al., 2010; Avirah et al., 2012; Peceli et al., 2013). This can be a way of increasing the intersystem crossing for the efficient application of squaraine triplet states, for instance, in photodynamic therapy or in phosphorescent molecular devices.

Figure 10 presents the frontier orbital energy levels for the isolated squaraine and thiosquaraine in comparison to their N,N′-diphenyl(thio)urea complexes. One needs to carefully pay attention to the orbitals involved in the most intensive long-range singlet–singlet transition—not for all of these systems, these are simply HOMO and LUMO orbitals, as pointed out by dashed lines in **Figure 10**. However, in all of the investigated systems, the type of the orbitals involved in excitations is preserved: the transition occurs from and to the orbitals localized mainly or even exclusively in the (thio)squaraine moiety (compare **Figure 12**). In the case of the thiosquaraines, additionally the lone pairs on the sulfur are involved. Although their contribution to the excitation is minor, the energy difference between this lone pair and LUMO undergo the uplifting upon the (thio)urea substitution even by one order of magnitude more that for the corresponding $\pi \rightarrow \pi^*$ transitions (0.003 49 vs. 0.000 32 Hartree for SSQ-DPU complexes, compare **Figure 11**). The highest occupied molecular orbitals for the -NMe_2 substituted systems correspond to the N,N′-diphenyl(thio)urea π orbitals involved in the transitions in the UV range below 300 nm (see **Figure 12**). Taking the orbital energy difference as the first approach to the electronic excitation spectrum, one can clearly foresee the differences between the oxosquaraine and thiosquaraine complexes absorption upon introduction of EW/ED groups in the N,N′-diphenyl(thio)urea.

3.5 Substituent Effect on the One-Photon Absorption

One-photon absorption for the investigated complexes is expected to undergo modifications arising from the introduction of the heavy (sulfur) atom and/or EW/ED substituents due to the changes in the orbital picture presented in the previous section. The electronic absorption spectrum of thiosquaraine differs from the oxygenated squaraine by the broader and red-shifted lower intensity absorption band in the long wavelength range (582 nm for isolated thiosquaraine vs. 500 nm for oxosquaraine, as presented in **Figure 12**), and the signal at 303 nm observed for isolated oxosquaraine is shifted to 377 nm for isolated thiosquaraine (compare **Supplementary Material**).

Figure 13 depicts the general tendencies in the OPA spectrum for investigated complexes upon substitution [panels (A–C)] and the comparison of the (thio)squaraine-in-complex with the isolated (thio)squaraine [panel (D)]. The maximum absorption wavelength in complex is shifted with respect to the isolated species by only tiny margin by 7–11 nm for oxosquaraines and significantly by almost 50–70 nm for thiosquaraines [**Figure 13**, panel (D)]. Clearly, the spectrum modification with respect to the isolated (thio)squaraine upon N,N′-diphenyl(thio)urea substitution is also almost constant for oxosquaraines. On the other hand, for thiosquaraines, the maximum absorption wavelength shift depends strongly on the introduction of the substituents in the urea. Thus, the total effect of the spectrum adjustment upon interactions can be roughly divided into two components: one arising from the complex formation by hydrogen bonds to unsubstituted urea derivatives and the other from the substitution of the urea phenyl rings. The influence of the complex formation *via* hydrogen bonds with the unsubstituted N,N′-diphenyl(thio)urea causes the bathochromic shift of absorption spectrum by 9–10 nm for oxosquaraine and by 54.5–56 nm for thiosquaraine. Additionally, upon hydrogen bonding, a gentle increase of the band intensity is noticed. Further substitution of urea units leads to the additional shifts on less than 2 nm for oxosquaraines and up to 12 nm for thiosquaraines. Therefore, the O-by-S replacement in squaraine is expected to provide a way of

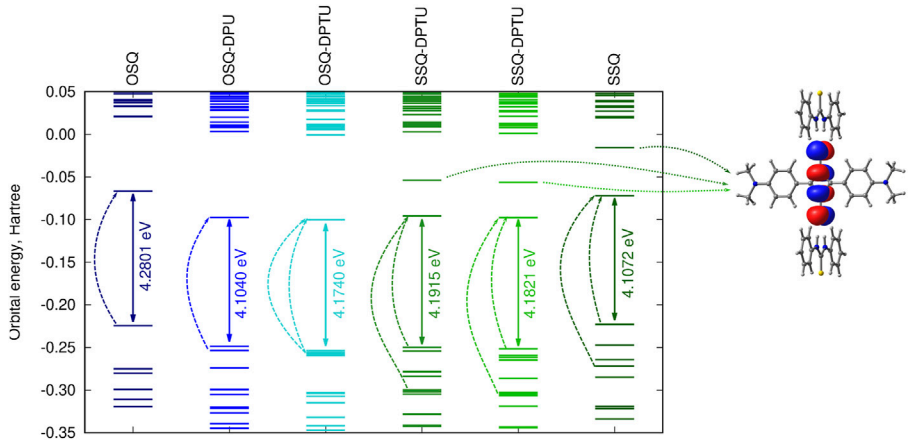


FIGURE 10 | Orbital energy levels [CAM-B3LYP/6-31+G(d)] for the investigated unsubstituted systems together with the LUMO+1 orbital for thiosquaraine (dashed lines indicate the orbitals involved in the most intensive one-photon singlet–singlet transitions; for the corresponding orbital shapes compare **Figure 12**).

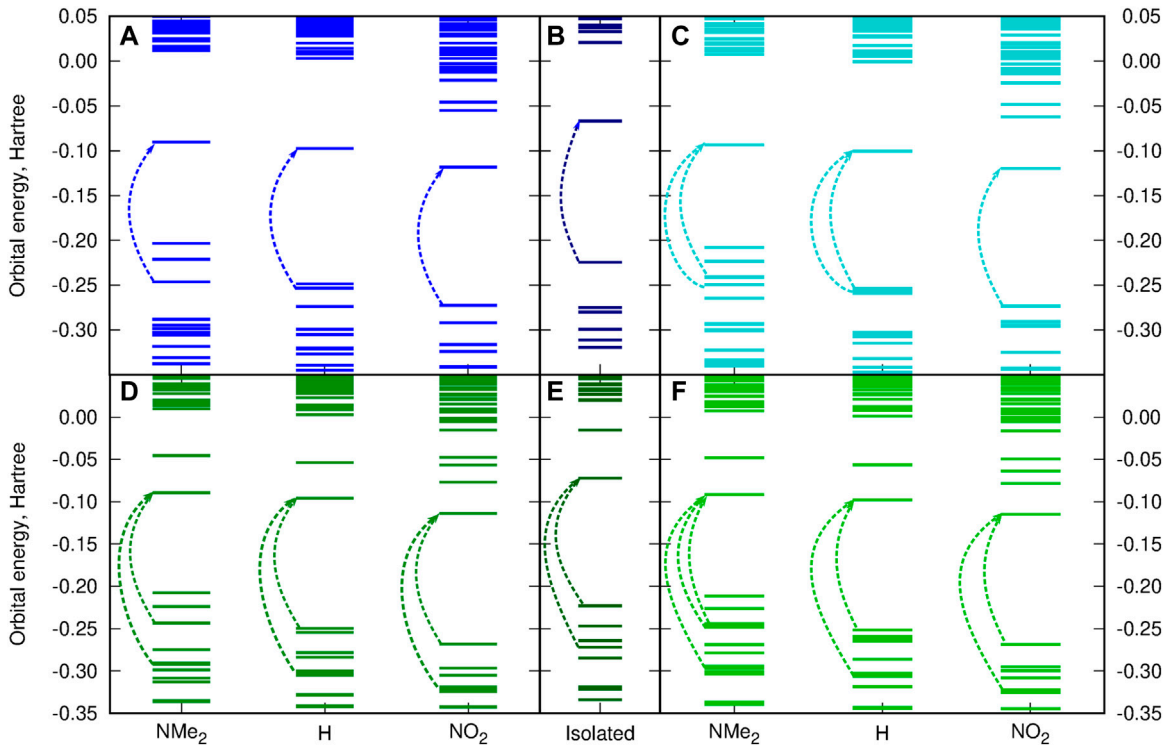


FIGURE 11 | Substitution effect on the orbital energy levels [CAM-B3LYP/6-31+G(d)] for the investigated systems in comparison to the isolated (thio)squaraine **(A)** OSQ-DPU, **(B)** isolated OSQ, **(C)** OSQ-DPTU, **(D)** SSQ-DPU, **(E)** isolated SSQ, and **(F)** SSQ-DPTU (dashed arrows indicate orbitals involved in most intensive long-wavelength transitions).

achieving the strong bathochromic shift of absorption spectrum, the formation of hydrogen-bonded complexes particularly in the case of thiosquaraine produces also the significant red-shift, and the tiny tailoring of the absorption spectra can be reached by the proper substitution on the phenyl rings of urea. Moreover, the thionylation of the diphenylurea is negligible both for oxosquaraine and thiosquaraine absorption spectrum.

The stronger shift in the thiosquaraine than for oxosquaraine can be ascribed to the lone-pair sulfur orbitals more prone to the influence made by the EW/ED substituents than the regular π orbitals involved in the excitations in the case of oxosquaraine.

Moreover, it should be noticed that the O-by-S replacement in squaric ring also affects the oscillator strength *f*, determining the probability of the transition [compare panel (B) of **Figure 13**]. The

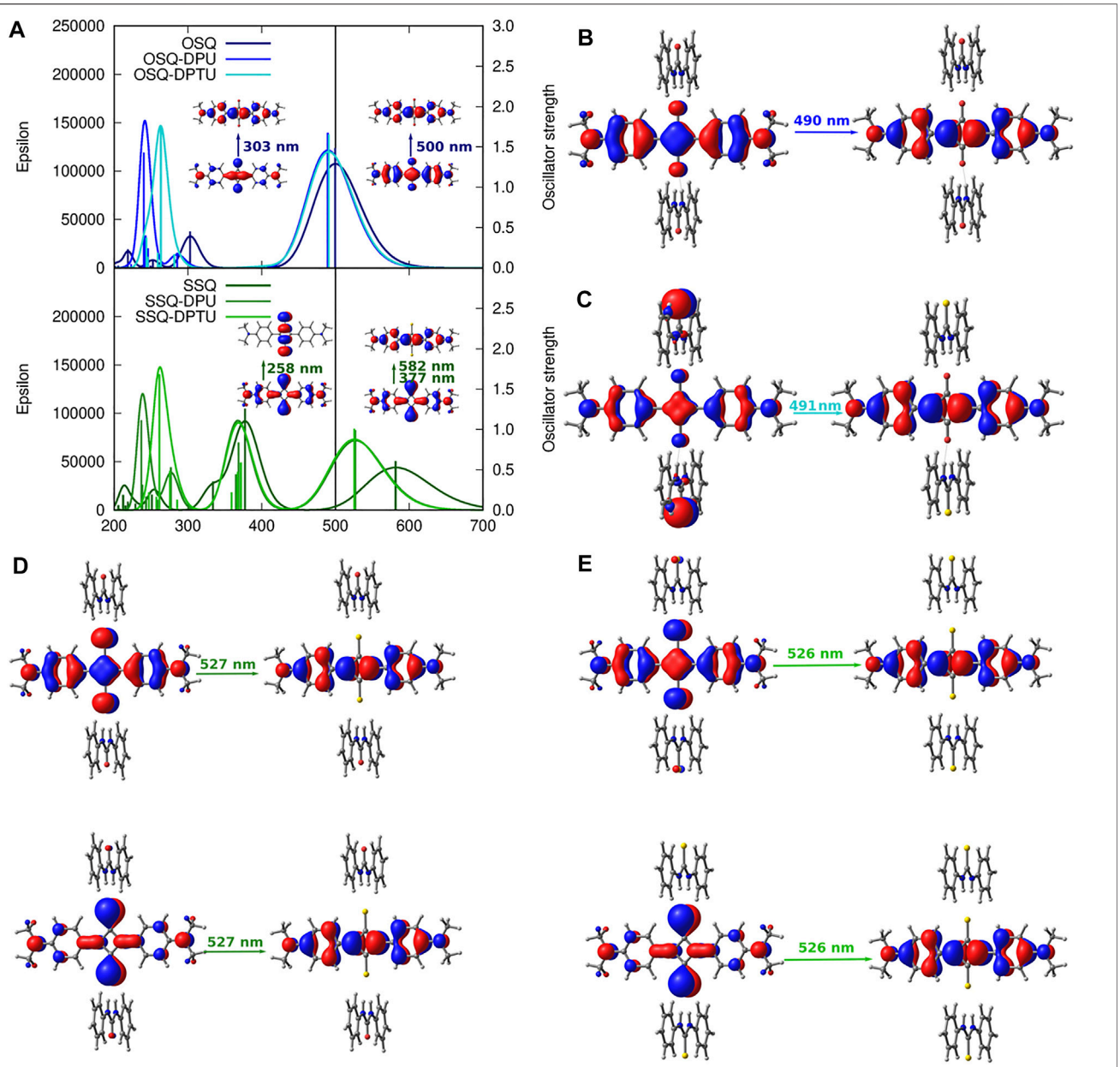


FIGURE 12 | One-photon absorption [CAM-B3LYP/6-31+G(d)] together with the corresponding orbitals for investigated systems: **(A)** vertical absorption spectrum for oxosquaraine and its complexes **(upper panel)** and thiosqaraine **(lower panel)**, **(B)** OSQ-DPU frontier orbitals, **(C)** OSQ-DPTU frontier orbitals, **(D)** SSQ-DPU frontier orbitals, and **(E)** SSQ-DPTU frontier orbitals (convolution of the theoretical oscillator strengths with Gaussian function for full-with half-maximum equal to 3,000 cm^{-1}).

complexes of thiosquaraine exhibit significantly higher values of *f*, additionally increasing with the growing substituent Hammett constant for the groups present in N,N′-diphenyl(thio)urea units.

Since the sulfur presence in the squaric ring is known to invert the ordering of the states and reduce the singlet–triplet energy gap, the vertical triplet state energy and corresponding singlet–triplet energy gaps have been determined for the considered systems and depicted in **Figure 14** (compare also **Supplementary Material**). The singlet excited state for the thiosquaraine complexes is characterized by the lower energy than for the oxosquaraine ones, and the tendency observed for the triplet states is opposite: the O-by-S replacement in the squaric ring causes the elevation of the excitation energy. This remains in agreement with the expectations based on the literature reports (Webster et al., 2010; Avirah et al., 2012; Peceli et al., 2013; Pristash et al., 2020; Chetti et al., 2021). One should also notice that the excitation energy for the oxosquaraine systems both of singlet and triplet states persists almost constant independently on the substitution pattern in N,N′-diphenyl(thio)urea units, while in the case of the thiosquaraine complexes, the growing

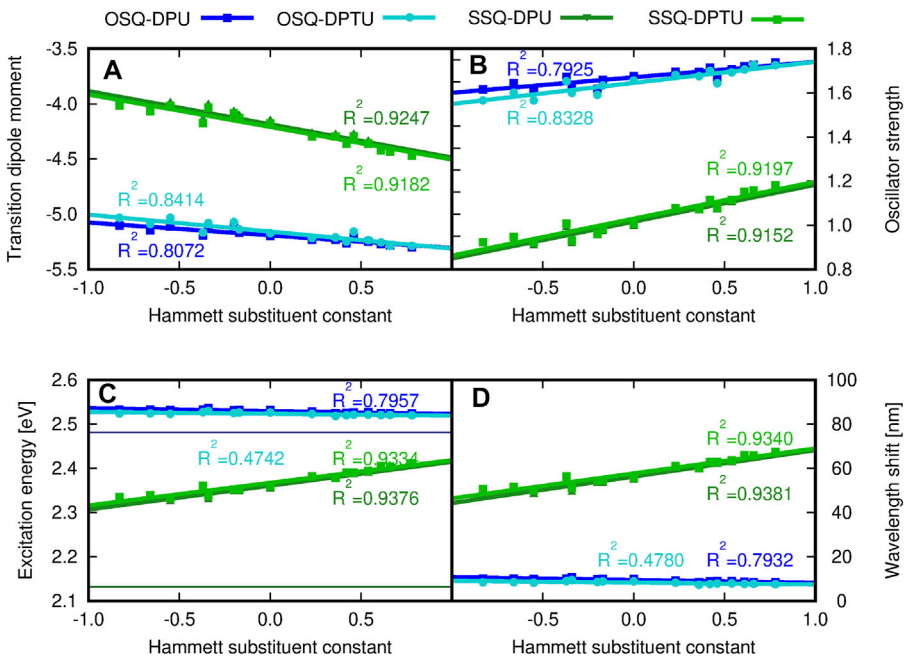


FIGURE 13 | (A) Transition dipole moment component, **(B)** oscillator strengths, **(C)** excitation energy (eV), and **(D)** maximum wavelength shift (nm) with respect to the isolated oxo- or thiosquaraine spectrum, respectively, for one-photon absorption in the investigated complexes versus the Hammett substituent constant for substituents in N,N′-diphenyl(thio)urea estimated within the CAM-B3LYP/6-31+G(d) approach. The reference excitation energy for isolated squaraines is shown for comparison in panel **(C)** as a thin navy horizontal line for oxosquaraine and thin dark green horizontal line for thiosquaraine.

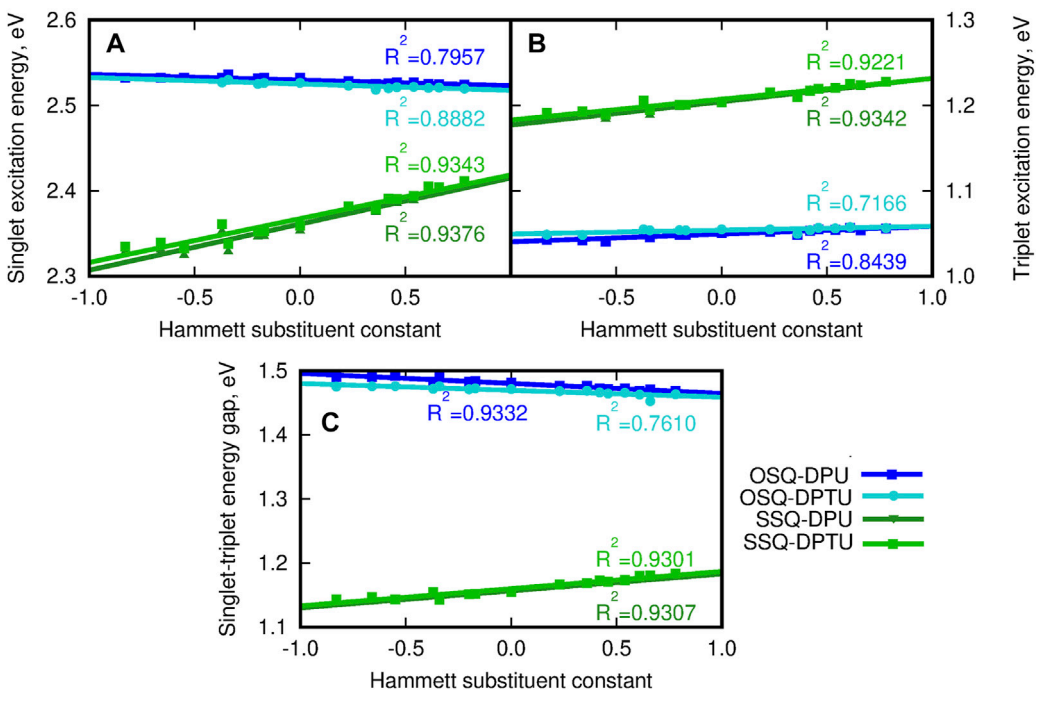


FIGURE 14 | (A) Vertical singlet excitation energy (eV), **(B)** vertical triplet excitation energy (eV), **(C)** singlet–triplet energy gap (eV) for the investigated complexes versus the Hammett substituent constant for substituents in N,N′-diphenyl(thio)urea estimated within the CAM-B3LYP/6-31+G(d) approach. The reference triplet excitation energies for isolated squaraines are equal 0.805 4 eV and 1.028 0 eV, respectively, for oxo- and thiosquaraine; for the reference singlet excitation energy see panel **(C)** of **Figure 13**.

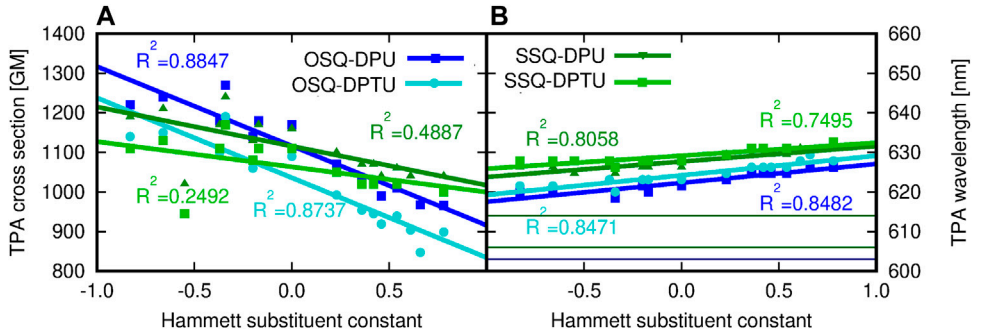


FIGURE 15 | (A) Two-photon absorption cross section (GM) and **(B)** two-photon absorption wavelength (nm) for the investigated complexes versus the Hammett substituent constant, determined within the CAM-B3LYP/6-31+G(d) approach (for comparison, the TPA wavelength of isolated species are equal to 603 and 614 nm for OSQ and SSQ, respectively).

electron-withdrawing character of the substituent leads to the stronger increase of the excitation energy for singlet and triplet states. Thus, the urea substitution can be perceived also as the way of fine modification of the singlet–triplet energy gap for hydrogen-bonded thiosquaraines and—in this way—the tool to manipulate the probability of the intersystem crossing.

3.6 Two-Photon Absorption

Comparison of the reference values for the isolated oxo- and thiosqaraine indicates the bathochromic shift of the TPA wavelength by 11 nm upon O-by-S replacement in the squaric ring accompanied by the slight decrease of the high TPA cross section which remains in agreement with literature reports considering similar oxo- and thiosquaraine derivatives Webster et al. (2010); Avirah et al. (2012). However, for both isolated systems, the σ^{TPA} values are still tremendous and this makes them interesting agents for TPA applications. Therefore, the possibility of the fine-tuning of the TPA wavelength together with controlling the intensity of the signals remains vital. Hydrogen bonding of squaraine derivatives with N,N′-diphenylurea influences its TPA, as it is illustrated in **Figure 15**. The hydrogen bond leads to the shift of 2λ to 622 nm for oxosquaraine (by 19 nm) and 614–626 nm (by 12 nm) for thiosquaraine, thus shifting these values more into the therapeutic window. Such supramolecular architecture involving the hydrogen bonds additionally strongly diminishes the signal intensity (namely, TPA cross section from 2010 to 1090 GM for oxosquaraine and from 1540 to 1160 GM for thiosquaraine); however, these values are still large and promising for applications. The TPA wavelength (2λ) is red-shifted with the increasing electron-withdrawing character of the diphenylurea substituent, and the shift is mild both for oxosquaraine (6 nm: from 620 nm for NMe_2 derivative by 622 nm for unsubstituted diphenylurea to 626 nm for NO_2 derivative) and for thiosquaraine (5 nm: 626 nm, 626 and 631 nm, respectively). Additionally, the intensity of the TPA increases with the growing electron-donating character of the substituent from 966 GM for nitro-diphenylurea to 1220 GM for dimethyloamino-diphenylurea derivative hydrogen bonded to oxosquaraine. The striking difference in the substituent influence on the OPA and TPA of oxosquaraine can be noticed in comparison to the thiosquaraine case. For oxosquaraine, modification of the EW/ED character of N,N′-diphenyl(thio)urea substituent in fact does not affect the singlet excitation energy (λ shift smaller than 2 nm), while the TPA wavelength (2λ) is shifted more than 6 nm. These effects remain tiny, however, opposite to those observed thiosquaraine complexes, where OPA wavelength can be modified as much as by 17 nm, and the effect of substitution on TPA wavelength becomes smaller than for oxosquaraine systems (below 5 nm).

From **Figure 15**, it may be concluded that the oxygen in squaraine is preferred when one need to tune the TPA cross-sectional value by hydrogen bonding. That is caused not only by higher electronegativity of the oxygen atom and higher CT character toward the center of OSQ than that in SSQ but also by the stronger H-bonding of oxygen as an acceptor of hydrogen bond. On the other hand the tuning of the position of the absorption can be realized by both compounds with preference for sulfur if one aims at the red-shifted spectra. Final remark and the challenge for synthesis are to obtain the monooxo derivatives containing one O and one S atom.

4 DISCUSSION

The present study is devoted to the development of the controlled modifications of photophysical properties of model squaraine dye. The investigated features such as one-photon absorption, two-photon absorption, or singlet–triplet energy gap affecting the intersystem crossing efficiency are shown to be prone to the fine-tuning by O-by-S replacement in the squaric ring of the chromophore by symmetric hydrogen bonding, by N,N′-diphenyl(thio)urea, and even by systematic modification of the character of substituents present in this N,N′-diphenyl-(thio)urea units. Meaningful differences are observed yet for isolated oxo- and thiosquaraines, namely, the strong bathochromic shift of one-photon absorption or the singlet–triplet energy gap arising from the reversed order of the orbitals involved due to the presence of sulfur, which has been reported earlier. However, this O-by-S replacement is shown to also affect the susceptibility of the central squaraine moiety to supramolecular adjustments arising from hydrogen bonding. Thiosquaraines exhibit significant one-photon absorption blueshift upon hydrogen

bonding with the growing electron-withdrawing character of the substituent present in position 4 of N,N′-diphenyl(thio)urea units, in opposition to the weak effects observed for oxosquaraines. On the other hand, the substituent effect in the hydrogen bond donor is similar in case of two-photon absorption for both oxo- and thiosquaraines. A strong reduction of singlet–triplet gap upon O-by-S replacement in squaraines can be further adjusted by hydrogen bonding and introduction of electron-donating substituents in urea. Thus, the present study provides a controlled way of modifying photophysical properties of squaraine dyes by the series of factors (O-by-S replacement, hydrogen bonding, substituent effects in hydrogen bond donor) and indicates the strong and gentle outcomes which allow to precisely adjust the desired features of system of interest, when properly combined.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

BO, PŻ, and AK-K contributed to conception and design of the study. AK-K and DK performed all the calculations and data analysis. AK-K wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

Wrocław Supercomputing and Networking Center is gratefully acknowledged for the generous allotment of computational resources. PZ and DK acknowledge the National Science Center for support (Sonata Bis 9, Grant No. 2019/34/E/ST4/00 407).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.800541/full#supplementary-material

REFERENCES

Aidas, K., Angeli, C., Bak, K. L., Bakken, V., Bast, R., and Boman, L. (2014). The Dalton Quantum Chemistry Program System. *Wires Comput. Mol. Sci.* 4, 269–284. doi:10.1002/wcms.1172

Aidas, K., Angeli, C., Bak, K. L., Bakken, V., Bast, R., Boman, L., et al. (2018). *Dalton, a Molecular Electronic Structure Program*. Release v2018.2 http://daltonprogram.orgDataset.

Alberto, M. E., Mazzone, G., Quartarolo, A. D., Sousa, F. F. R., Sicilia, E., and Russo, N. (2014). Electronic Spectra and Intersystem Spin-Orbit Coupling in 1,2- and 1,3-squaraines. *J. Comp. Chem.* 35, 2107–2113. doi:10.1002/jcc.23725

Allen, F. H., Bird, C. M., Rowland, R. S., and Raithby, P. R. (1997). Hydrogen-Bond Acceptor and Donor Properties of Divalent Sulfur (Y-S-Z and R-S-H). *Acta Cryst. B53* 53, 696–701. doi:10.1107/S0108768197002644

Altun, A., Izsák, R., and Bistoni, G. (2021). Local Energy Decomposition of Coupled-Cluster Interaction Energies: Interpretation, Benchmarks, and Comparison with Symmetry-Adapted Perturbation Theory. *Int. J. Quant. Chem.* 121, e26339. doi:10.1002/qua.26339

Arunkumar, E., Sudeep, P. K., Kamat, P. V., Noll, B. C., and Smith, B. D. (2007). Singlet Oxygen Generation Using Iodinated Squaraine and Squaraine-Rotaxane Dyes. *New J. Chem.* 31, 677–683. doi:10.1039/b616224j

Avirah, R. R., Jayaram, D. T., Adarsh, N., and Ramaiah, D. (2012). Squaraine Dyes in PDT: from Basic Design to *In Vivo* Demonstration. *Org. Biomol. Chem.* 10, 911. doi:10.1039/c1ob06588b

Azarias, C., Budzák, Š., Laurent, A. D., Ulrich, G., and Jacquemin, D. (2016). Tuning ESIPT Fluorophores into Dual Emitters. *Chem. Sci.* 7, 3763–3774. doi:10.1039/C5SC04826E

Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*. Oxford: Oxford University Press.

Bahers, T. L., Adamo, C., and Ciofini, I. (2011). A Qualitative index of Spatial Extent in Charge-Transfer Excitations. *J. Chem. Theor. Comput.* 7, 2498–2506. doi:10.1021/ct200308m

Barcenas, G., Biaggne, A., Mass, O. A., Wilson, C. K., Obukhova, O. M., Kolosova, O. S., et al. (2021). First-principles Studies of Substituent Effects on Squaraine Dyes. *RSC Adv.* 11, 19029. doi:10.1039/D1RA01377G

Bassal, F., Laurent, A. D., Guennic, B. L., and Jacquemin, D. (2017). Exploring the Excited-States of Squaraine Dyes with Td-Dft, Sos-Cis(d) and Adc(2). *Dyes Pigm.* 138, 169–175. doi:10.1016/j.dyepig.2016.11.046

Beerepoot, M. T. P., Alam, M. M., Bednarska, J., Bartkowiak, W., Ruud, K., and Zaleśny, R. (2018). Benchmarking the Performance of Exchange-Correlation Functionals for Predicting Two-Photon Absorption Strengths. *J. Chem. Theor. Comput.* 14, 3677–3685. doi:10.1021/acs.jctc.8b00245

Beerepoot, M. T. P., Friese, D. H., List, N. H., Kongsted, J., and Ruud, K. (2015). Benchmarking Two-Photon Absorption Cross Sections: Performanceof CC2 and CAM-B3lyp. *Phys. Chem. Chem. Phys.* 17, 19306–19314. doi:10.1039/C5CP03241E

Boto, R. A., Peccati, F., Laplaza, R., Quan, C., Carbone, A., Piquemal, J.-P., et al. (2020). NCIPLOT4: A New Step towards a Fast Quantification of Noncovalent Interactions. Available at: https://github.com/juliacontrerasgarcia/nciplot. (accessed: June 20, 2021 [Dataset].

Chai, J.-D., and Head-Gordon, M. (2008a). Long-range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* 10, 6615. doi:10.1039/b810189b

Chai, J.-D., and Head-Gordon, M. (2008b). Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* 128, 084106. doi:10.1063/1.2834918

Chałasinski, G., Szczęśniak, M. M., Cieplak, P., and Scheiner, S. (1991). *Ab Initio* study of Intermolecular Potential of H_20 Trimer. *J. Chem. Phys.* 94, 2873. doi:10.1063/1.459809

Charaf-Eddin, A., Planchat, A., Mennucci, B., Adamo, C., and Jacquemin, D. (2013). Choosing a Functional for Computing Absorption and Fluorescence Band Shapes with TD-DFT. *J. Chem. Theor. Comput.* 9, 2749–2760. doi:10.1021/ct4000795

Chetti, P., Tripathi, A., Mittal, R., and Chaskar, A. (2021). A DFT Study on Electronic Excitations, Charge Transfer and NLO Properties of Visible Absorbing Squaraine and Thiosquaraine Dyes. *Ind. J. Chem.* 60A, 370–377.

Contreras-Garcia, J., Johnson, E. R., Keinan, S., Chaudret, R., Piquemal, J.-P., Beratan, D. N., et al. (2011). NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theor. Comput.* 7, 625–632. doi:10.1021/ct100641a

Corpinot, M. K., Guo, R., Tocher, D. A., Buanz, A. B. M., Gaisford, S., Price, S. L., et al. (2017). Are Oxygen and Sulfur Atoms Structurally Equivalent in Organic Crystals. *Cryst. Growth Des.* 17, 827–833. doi:10.1021/acs.cgd.6b01669

El-Sayed, M. A. (1963). Spin-Orbit Coupling and the Radiationless Processes in Nitrogen Heterocycles. *J. Chem. Phys.* 38, 2834–2837.

Elrod, M. J., and Saykally, R. J. (1994). Many-body Effects in Intermolecular Forces. *Chem. Rev.* 94, 1975–1997. doi:10.1021/cr00031a010

Espinosa, E., Mollins, E., and Lecomte, C. (1998). Hydrogen Bond Strengths Revealed by Topological Analyses of Experimentally Observed Electron Densities. *Chem. Phys. Lett.* 285, 170–173. doi:10.1016/S0009-2614(98)00036-0

Ferrer, F. J. A., Angeli, C., Cerezo, J., Coriani, S., Ferretti, A., and Santoro, F. (2019). The Intriguing Case of the One-Photon and Two-Photon Absorption of a Prototypical Symmetric Squaraine: Comparison of TDDFT and Wave Function Methods. *Chem. Photo. Chem.* 3, 1–7. doi:10.1002/cptc.201900121

Foster, J. P., and Weinhold, F. (1980). Natural Hybrid Orbitals. *J. Am. Chem. Soc.* 102, 7211–7218. doi:10.1021/ja00544a007

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 16 Revision C.01*. Wallingford CT [Dataset: Gaussian Inc.

Góra, U., Podeszwa, R., Cencek, W., and Szalewicz, K. (2011). Interaction Energies of Large Clusters from many-body Expansion. *J. Chem. Phys.* 135, 224102. doi:10.1063/1.3664730

Hansch, C., Leo, A., and Taft, R. W. (1991). A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* 91, 165–195. doi:10.1021/cr00002a004

Hinchliffe, A. (1984). Ab Initio study of the Hydrogen-Bonded Complexes $H_2X.HY$ (X = O, S, Se: Y = F, Cl, Br). *J. Mol. Struct. THEOCHEM* 106, 361–366. doi:10.1016/0166-1280(84)85040-X

Hirshfeld, F. L. (1977). Bonded-atom Fragments for Describing Molecular Charge Densities. *Theor. Chem. Acc.* 44, 129–138. doi:10.1007/BF00549096

Houari, Y., Charaf-Eddin, A., Laurent, A. D., Massue, J., Ziessel, R., Ulrich, G., et al. (2014). Modeling Optical Signatures and Excited-State Reactivities of Substituted Hydroxyphenylbenzoxazole (HBO) ESIPT Dyes. *Phys. Chem. Chem. Phys.* 16, 1319–1321. doi:10.1039/C3CP54703E

Isegawa, M., Peverati, R., and Truhlar, D. G. (2012). Performance of Recent and High-Performance Approximate Density Functionals for Time-dependent Density Functional Theory Calculations of Valence and Rydberg Electronic Transition Energies. *J. Chem. Phys.* 137, 244104. doi:10.1063/1.4769078

Jabłoński, M., Kaczmarek, A., and Sadlej, A. J. (2006). Estimates of the Energy of Intramolecular Hydrogen Bonds. *J. Phys. Chem. A.* 110, 10890–10898. doi:10.1021/jp062759o

Jacquemin, D., Planchat, A., Adamo, C., and Mennucci, B. (2012). TD-DFT Assessment of Functionals for Optical 0–0 Transitions in Solvated Dyes. *J. Chem. Theor. Comput.* 8, 2359–2372. doi:10.1021/ct300326f

Jedrzejewska, B., Zakrzewska, A., Mloston, G., Budzak, S., Mroczyńska, K., Grabarz, A. M., et al. (2016). Synthesis and Photophysical Properties of Novel Donor-Acceptor N-(Pirydyn-2-Yl) Substituted Benzo(thio)amides and Their Difluoroboranyl Derivatives. *J. Phys. Chem. A.* 120, 4116–4123. doi:10.1021/acs.jpca.6b04004

Johnson, E. R., Keinan, S., Mori-Sánchez, P., Contreras-García, J., Cohen, A. J., and Yang, W. (2010). Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* 132, 6498–6506. doi:10.1021/ja100936w

Karpenko, I. A., Klymchenko, A. S., Gioria, S., Kreder, R., Shulov, I., Villa, P., et al. (2015). Squaraine as a Bright, Stable and Environment-Sensitive Far-Red Label for Receptor-specific Cellular Imaging. *Chem. Commun.* 51, 2960–2963. doi:10.1039/C4CC09113B

Keith, T. A. (2019). Aimall (Version 19.10.12). TK Gristmill Software. Available at: http://aim.tkgristmill.com [Dataset] (accessed July 20, 2021).

Krepps, M. K., Parkin, S., and Atwood, D. A. (2001). Hydrogen Bonding with Sulfur. *Cryst. Growth Des.* 1, 291–297. doi:10.1021/cg015505v

Leang, S. S., Zahariev, F., and Gordon, M. S. (2012). Benchmarking the Performance of Time-dependent Density Functional Methods. *J. Chem. Phys.* 136, 104101. doi:10.1063/1.3689445

Liu, K., Wang, Y., Tu, Y., Å gren, H., and Luo, Y. (2007). Aggregation Effects on Two-Photon Absorption Spectra of Octupolar Molecules. *J. Chem. Phys.* 127, 026101. doi:10.1063/1.2751190

Liu, K., Wang, Y., Tu, Y., Å gren, H., and Luo, Y. (2008). Two-Photon Absorption of Hydrogen-Bonded Octupolar Molecule Cluster. *J. Phys. Chem. B* 112, 4387–4392. doi:10.1021/jp712118w

Lu, T., and Chen, F. (2021). Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comp. Chem.* 33, 580–592. doi:10.1002/jcc.22885

Neese, F., Aravena, D., Atanasov, M., Auer, A. A., Becker, U., Bistoni, G., et al. (2020a). ORCA: An Ab Initio, DFT and Semiempirical Electronic Structure Package, Program Version 4.2.1. Available at: https://www.faccts.de/orca/. (accessed: August 24, 2021 [Dataset].

Neese, F. (2018). Software Update: the ORCA Program System, Version 4.0. *Wires Comput. Mol. Sci.* 8, e1327. doi:10.1002/wcms.1327

Neese, F. (2012). The ORCA Program System. *Wires Comput. Mol. Sci.* 2, 73–78. doi:10.1002/wcms.81

Neese, F., Wennmohs, F., Becker, U., and Riplinger, C. (2020b). The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* 152, 224108. doi:10.1063/5.0004608

Ośmiałowski, B., Mroczyńska, K., Kolehmainen, E., Kowalska, M., Valkonen, A., Pietrzak, M., et al. (2013). Association of N-(Pyridin-2-yl),N'-substituted Ureas with 2-Amino-1,8-Naphthyridines and Benzoates: NMR and Quantum Chemical Studies of the Substituent Effect on Complexation. *J. Org. Chem.* 78, 7582–7593. doi:10.1021/jo4011393

Pawlicki, M., Colins, H. A., Denning, R. G., and Anderson, H. L. (2009). Two-photon Absorption and the Design of Two-Photon Dyes. *Angew. Chem. Int. Ed.* 48, 3244–3266. doi:10.1002/anie.200805257

Peceli, D., Hu, H., Fishman, D. A., Webster, S., Przhonska, O. V., Kurdyukov, V. V., et al. (2013). Enhanced Intersystem Crossing Rate in Polymethine-like Molecules: Sulfur-Containing Squaraines versus Oxygen-Containing Analogues. *J. Phys. Chem. A.* 117, 2333–2346. doi:10.1021/jp400276g

Platts, J. A., Howard, S. T., and Bracke, B. R. F. (1996). Directionality of Hydrogen Bonds to Sulfur and Oxygen. *J. Am. Chem. Soc.* 118, 2726–2733. doi:10.1021/ja952871s

Popelier, P. L. (2000). *Atoms in Molecules: An Introduction*. London: Prentice-Hall.

Prabhakar, C., Bhanuprakash, K., Rao, V. J., Balamuralikrishna, M., and Rao, D. N. (2010). Third Order Nonlinear Optical Properties of Squaraine Dyes Having Absorption below 500 Nm: A Combined Experimental and Theoretical Investigation of Closed Shell Oxyallyl Derivatives. *J. Phys. Chem. C* 114, 6077–6089. doi:10.1021/jp908475n

Prabhakar, C., Chaitanya, G. K., Sitha, S., Bhanuprakash, K., and Rao, V. J. (2005a). Role of the Oxyallyl Substructure in the Near Infrared (NIR) Absorption in Symmetrical Dye Derivatives: A Computational Study. *J. Phys. Chem. A.* 109, 2614–2622. doi:10.1021/jp044954d

Prabhakar, C., Yesudas, K., Krishna Chaitanya, G., Sitha, S., Bhanuprakash, K., and Rao, V. J. (2005b). Near-Infrared Absorption in Symmetric Squarylium and Croconate Dyes: A Comparative Study Using Symmetry-Adapted Cluster-Configuration Interaction Methods. *J. Phys. Chem. A.* 109, 8604–8616. doi:10.1021/jp0516184

Pristash, S. R., Corp, K. L., Rabe, E. J., and Schlenker, C. W. (2020). Heavy-Atom-Free Red-To-Yellow Photon Upconversion in a Thiosquaraine Composite. *ACS Appl. Energ. Mater.* 3, 19. doi:10.1021/acsaem.9b01808

Quartarolo, A. D., Sicilia, E., and Russo, N. (2009). On the Potential Use of Squaraine Derivatives as Photosensitizers in Photodynamic Therapy: A Tddft and Ricc2 Survey. *J. Chem. Theor. Comput.* 5, 1849–1857. doi:10.1021/ct900199j

Rablen, P. R., Lockman, J. W., and Jorgensen, W. L. (1998). Ab Initio Study of Hydrogen-Bonded Complexes of Small Organic Molecules with Water. *J. Phys. Chem. A.* 102, 3782–3797. doi:10.1021/jp980708o

Reed, A. E., Curtiss, L. A., and Weinhold, F. (1988). Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.* 88, 899–926. doi:10.1021/cr00088a005

Řezáč, J., Huang, Y., Hobza, P., and Beran, G. J. O. (2015). Benchmark Calculations of Three-Body Intermolecular Interactions and the Performance of Low-Cost Electronic Structure Methods. *J. Chem. Theor. Comput.* 11, 3065–3079. doi:10.1021/acs.jctc.5b00281

Richard, R. M., Bakr, B. W., and Sherrill, C. D. (2018). Understanding the Many-Body Basis Set Superposition Error: Beyond Boys and Bernardi. *J. Chem. Theor. Comput.* 14, 2386–2400. doi:10.1021/acs.jctc.7b01232

Ros-Lis, J. V., García, B., Diego, J., Martínez-Máñez, R., Sancenón, F., Soto, J., et al. (2004). Squaraines as Fluoro-Chromogenic Probes for Thiol-Containing Compounds and Their Application to the Detection of Biorelevant Thiols. *J. Am. Chem. Soc.* 126, 4064–4065. doi:10.1021/ja03198710.1021/ja031987i

Ros-Lis, J. V., Martínez-Máñez, R., and Soto, J. (2002). A Selective Chromogenic Reagent for Cyanide Determination. *Chem. Commun.* 34, 2248–2249. doi:10.1039/B206500B

Srinivas, K., Prabhakar, C., Devi, C. L., Yesudas, K., Bhanuprakash, K., and Rao, V. J. (2007). Enhanced Diradical Nature in Oxyallyl Derivatives Leads to Near

Infra Red Absorption: A Comparative Study of the Squaraine and Croconate Dyes Using Computational Techniques. *J. Phys. Chem. A.* 111, 3378–3386. doi:10.1021/jp067410f

Steiner, T. (2000). S-H.S Hydrogen-Bond Chain in Thiosalicylic Acid. *Acta Cryst.* C56, 876–877. doi:10.1107/s0108270100005898

Valdés-Martínez, J., Hernández-Ortega, S., Rubio, M., Li, D. T., Swearingen, J. K., Kaminsky, W., et al. (2004). Study of the Sulfur Atom as Hydrogen Bond Acceptor in N(2)-pyridylmethyl-N-arylthioureas. *J. Chem. Crystall.* 34, 533–540. doi:10.1023/B:JOCC.0000042022.16392.99

Webster, S., Peceli, D., Hu, H., Padilha, L. A., Przhonska, O. V., Masunov, A. E., et al. (2010). Near-Unity Quantum Yields for Intersystem Crossing and Singlet Oxygen Generation in Polymethine-like Molecules: Design and Experimental Realization. *J. Phys. Chem. Lett.* 1, 2354–2360. doi:10.1021/jz100381v

Wells, B. H., and Wilson, S. (1983). van der Waals interaction potentials: Many-body basis set superposition effects. *Chem. Phys. Lett.* 101, 429. doi:10.1016/0009-2614(83)87508-3

Wennmohs, F., Staemmler, V., and Schindler, M. (2003). Theoretical Investigation of Weak Hydrogen Bonds to Sulfur. *J. Chem. Phys.* 119, 3208. doi:10.1063/1.1588291

Yesudas, K., Chaitanya, G. K., Prabhakar, C., Bhanuprakash, K., and Rao, V. J. (2006). Structure, Bonding, and Lowest Energy Transitions in Unsymmetrical Squaraines: A Computational Study. *J. Phys. Chem. A.* 110, 11717–11729. doi:10.1021/jp064074u

Zhao, Y., and Truhlar, D. G. (2008). The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* 120, 215–241. doi:10.1007/s00214-007-0310-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 07 January 2022
doi: 10.3389/fchem.2021.801426

Infrared Spectra of Hydrogen-Bonded Molecular Complexes Under Spatial Confinement

Marta Chołuj[1]*, *Josep M. Luis*[2], *Wojciech Bartkowiak*[1] and *Robert Zaleśny*[1]*

[1]Faculty of Chemistry, Wrocław University of Science and Technology, Wrocław, Poland, [2]Institute of Computational Chemistry and Catalysis and Department of Chemistry, University of Girona, Catalonia, Spain

Infrared (IR) spectroscopy is commonly used in chemical laboratories to study the geometrical structure of molecules and molecular complexes. The analysis of experimental IR spectra can nowadays be reliably supported by the results of quantum-chemical computations as vibrational frequencies and corresponding vibrational transition intensities are routinely calculated using harmonic approximation by virtually all quantum chemistry packages. In the present study we combine the methodology of computing vibrational spectra using high-level electron correlation treatments with an analytical potential-based approach to take into account spatial confinement effects. Using this approach, we perform a pioneering analysis of the impact of the spatial confinement caused by a cylindrical harmonic oscillator potential on the harmonic vibrational transition intensities and frequencies of two hydrogen-bonded complexes: HCN...HCN and HCN...HNC. The emphasis is put on the largest-intensity bands, which correspond to the stretching vibrations. The obtained results demonstrate that embedding the molecular complexes in an external confining potential causes significant changes of transition intensities and vibrational frequencies.

Keywords: infrared spectrum (IR), spatial confinement, harmonic oscillator potential, hydrogen-bonded complexes, vibrational transition intensity

OPEN ACCESS

Edited by:
Sonja Grubisic,
University of Belgrade, Serbia

Reviewed by:
Malgorzata Biczysko,
Shanghai University, China
Roberto Cammi,
University of Parma, Italy

***Correspondence:**
Marta Chołuj
marta.choluj@pwr.edu.pl
Robert Zaleśny
robert.zalesny@pwr.edu.pl

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *25 October 2021*
Accepted: *30 November 2021*
Published: *07 January 2022*

Citation:
Chołuj M, Luis JM, Bartkowiak W and Zaleśny R (2022) Infrared Spectra of Hydrogen-Bonded Molecular Complexes Under Spatial Confinement. Front. Chem. 9:801426. doi: 10.3389/fchem.2021.801426

1 INTRODUCTION

Studies of molecular structure using infrared spectroscopy have become a fairly routine task and nowadays IR spectrometers are part of virtually all chemical laboratories. This technique (as well as other complementary vibrational spectroscopies) allows to gain an insight into the geometrical structure of molecules and molecular complexes. In fact, changes in frequencies and intensities of some key vibrational transitions are used as fingerprints to monitor the hydrogen or halogen bonds of molecular complexes. The analysis of experimental IR spectra can nowadays be reliably supported by the results of quantum-chemical computations as vibrational frequencies and corresponding transition intensities are routinely calculated using harmonic approximation by practically all packages. High-level electron-correlation methods can be used for band assignments in vibrational spectra of small- and medium-sized molecules and molecular complexes. Moreover, they can also be used to map the changes in infrared vibrational spectroscopic signatures to structural changes. The present study contributes to these theoretical efforts and its goal is to apply the methodology to compute the harmonic infrared vibrational spectra using high-level electron correlation treatments to spatially confined hydrogen-bonded molecular complexes.

The term "spatial confinement" is used throughout this work to refer to atoms and molecules trapped inside chemical cages (such as zeolites, mesoporous silica/organosilica, metal-organic frameworks, nanotubes and fullerenes), matter under high pressure, quantum wells, wires and dots as well as other constraining environments (Jaskólski (1996); Sen (2014)). On the theoretical basis, the spatial confinement is often simulated by external repulsive potentials, which reproduce the effect of orbital compression (connected with valence repulsion). This approach describes the interaction of objects with chemically and electronically inert environments (Jaskólski (1996)). Over the years, much effort has been devoted to the studies of the spatial confinement phenomenon, demonstrating its significant effects on the physical and chemical properties of atomic and molecular systems (Bartkowiak et al. (2021); Buchachenko (2001); Kozłowska et al. (2014); Chattaraj and Sarkar (2003); Cammi et al. (2008, 2012); Zaleśny et al. (2013); Góra et al. (2012); Lo and Klobukowski (2006)). Numerous interesting results were obtained, e.g., it was shown that the spatial confinement, in the form of analytical potential, causes an increase of the total energy (Buchachenko (2001); Bartkowiak and Strasburger (2010); Bielińska-Wąż et al. (2001)) and the separation of HOMO and LUMO orbitals (Borgoo et al. (2008, 2009); Kozłowska et al. (2014)). Upon increasing the confinement strength, the polarizability of the spatially restricted systems decreases (Holka et al. (2005); Góra et al. (2012); Lo and Klobukowski (2006); Zaleśny et al. (2015); Kozłowska et al. (2015); Zaleśny et al. (2017)). Moreover, a shortening of the bond lengths was reported for molecules and molecular complexes embedded in confining potentials (Cammi et al. (2008, 2012); Zaleśny et al. (2013); Zaleśny et al. (2015)). Nevertheless, despite recent progress, there are still many aspects of spatial confinement which have not received sufficient attention yet. One of them is the modeling of IR spectra of spatially confined molecules and molecular complexes. In this context, one should not overlook the paper by Bonfim and Pilling who studied the effect of chemical environment on some properties of "trapped molecules" by using Polarizable Continuum Model (Bonfim and Pilling (2017)). These authors considered the effect of dielectric constant ($3 < \epsilon < 180$) on vibrational frequencies and transition intensities. It was demonstrated that in the case of all analyzed molecules, i.e., CO, CO_2, H_2O—components of the astrophysical ices matrix, the increase of the value of dielectric constant is accompanied by larger band strengths. Other literature reports are largely limited to the analysis of the spatial confinement effect on shifts of vibrational frequencies (Shameema et al. (2006); Jana and Bandyopadhyay (2011); Song et al. (2005); Pagliai et al. (2014); Cammi et al. (2012); Boccalini et al. (2021)). Therefore, the goal of this project is to fill this gap by performing a pioneering analysis of the impact of spatial restriction, modeled by confining cylindrical analytical harmonic potential, on the infrared spectra (vibrational frequencies and transition intensities) of hydrogen-bonded molecular complexes.

2 METHODS

In this work, the effect of spatial confinement on the vibrational properties of HCN. . .HCN and HCN. . .HNC complexes, with symmetry axis parallel to Cartesian z direction, was modeled by the cylindrical harmonic oscillator potential:

$$V_{conf}(\vec{r}_i) = \frac{1}{2}\omega^2 (x_i^2 + y_i^2), \tag{1}$$

where the ω parameter controls the strength of spatial confinement, which is obtained by changing the curvature of the harmonic potential and the $\vec{r}_i$ refer *only* to the electronic coordinates. Thus, the effect of confinement on nuclei is accounted for indirectly by its effect on the electrons. Because of its symmetry, the cylindrical harmonic oscillator potential can be used as simplified representation of nanotube-like confining cages. V_{conf} as given by **Eq. 1** was added to the Hamiltonian of the isolated complexes and subsequently vibrational-structure calculations were performed. In particular, the calculations of the infrared intensity for i-th mode (I_i) under the influence of spatial confinement were performed based on the following equation:

$$I_i = \frac{N\pi}{3c^2} \sum_{j=x,y,z} \left(\frac{\partial \mu_j}{\partial Q_i}\right)^2 \tag{2}$$

where N is Avogadro's number, c is speed of light and $\frac{\partial \mu_j}{\partial Q_i}$ is derivative of j-th Cartesian component of dipole moment with respect to normal mode Q_i.

Vibrational-structure computations were performed using MP2 and CCSD(T) methods and aug-cc-pVTZ basis set. MP2 method was used to determine gas-phase equilibrium geometries, vibrational frequencies and transition intensities analytically, as implemented in the GAUSSIAN 2009 package (Frisch et al. (2009)). On the other hand, at the CCSD(T)/aug-cc-pVTZ level the geometries of the studied complexes were optimized in vacuum and in the presence of cylindrical harmonic oscillator potential using custom computer routines based on total energies computed with the aid of the GAUSSIAN 2009 package (Frisch et al. (2009)). The calculations of infrared intensities of the spatially confined molecular complexes were performed at the CCSD(T)/aug-cc-pVTZ level using custom computer routines. To that end, we employed double numerical differentiation, i.e., the dipole moment components were computed numerically for a mesh of electric fields F ($\pm 2^n \times F$, where n = 0, . . ., 8, F = 0.000 2 au) at a set of displaced geometries (with mesh $\pm 2^m \times \Delta$, where $m = 0, \ldots, 4$, $\Delta = 0.01$ bohr). The numerical accuracy of both numerical diffferentiation procedures was controlled with the aid of Romberg-Rutishauser scheme (Medved et al. (2007)).

3 RESULTS AND DISCUSSION

To investigate the effect of spatial confinement on vibrational transition intensities, we chose two hydrogen-bonded complexes, i.e., HCN. . .HCN and HCN. . .HNC. The rationale behind

TABLE 1 | Harmonic frequencies (ν_i, in cm^{-1}) and infrared intensities (I_i, in km/mol) for vibrational normal modes i of HCN...HCN and HCN...HNC complexes computed at the MP2/aug-cc-pVTZ level of theory.

	HCN...HCN		HCN...HNC	
i	ν_i	I_i	ν_i	I_i
1	51.42	6.55	66.36	4.14
2	51.42	6.55	66.36	4.14
3	123.48	2.00	162.97	26.14
4	146.20	42.13	162.97	26.14
5	146.20	42.13	165.28	5.52
6	728.48	32.46	732.06	8.44
7	728.48	32.46	732.06	8.44
8	820.01	34.11	751.31	136.00
9	820.01	34.11	751.31	136.00
10	2019.47	8.86	2013.80	1.15
11	2038.53	0.97	2050.26	1.32
12	3380.56	389.39	3449.34	147.88
13	3452.89	76.78	3560.10	1190.32

selection of these two complexes stems from the results presented in the earlier studies by some of the present authors, devoted to the analysis of the effect of spatial confinement on the electronic and nuclear relaxation polarizability as well as first and second hyperpolarizability of molecules and molecular complexes (Zaleśny et al. (2015); Zaleśny et al. (2017)). The nuclear relaxation (hyper) polarizabilities are the major contribution to the vibrational (hyper) polarizabilities and include all their low-order corrections (Bishop (1998); Kirtman et al. (2000)). There are two main conclusions of these studies. Firstly, the effect of spatial confinement increases with the order of electrical property, both for electronic and vibrational counterparts, and it is much more pronounced for the former. In other words, the least significant effect was observed for electronic and nuclear relaxation polarizability. Secondly, the decrease of vibrational first hyperpolarizability was largely due to the harmonic part. The nuclear relaxation polarizability can be directly linked with infrared intensity given by **Eq. 2** as both involve electric dipole moment (μ) derivatives with respect to normal modes (Q). Average nuclear relaxation polarizability is given by:

$$\langle \alpha^{\mathrm{nr}} \rangle = \frac{1}{3} \sum_{j=x,y,z} \alpha_{jj}^{\mathrm{nr}} = \frac{1}{3} \sum_{i=1} \frac{1}{\omega_i^2} \left[\sum_{j=x,y,z} \left(\frac{\partial \mu_j}{\partial Q_i} \right)^2 \right] = \sum_{i=1} \langle \alpha^{\mathrm{nr}} \rangle_i \quad (3)$$

where i labels normal modes with angular frequencies ω_i. We may re-write the equation defining the nuclear relaxation polarizability in terms of the IR intensity I_i:

$$\langle \alpha^{\mathrm{nr}} \rangle = \sum_i \langle \alpha^{\mathrm{nr}} \rangle_i = \frac{c^2}{N\pi} \sum_i \frac{I_i}{\omega_i^2} \quad (4)$$

Then for each mode we have the following relationship between the nuclear relaxation contribution to mode i and the corresponding intensity I_i:

$$I_i \sim \omega_i^2 \langle \alpha^{\mathrm{nr}} \rangle_i \quad (5)$$

It follows from the above relations that the effect of spatial confinement on nuclear relaxation polarizabilities can be predominantly manifested by the contributions from low-frequency normal modes, provided that corresponding dipole moment derivatives are significant. These expressions also allow for rationalizing why the absolute value of α^{nr} is more than order of magnitude larger for HCN...HCN than for HCN (Zaleśny et al. (2015); Zaleśny et al. (2017)). In the case of hydrogen-bonded complexes there are X-H stretching modes corresponding to significant dipole moment changes which give rise to large-intensity bands. One may thus expect a significant effect of spatial confinement on vibrational intensities for such modes. Taken together, these evidences support the choice of HCN...HCN and HCN...HNC for the pioneering analysis performed in this study.

In what follows we will analyze the effect of confinement by choosing two values of ω parameter, i.e., 0.1 and 0.2 au. It was shown that these values approximately correspond to the realistic chemical environment, i.e., they reproduce exchange repulsion for linear weakly-bound complexes enclosed in carbon nanotubes (Zaleśny et al. (2017)). Such link can be established by comparison of the Hartree-Fock interaction energies between both studied complexes and the (4,4) carbon natotube with the results obtained using the analytic cylindrical confining potential. The results presented in Ref (Zaleśny et al. (2017)) demonstrate

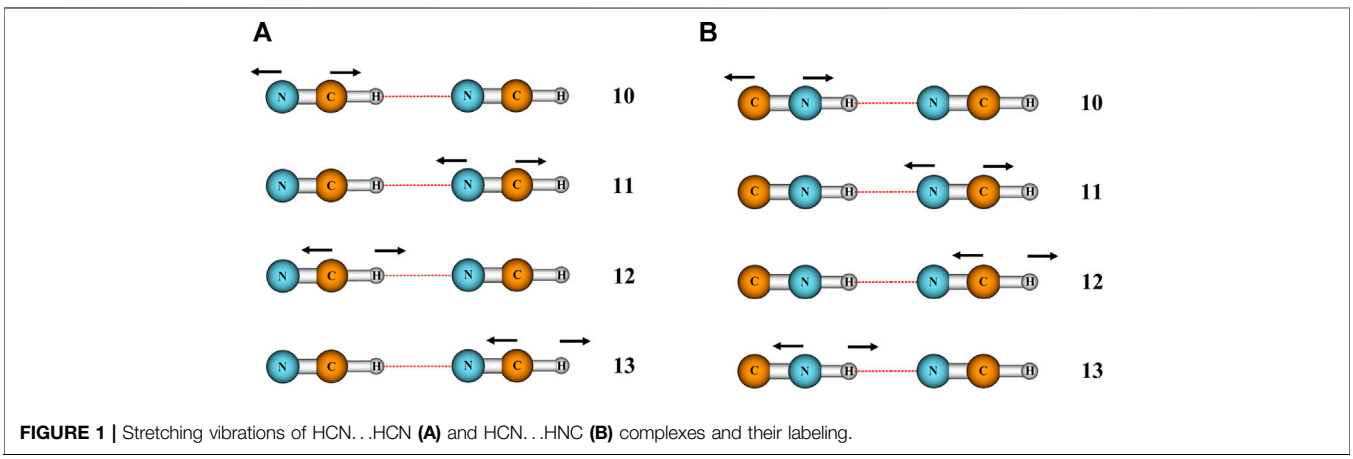


FIGURE 1 | Stretching vibrations of HCN...HCN **(A)** and HCN...HNC **(B)** complexes and their labeling.

TABLE 2 | Frequencies (ν_i, in cm^{-1}) and infrared intensities (I_i, in km/mol) for vibrational normal modes *i* of HCN...HCN and HCN...HNC complexes computed at the CCSD(T)/aug-cc-pVTZ level of theory.

i	ν_i			I_i		
	HCN...HCN					
	ω = **0.0 au**	ω = **0.1 au**	ω = **0.2 au**	ω = **0.0 au**	ω = **0.1 au**	ω = **0.2 au**
10	2103.1	2138.8	2225.1	15.10	15.45	16.54
11	2123.8	2157.9	2243.2	4.75	5.20	5.89
12	3363.8	3418.0	3556.9	346.96	336.32	311.94
13	3421.1	3477.2	3618.4	58.50	54.80	46.52
	HCN...HNC					
10	2044.7	2079.5	2166.4	11.80	11.14	9.86
11	2134.2	2167.9	2251.9	3.80	4.14	5.00
12	3418.8	3474.7	3616.9	109.16	102.52	89.37
13	3590.1	3641.1	3766.9	1119.50	1109.96	1098.96

that the HF interaction energy between the (4,4) nanotube and HCN...HCN (HCN...HNC) is 133 kcal/mol (134 kcal/mol). The amplitude of the confining potential (ω) can be adjusted to represent this repulsive interaction:

$$\Delta E_{\text{int}}^{\text{HF}} \approx E(\omega) - E(\omega = 0.0)$$

The amplitude ω = 0.1 a.u., through the above equation, roughly corresponds to Hartree-Fock interaction energy of HCN...HCN and HCN...HNC complexes trapped inside (4,4) carbon nanotube. The second chosen value of the amplitude thus corresponds to much larger confining environment caused by a carbon nanotube with smaller diameter.

In order to shed light on the vibrational structure of isolated complexes, we evaluated the infrared intensities for all vibrational modes of HCN...HCN and HCN...HNC complexes, hereafter shortly referred to as (A(*i*)) and (B(*i*)). The results of calculations, performed at the cost-effective MP2/aug-cc-pVTZ level of theory, are presented in **Table 1**. Modes 12 and 13 have the largest IR intensities among all vibrations and this holds for both studied complexes. In particular, the IR intensities of the C-H and N-H stretching vibrations involving the hydrogen of hydrogen bond (A (12) and B (13)) are equal to 389 and 1,190 km/mol, respectively, and dominate the IR spectrum. In the reminder of the analysis, we will use highly accurate CCSD(T)/aug-cc-pVTZ level of theory to analyze the effect of confinement on four stretching modes presented in **Figure 1**, corresponding to the two C-N stretchings (A (10,11) and B (10,11)) and to the C-H (A (12,13), B (12)) and N-H stretchings vibrations (B (13)). We will start with the analysis of the influence of spatial confinement on the vibrational frequencies. **Table 2** reports the computed vibrational frequencies of HCN...HCN and HCN...HNC complexes in vacuum and in the presence of confinement represented by the cylindrical harmonic oscillator potential. It is clear from these data that ν_i for modes 10-13 of both complexes are shifted to higher values by the spatial confinement, which agrees with the fact that the energy of a chemical systems increases under confinement. The increase of all ν_i values upon growing confinement strength from ω = 0.0 to ω = 0.1 au falls in between 33.7 and 56.1 cm^{-1}. On the other hand, the changes caused by the confinement when its strength is equal to 0.2 au are much more pronounced, as the increase of ν_i ranges from 117.7 to 198.1 cm^{-1}. Although the absolute shift of the frequencies caused by the spatial restriction is larger for the C-H and N-H stretchings (modes A (12,13) and B (12,13)) than the frequencies of the other two modes (A (10,11) and B (10,11)), analyzing the relative shift respect to the vacuum frequency is very similar for the four vibrations (i.e., 1.4–1.7% for ω = 0.1 au and 4.9–5.8% for ω = 0.2 au). Similar findings, i.e., an increase of vibrational frequencies under pressure, were reported for various molecular systems, such as diborane (Cammi et al. (2012)), C_{60} and C_{70} fullerenes (Pagliai et al. (2014)) and SF_6 molecule (Boccalini et al. (2021)).

We will now turn to the infrared intensities (see **Table 2**). Similarly to what we reported for vibrational frequencies, there is a monotonic perturbation in I_i for all four modes as the amplitude of confinement increases. The absolute changes in the IR intensities of modes A (10,11) and B (10,11) are very small and do not exceed 2 km/mol. Nevertheless, we highlight that for HCN...HCN the intensities of modes 10 and 11 *increase* on passing from ω = 0 through ω = 0.1 to ω = 0.2 au. The same is observed for mode 11 for HCN...HNC complex. In contrast, the modes A (12,13) and B (12,13) suffer a much stronger impact due to the spatial confinement, as there is a quite substantial decrease in I_i computed for unconfined and confined complexes that ranges from 3.7 to 20.54 km/mol. The largest-intensity bands in infrared spectra of the studied molecular complexes are the C-H and N-H stretchings which hydrogens are involved in the hydrogen bond (modes A (12) and B (13)). These two bands are the most significantly influenced by the spatial confinement. These observations can be further elucidated based on simple geometrical arguments. For ω = 0.1 a.u. there is shortening of all covalent bonds in HCN...HCN and HCN...HNC approximately by 0.006 Å, while the N...H bond length is shorter by 0.011 Å. Larger strength of confinement, i.e., ω = 0.2 a.u., leads to further shortening of covalent bonds (ranging from 0.020 to 0.022 Å with respect to the unconfined systems). Moreover, there is a significant decrease in the N...H bond length ranging from 0.035 Å (HCN...HCN) to 0.039 Å

(HCN. . .HNC). The latter data explain the significant changes in intensities for modes A (12) and B (13). The confinement-induced geometrical changes observed for HCN. . .HCN and HCN. . .HNC complexes are in line with reports for other systems, e.g., CO_2, C_2H_2, HCN and HCCCN (Zaleśny et al. (2013); Zaleśny et al. (2015)) Similar trends were also observed for the hydrogen bond length in other theoretical as well as experimental works concerning the effect of spatial confinement on the HB complexes (Lipkowski et al. (2014); Roztoczyńska et al. (2014); Ajami et al. (2011); Dougherty (1998)). As highlighted by Cammi et al. the bond shortening upon confinement can be linked with the deformation of the molecular electronic charge density, due to the Pauli repulsive interaction with the environment (Cammi et al. (2008; 2012)).

4 SUMMARY

In summary, in the present study we demonstrated that the spatial confinement simulated by a cylindrical harmonic oscillator potential significantly affects vibrational spectra of hydrogen-bonded molecular complexes. Based on the obtained results we can draw several conclusions concerning the impact of spatial confinement on four stretching vibrations along molecular axis of HCN. . .HCN and HCN. . .HNC complexes:

- the changes in the vibrational transition intensities and corresponding frequencies are monotonic, and yet they are much more pronounced upon increasing the confinement strength from $\omega = 0.0$ to $\omega = 0.2$ au than on passing from $\omega = 0.0$ to $\omega = 0.1$ au;
- the presence of cylindrical harmonic oscillator potential leads to an increase of vibrational frequencies for all studied stretching vibrations;
- the vibrational transition intensities of modes corresponding to the C-H and N-H stretchings significantly decrease when the strength of the spatial confinement grows;
- the spatial confinement causes very small absolute changes in the IR intensities of C-N stretching vibrations and in the case of HCN. . .HCN complex the increase of confinement leads to increase in the corresponding vibrational transition intensities.

We believe that our results will not only contribute to broadening the knowledge on the spatial confinement phenomenon but they will also complement and support the experimental efforts in this field, by giving valuable insight into the nature of changes in the infrared spectra caused by spatial confinement. Our findings are particularly important from the point of view of astrochemistry, as the astrochemical measurements probe molecular species that experience extreme conditions, such as high pressure, and one of the most fundamental techniques used for exploration of cosmic space is the infrared spectroscopy. Finally, we highlight that in order to predict IR spectra of confined species with high accuracy it is mandatory to account for anharmonicity effects.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

The analysis and design of data and the writing of the article was carried out predominantly by MC. RZ performed the calculations. RZ, WB, and JML helped in the design and review of data. All authors contributed to the writing and review of the article.

FUNDING

This work was supported by MINECO-Spain for project PGC 2018-098212-B-C22 and Generalitat de Catalunya for project 2017SGR264 to JML.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Wroclaw Centre for Networking and Supercomputing for the generous allotment of computer time.

REFERENCES

Ajami, D., Tolstoy, P. M., Dube, H., Odermatt, S., Koeppe, B., Guo, J., et al. (2011). Encapsulated Carboxylic Acid Dimers with Compressed Hydrogen Bonds. *Angew. Chem. Int. Ed.* 50, 528–531. doi:10.1002/anie.201002182

Bartkowiak, W., Chołuj, M., and Kozłowska, J. (2021). "Effect of Confinement on the Optical Response Properties of Molecules," in *Chemical Reactivity in Confined Systems: Theory, Modelling and Applications.* Editors P. K. Chattaraj and D. Chakraborty (Wiley). doi:10.1002/9781119683353.ch12

Bartkowiak, W., and Strasburger, K. (2010). Linear and Nonlinear Electric Properties of Spatially Confined LiH Molecule, Studied with the Finite Field Method. *J. Mol. Struct. THEOCHEM* 960, 93–97. doi:10.1016/j.theochem.2010.08.028

Bielińska-Wąż, D., Diercksen, G. H. F., and Klobukowski, M. (2001). Quantum Chemistry of Confined Systems: Structure and Vibronic Spectra of a Confined Hydrogen Molecule. *Chem. Phys. Lett.* 349, 215–219. doi:10.1016/S0009-2614(01)01226-X

Bishop, D. M. (1998). Molecular Vibration and Nonlinear Optics. *Adv. Chem. Phys.* 104, 1–40.

Boccalini, M., Cammi, R., Pagliai, M., Cardini, G., and Schettino, V. (2021). Toward an Understanding of the Pressure Effect on the Intramolecular Vibrational Frequencies of Sulfur Hexafluoride. *J. Phys. Chem. A.* 125, 6362–6373. doi:10.1021/acs.jpca.1c02595

Bonfim, V. S., and Pilling, S. (2017). The Influence of Chemical Environment on the Infrared Spectra of Embedded Molecules in Astrophysical Ices. *Proc. IAU* 13, 346–352. doi:10.1017/s1743921317008110

Borgoo, A., Tozer, D. J., Geerlings, P., and De Proft, F. (2009). Confinement Effects on Excitation Energies and Regioselectivity as Probed by the Fukui Function and the Molecular Electrostatic Potential. *Phys. Chem. Chem. Phys.* 11, 2862–2868. doi:10.1039/b820114e

Borgoo, A., Tozer, D. J., Geerlings, P., and De Proft, F. (2008). Influence of Confinement on Atomic and Molecular Reactivity Indicators in DFT. *Phys. Chem. Chem. Phys.* 10, 1406–1410. doi:10.1039/b716727j

Buchachenko, A. L. (2001). Compressed Atoms. *J. Phys. Chem. B* 105, 5839–5846. doi:10.1021/jp003852u

Cammi, R., Cappelli, C., Mennucci, B., and Tomasi, J. (2012). Calculation and Analysis of the Harmonic Vibrational Frequencies in Molecules at Extreme Pressure: Methodology and Diborane as a Test Case. *J. Chem. Phys.* 137, 154112. doi:10.1063/1.4757285

Cammi, R., Verdolino, V., Mennucci, B., and Tomasi, J. (2008). Towards the Elaboration of a QM Method to Describe Molecular Solutes under the Effect of a Very High Pressure. *Chem. Phys.* 344, 135–141. doi:10.1016/j.chemphys.2007.12.010

Chattaraj, P. K., and Sarkar, U. (2003). Chemical Reactivity of the Spherically Confined Atoms. *Chem. Phys. Lett.* 372, 805–809. doi:10.1016/s0009-2614(03)00516-5

Dougherty, R. C. (1998). Temperature and Pressure Dependence of Hydrogen Bond Strength: A Perturbation Molecular Orbital Approach. *J. Chem. Phys.* 109, 7372–7378. doi:10.1063/1.477343

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (20092009). *Gaussian 09 Revision D.01*. Wallingford CT: Gaussian Inc [Dataset].

Góra, R. W., Zaleśny, R., Kozłowska, J., Naciążek, P., Roztoczyńska, A., Strasburger, K., et al. (2012). Electric Dipole (Hyper)Polarizabilities of Spatially Confined LiH Molecule. *J. Chem. Phys.* 137, 094307. doi:10.1063/1.4748144

Holka, F., Neogrády, P., Kellö, V., Urban, M., and Diercksen, G. H. F. (2005). Polarizabilities of Confined Two-Electron Systems: the 2-Electron Quantum Dot, the Hydrogen Anion, the Helium Atom and the Lithium Cation. *Mol. Phys.* 103, 2747–2761. doi:10.1080/00268970500181160

Jana, M., and Bandyopadhyay, S. (2011). Vibrational Spectrum of Water Confined in and Around Cyclodextrins. *Chem. Phys. Lett.* 509, 181–185. doi:10.1016/j.cplett.2011.04.103

Jaskólski, W. (1996). Confined Many-Electron Systems. *Phys. Rep.* 27, 1–66.

K. D. Sen (Editor) (2014). *Electronic Structure of Quantum Confined Atoms and Molecules* (Springer International Publishing).

Kirtman, B., Champagne, B., and Luis, J. M. (2000). Efficient Treatment of the Effect of Vibrations on Electrical, Magnetic, and Spectroscopic Properties. *J. Comput. Chem.* 21, 1572–1588. doi:10.1002/1096-987x(200012)21:16<1572:aid-jcc14>3.0.co;2-8

Kozłowska, J., Roztoczyńska, A., and Bartkowiak, W. (2015). About Diverse Behavior of the Molecular Electric Properties upon Spatial Confinement. *Chem. Phys.* 456, 98–105. doi:10.1016/j.chemphys.2014.12.003

Kozłowska, J., Zaleśny, R., and Bartkowiak, W. (2014). On the Nonlinear Electrical Properties of Molecules in Confined Spaces - from Cylindrical Harmonic Potential to Carbon Nanotube Cages. *Chem. Phys.* 428, 19–28. doi:10.1016/j.chemphys.2013.10.007

Lipkowski, P., Kozłowska, J., Roztoczyńska, A., and Bartkowiak, W. (2014). Hydrogen-Bonded Complexes upon Spatial Confinement: Structural and Energetic Aspects. *Phys. Chem. Chem. Phys.* 16, 1430–1440. doi:10.1039/c3cp53583e

Lo, J. M. H., and Klobukowski, M. (2006). Computational Studies of One-Electron Properties of Lithium Hydride in Confinement. *Chem. Phys.* 328, 132–138. doi:10.1016/j.chemphys.2006.06.019

Medved, M., Stachovà, M., Jacquemin, D., Andrè, J. M., and Perpeté, E. A. (2007). A Generalized Romberg Differentiation Procedure for Calculation of Hyperpolarizabilities. *J. Mol. Struc. (Theochem)* 847, 39–46. doi:10.1016/j.theochem.2007.08.028

Pagliai, M., Cardini, G., and Cammi, R. (2014). Vibrational Frequencies of Fullerenes C60 and C70 under Pressure Studied with a Quantum Chemical Model Including Spatial Confinement Effects. *J. Phys. Chem. A.* 118, 5098–5111. doi:10.1021/jp504173k

Roztoczyńska, A., Kozłowska, J., Lipkowski, P., and Bartkowiak, W. (2014). Does the Spatial Confinement Influence the Electric Properties and Cooperative Effects of the Hydrogen Bonded Systems? HCN Chains as a Case Study. *Chem. Phys. Lett.* 608, 264–268. doi:10.1016/j.cplett.2014.05.102

Shameema, O., Ramachandran, C. N., and Sathyamurthy, N. (2006). Blue Shift in X−H Stretching Frequency of Molecules Due to Confinement. *J. Phys. Chem. A.* 110, 2–4. doi:10.1021/jp056027s

Song, Y., Liu, Z., Mao, H.-k., Hemley, R. J., and Herschbach, D. R. (2005). High-Pressure Vibrational Spectroscopy of Sulfur Dioxide. *J. Chem. Phys.* 122, 174511. doi:10.1063/1.1883405

Zaleśny, R., Góra, R. W., Kozłowska, J., Luis, J. M. J., Ågren, H., and Bartkowiak, W. (2013). Resonant and Nonresonant Hyperpolarizabilities of Spatially Confined Molecules: A Case Study of Cyanoacetylene. *J. Chem. Theor. Comput.* 9, 3463–3472. doi:10.1021/ct400410m

Zaleśny, R., Góra, R. W., Luis, J. M., and Bartkowiak, W. (2015). On the Particular Importance of Vibrational Contributions to the Static Electrical Properties of Model Linear Molecules under Spatial Confinement. *Phys. Chem. Chem. Phys.* 17, 21782–21786. doi:10.1039/c5cp02865e

Zaleśny, R., Chołuj, M., Kozłowska, J., Bartkowiak, W., and Luis, J. M. (2017). Vibrational Nonlinear Optical Properties of Spatially Confined Weakly Bound Complexes. *Phys. Chem. Chem. Phys.* 19, 24276–24283. doi:10.1039/C7CP04259K

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 24 January 2022
doi: 10.3389/fchem.2021.801551

Accelerate the Electrolyte Perturbed-Chain Statistical Associating Fluid Theory–Density Functional Theory Calculation With the Chebyshev Pseudo-Spectral Collocation Method. Part II. Spherical Geometry and Anderson Mixing

Yunhao Sun[1,2], *Zhengxing Dai*[3], *Gulou Shen*[4], *Xiaohua Lu*[3], *Xiang Ling*[1]* and *Xiaoyan Ji*[2]*

[1]Jiangsu Key Laboratory of Process Enhancement and New Energy Equipment Technology, School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing, China, [2]Division of Energy Science/Energy Engineering, Lulea° University of Technology, Lulea°, Sweden, [3]State Key Laboratory of Materials-Oriented Chemical Engineering, Nanjing Tech University, Nanjing, China, [4]National and Local Joint Engineering Research Center for Deep Utilization Technology of Rock-Salt Resource, Faculty of Chemical Engineering, Huaiyin Institute of Technology, Huaian, China

To improve the efficiency of electrolyte perturbed-chain statistical associating fluid theory–density functional theory (ePC-SAFT-DFT) calculation of the confined system, in this work, first, the Chebyshev pseudo-spectral collocation method was extended to the spherical pores. Second, it was combined with the Anderson mixing algorithm to accelerate the iterative process. The results show that the Anderson mixing algorithm can reduce the computation time significantly. Finally, based on the accelerated ePC-SAFT-DFT program, a systematic study of the effects of the temperature, pressure, pore size, and pore shape on the CO_2 solubilities in the ionic liquids (ILs) confined inside the silica nanopores was conducted. Based on the simulation results, to obtain high CO_2 solubilities in the ILs confined in silica, a better option is to use the silica material with a narrow spherical pore, and the IL-anion should be selected specifically considering that it has a more significant impact on the absorption enhancement effect.

Keywords: density functional theory, algorithm, ionic liquids, CO_2, electrolyte perturbed-chain statistical associating fluid theory (ePC-SAFT)

OPEN ACCESS

Edited by:
Heribert Reis,
National Hellenic Research Foundation, Greece

Reviewed by:
Rolf Stierle,
University of Stuttgart, Germany
Ranjan Dey,
Birla Institute of Technology and Science, India

***Correspondence:**
Xiang Ling
xling@njtech.edu.cn
Xiaoyan Ji
xiaoyan.ji@ltu.se

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *25 October 2021*
Accepted: *14 December 2021*
Published: *24 January 2022*

Citation:
Sun Y, Dai Z, Shen G, Lu X, Ling X and Ji X (2022) Accelerate the Electrolyte Perturbed-Chain Statistical Associating Fluid Theory–Density Functional Theory Calculation With the Chebyshev Pseudo-Spectral Collocation Method. Part II. Spherical Geometry and Anderson Mixing. Front. Chem. 9:801551. doi: 10.3389/fchem.2021.801551

1 INTRODUCTION

Mitigating CO_2 emission from fossil-fueled power plants as well as from transports has become an urgent and worldwide research topic, in which CO_2 separation is often needed (MacDowell et al., 2010; Boot-Handford et al., 2014). Ionic liquids (ILs) are promising absorbents for CO_2 separation due to their extremely low vapor pressure, high CO_2 solubility, as well as low-energy usage for solvent regeneration (Brennecke and Gurkan, 2010; Ramdin et al., 2012; Zhang et al., 2012). However, the viscosity of pure ILs is relatively high compared with common organic solvents, causing a significant decrease in the mass and heat-transfer rates. Using supported ILs has been proposed as a promising solution for practical applications. This can take the

advantage of high gas selectivity in ILs, and also, the high surface area of the supported materials can reduce the impact of high viscosity, improve the gas transfer, and hence increase the absorption rate (Zhang et al., 2006; Ren et al., 2012; Romanos et al., 2014).

Research has been conducted to address the confinement effect on the gas solubility in ILs via experiments and molecular simulations (Baltus et al., 2005; Ilconich et al., 2007; Ilconich et al., 2007; Zhang et al., 2010; Iarikov et al., 2011; Ren et al., 2012; Banu et al., 2013; Shi and Luebke, 2013; Romanos et al., 2014; Santos et al., 2014; Budhathoki et al., 2017; Shen and Hung, 2017). According to previous research, several factors will affect the CO_2 solubility inside the confined ILs, for example, temperature, pressure, pore size, and shape of porous materials (Baltus et al., 2005; Zhang et al., 2006; Ilconich et al., 2007; Zhang et al., 2010; Iarikov et al., 2011; Ren et al., 2012; Banu et al., 2013; Shi and Luebke, 2013; Romanos et al., 2014; Santos et al., 2014; Budhathoki et al., 2017; Shen and Hung, 2017). However, to screen a suitable IL, optimizing the structure of supported material and operation conditions by experiment or molecular simulations is time and cost consuming, considering the fact that the huge number (10^{18}) of possible ILs can be synthesized (Wasserscheid and Thomas, 2008), as well as the wide temperature and/or pressure range in applications. Therefore, it is desirable to develop a theoretical model to predict the properties of confined IL–CO_2 systems.

The classical density functional theory (DFT) is considered as an efficient theoretical method for studying the confined properties (Tripathi and Chapman, 2005; Qiao et al., 2018; Shen et al., 2013; Xu et al., 2008; Sauer and Gross, 2017; Camacho Vergara et al., 2019). In addition, in our previous work (Ji et al., 2012; Ji et al., 2014; Shen et al., 2015; Ji and Held, 2016; Sun et al., 2019), electrolyte-perturbed-chain statistical associating fluid theory (ePC-SAFT) (Cameretti et al., 2005) has been developed to represent the thermodynamic properties of IL systems. Moreover, the developed ePC-SAFT has been combined with DFT (ePC-SAFT-DFT) to describe the properties of IL and CO_2/IL confined in nanopores with acceptable results (Shen et al., 2018). Recently, in order to calculate the properties of the confined ILs with ePC-SAFT-DFT efficiently, the Chebyshev pseudo-spectral collocation method (Yatsyshin et al., 2012; Nold et al., 2017) was implemented to accelerate the ePC-SAFT-DFT calculation (Sun et al., 2021). However, only the slit-shaped and cylindrical pores have been considered previously, while for the spherical cavity, the corresponding method has not been available. In addition, an advanced iteration method is required for replacing the simple Picard iteration to accelerate ePC-SAFT-DFT calculation further. Anderson mixing is an elaborate iteration method that has been used in the work by Mairhofer and Gross (2017) and Shen et al. (2021), showing desirable performance in accelerating DFT computing. Therefore, replacing the Picard iteration with Anderson mixing can be an effective strategy.

In this work, the Chebyshev pseudo-spectral collocation method was extended to the spherical geometry, where an expression of 9-3 Lennard–Jones potential for the spherical cavity was derived. In addition, Anderson mixing was used to accelerate the ePC-SAFT-DFT calculation further. Based on the modified program, the CO_2 solubility of ILs confined in the silica nanopores was chosen as the representative to conduct the investigation, considering silica is a promising supporting material for ILs (Shi and Luebke, 2013), and the effects of temperature, pressure, pore structures, and IL-ions were investigated systematically.

2 THEORY

2.1 Electrolyte Perturbed-Chain Statistical Associating Fluid Theory-Density Functional Theory

According to DFT, in the presence of a solid surface, the grand potential Ω at grand canonical ensemble is given by the equation:

$$\Omega[\rho_i(\boldsymbol{r})] = A[\rho_i(\boldsymbol{r})] - \sum_i \int d\boldsymbol{r}'[\rho_i(\boldsymbol{r}')(\mu_i - m_i V_{i,ext}(\boldsymbol{r}'))] \quad (1)$$

where A is the Helmholtz free energy, $\rho_i(\boldsymbol{r})$ is the molecular density of component i at position $\boldsymbol{r}$, μ_i is the chemical potential, m_i is the number of segments in a chain for component i, and $V_{i,ext}(\boldsymbol{r}')$ is the nonelectrostatic external field acting on the segment of component i. The Helmholtz free energy A in ePC-SAFT-DFT for a system can be expressed as (Qiao et al., 2018; Shen et al., 2018):

$$A[\rho_i(\boldsymbol{r})] = A^{id}[\rho_i(\boldsymbol{r})] + A^{hs}[\rho_i(\boldsymbol{r})] + A^{chain}[\rho_i(\boldsymbol{r})] + A^{disp}[\rho_i(\boldsymbol{r})] + A^{ion}[\rho_i(\boldsymbol{r})] \quad (2)$$

where A^{id} is the ideal free energy, A^{hs}, A^{chain}, A^{disp}, and A^{ion} are the excess free energies due to hard-sphere repulsions, chain connectivity, dispersive, and electrostatic interactions, respectively. The model performance has been verified in a previous work (Shen et al., 2018), where the model predictions in the density profiles of ionic fluids in the charged pores were compared with the molecular simulation results with high consistency.

The details of A^{hs}, A^{chain}, and A^{disp} have been described elsewhere (A^{hs}: Barker and Henderson, 1967; Rosenfeld, 1989; Roth et al., 2002; Yu and Wu, 2002a; Yu and Wu, 2002b; A^{chain}: Tripathi and Chapman, 2005; A^{disp}: Shen et al., 2013). The A^{ion} is composed of two terms, the Coulomb term (A^{col}) and the Debye–Hückel term (A^{DH}). A^{DH} accounts for the short-range electrostatic interaction caused by the inhomogeneous distribution at the nearby region of ions, which is based on the Debye–Hückel (DH) theory. A^{col} accounts for the Helmholtz free energy caused by the long-range electrostatic interaction due to the unsymmetrical distribution of cation and anion. The expressions of A^{col} and A^{DH} can be found in Li and Wu (2006) and Shen et al. (2018), respectively. All these terms are involved in the ePC-SAFT-DFT calculations for the IL–CO_2 systems.

Minimization of the grand potential with respect to the density profile of component i yields the following Euler–Lagrange equation:

$$\frac{\delta\Omega[\rho(\boldsymbol{r})]}{\delta\rho_i(\boldsymbol{r}')} = \sum_c \frac{\delta A^c[\rho(\boldsymbol{r})]}{\delta\rho_i(\boldsymbol{r}')} + q_i\psi(\boldsymbol{r}') - (\mu_i - m_i V_{i,ext}(\boldsymbol{r}')) = 0 \quad (c = id,\ hs,\ chain,\ disp,\ DH) \tag{3}$$

where q_i is the charge of component i, and $\psi(\boldsymbol{r})$ is the mean electric potential at position $\boldsymbol{r}$, which can be obtained by solving the Poisson's equation (Li and Wu, 2006). The expressions of functional derivatives $\frac{\delta A^c[\rho(\boldsymbol{r})]}{\delta\rho_i(\boldsymbol{r})}$ are described in Rosenfeld et al. (1997), Gross (2009), Roth (2010), and Sun et al. (2021).

2.2 Evaluation of the Convolution-Like Integrals for Spherical Cavity

The expressions of the functional derivatives $\frac{\delta A^c[\rho(\boldsymbol{r})]}{\delta\rho_i(\boldsymbol{r})}$ are described in Rosenfeld et al. (1997), Gross (2009), Roth (2010), and Sun et al. (2021). All the convolution-like integrals involved in ePC-SAFT-DFT calculation can be classified as:

$$\begin{cases} I_1(\boldsymbol{r}) = \int d\boldsymbol{r}' f(\boldsymbol{r}')\delta(R_c - |\boldsymbol{r} - \boldsymbol{r}'|) \\ I_2(\boldsymbol{r}) = \int d\boldsymbol{r}' f(\boldsymbol{r}')\dfrac{\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'}{|\vec{\boldsymbol{r}} - \boldsymbol{r}'|}\delta(R_c - |\boldsymbol{r} - \boldsymbol{r}'|) \\ I_3(\boldsymbol{r}) = \int d\boldsymbol{r}' f(\boldsymbol{r}')\theta(R_c - |\boldsymbol{r} - \boldsymbol{r}'|) \end{cases} \tag{4a}$$

where R_c represents the weighting distances, which are:

$$\begin{cases} R_c = \dfrac{d_i}{2}, & for\ the\ hard\ sphere\ and\ association\ term \\ R_c = d_i, & for\ the\ chain\ term \\ R_c = \lambda\sigma_i, & for\ the\ dispersion\ term \\ R_c = \dfrac{d_{i,DH}}{2}, & for\ the\ DH\ term \end{cases} \tag{4b}$$

In **Eq. 4b**, θ is the Heaviside step function, and δ is the Dirac delta function. The σ_i and d_i in **Eq. 4b** represent the temperature-independent and -dependent hard sphere segment diameter (Barker and Henderson, 1967), respectively. The $d_{i,DH}$ is the weighting distance for the Debye–Hückel term, which can be referred to in Shen et al. (2018). The λ was set as 1.5 in order to be consistent with the bulk PC-SAFT model (Gross and Sadowski, 2001). The convolution integrals in Eq. 7a map $f(\boldsymbol{r}')$ to $I(\boldsymbol{r})$ with these integral kernels.

For spherical cavities, $f(\boldsymbol{r}')$ and $I(\boldsymbol{r})$ only vary with the radial direction, and the analytic expressions of $I(r)$ reads (Groh and Mulder, 2000):

$$\begin{cases} I_1(r) = \dfrac{2\pi R_c}{r}\displaystyle\int_{|r-R_c|}^{r+R_c} dr' f(r')r' \\ I_2(r) = \dfrac{\pi}{r^2}\vec{\boldsymbol{e}}_r\displaystyle\int_{|r-R_c|}^{r+R_c} dr' f(r')r'(r^2 - r'^2 + R_c^2) \\ I_3(r) = \dfrac{\pi}{r}\displaystyle\int_{|r-R_c|}^{r+R_c} dr' f(r')r'\left[R_c^2 - (r-r')^2\right] + 4\pi\theta(R_c - r)\int_0^{R_c - r} dr' f(r')r'^2 \end{cases} \tag{5}$$

2.3 Evaluation of Mean Electric Potential Distribution Inside Spherical Cavity

In the spherical symmetric distribution, the Poisson's equation reduces to:

$$\frac{d^2[r^2\psi(r)]}{r^2 dr^2} = -\frac{1}{\varepsilon_0\varepsilon_r}q(r) \tag{6}$$

where $q(r)$ is the total charge at position r.

For a spherical cavity with diameter R, by solving **Eq. 6** with the boundary conditions $\psi(R) = \psi^w$ and $\frac{d\psi(0)}{dr} = 0$, the expression for the mean electric potential reads:

$$\psi(r) = \psi^w + \frac{1}{\varepsilon_0\varepsilon_r}\left[\frac{R-r}{Rr}\int_0^r dr' q(r')r'^2 + \int_r^R dr'\frac{R-r'}{R}q(r')r'\right] \tag{7}$$

According to **Eq. 7**, the boundary electric potential ψ^w is required, which corresponds to a specific charge density on the internal surface of a nanopore. This value can be obtained *via* "trial and error" until the charge of ions in the nanopore is equal in magnitude to that with the opposite sign on the internal surface of the nanopore.ü

2.4 Chebyshev Pseudo-Spectral Collocation Method

The Chebyshev pseudo-spectral collocation method for DFT modeling was developed by Yatsyshin et al. (2012). In the Chebyshev pseudo-spectral collocation method, for one-dimensional DFT calculation with an N-point discretization scheme, the density or the weighted density profile over the whole computation domain is determined from the density or the weighted density at a prescribed set of collocation points $\{z_k\}$, $k = 1,2\ \ldots,\ N$ using the barycentric form (Baltensperger, 2002; Berrut and Trefethen, 2004):

$$\rho(z) = \frac{\sum_1^{'N}\frac{(-1)^k}{z - z_k}\rho(k)}{\sum_1^{'N}\frac{(-1)^k}{z - z_k}} \tag{8}$$

where the primes indicate that the first and last terms in the sums are divided by 2. The collocation points (z_k) can be obtained by a conformal map from the Chebyshev collocation points (x_k) (Baltensperger, 2002; Berrut and Trefethen, 2004):

$$x_k = cos\frac{(k-1)\pi}{N-1} \tag{9}$$

The integrals associated with the DFT calculation can be evaluated with the Clenshaw–Curtis quadrature (Clenshaw and Curtis, 1960).

2.5 Implementation of Chebyshev Pseudo-Spectral Collocation Method in Electrolyte Perturbed-Chain Statistical Associating Fluid Theory-Density Functional Theory for Spherical Cavity

In the ePC-SAFT-DFT calculation, three domains need to be discretized for interpolation. The first domain is for the density profile, the second one is for the weighted density function profiles in the hard-sphere term, and the third one is for the weighted density function profile in the Debye–Hückel term.

For the spherical cavity with diameter R, we defined a diameter R' as:

$$R' = R - \sigma_s \tag{10}$$

The density profile is considered in the domain (0, R') based on the coordinate system illustrated in **Figure 1**. In other words, this domain is discretized for interpolating the density profile function. In general, the density profile vibrates dramatically near the wall, which implies that more collocation points are required in these regions compared with the middle of the nanopore. In this work, the conformal map proposed by Bayliss and Turkel was used to map the Chebyshev collocation points $\{x_k\}$ to domain (0, R') (Bayliss and Turkel, 1992):

$$z_k = \frac{R'}{2}\left\{\beta + \frac{1}{\alpha} tan\left[\lambda\left(x_k - \mu\right)\right]\right\} + \frac{R'}{2} \tag{11a}$$

The λ and μ in **Eq. 11a** read:

$$\begin{cases} \lambda = \dfrac{\gamma + \eta}{2} \\ \mu = \dfrac{\gamma - \eta}{\gamma + \eta} \end{cases} \tag{11b}$$

where:

$$\begin{cases} \gamma = \arctan\left[\alpha\left(1 + \beta\right)\right] \\ \eta = \arctan\left[\alpha\left(1 - \beta\right)\right] \end{cases} \tag{11c}$$

Equation 11a maps the Chebyshev points in the transformed coordinate into the points that cluster in the physical coordinate near $\frac{R'}{2}(\beta + 1)$ with a density that is large when α is large. In this work, the β was set as:

$$\beta = 1 - \frac{\bar{d}}{R'} \tag{12}$$

where $\bar{d}$ is the mean diameter of all components.

Therefore, the collocation points are clustered in the physical coordinate near $(R' - \frac{\bar{d}}{2})$, where high peaks have occurred in the nearby region. The α was set as 2 in this work. Similar approaches were used for the discretization of another two domains, i.e., the

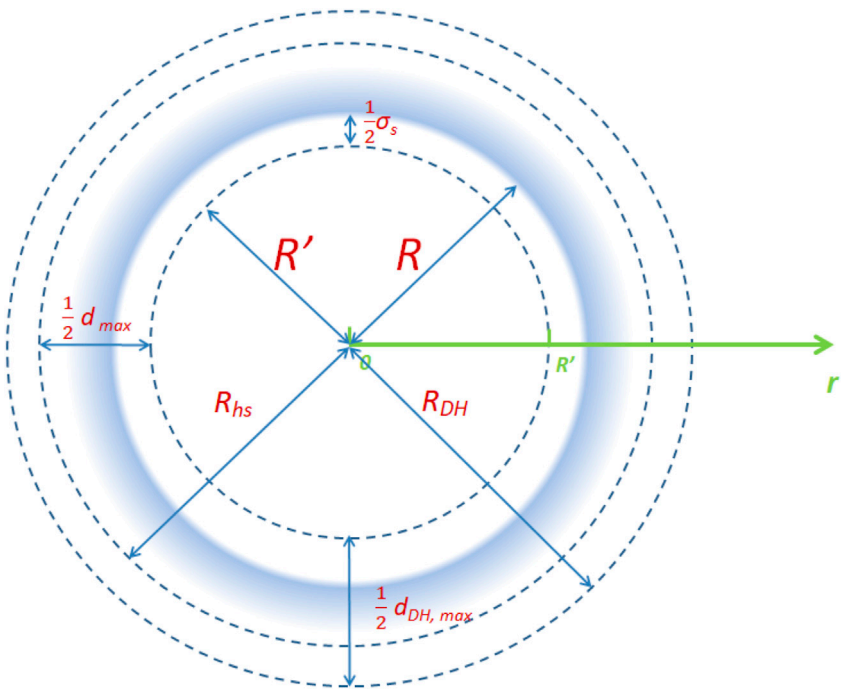


FIGURE 1 | Schematic diagram of the domains considered in a spherical cavity.

domains for weighted density function profiles of hard-sphere term (0, R_{hs}) and Debye–Hückel term (0, R_{DH}).

2.5.1 Evaluation of Convolution-Like Integral

As pointed out by Yatsyshin et al. (2012), the maps (**Eq. 4**) involved in DFT calculation can be carried out by matrix–vector products with discrete data based on interpolation (**Eq. 8**) and Clenshaw–Curtis quadrature. In the spherical cavities, for a map with original space discretized into N_2 points and the image space discretized into N_1 points, the map can be represented by $N_1 \times N_2$ matrix (Yatsyshin et al., 2012):

$$\boldsymbol{J}_{ij} = \begin{cases} \dfrac{\boldsymbol{J}'_{ij}}{2} & j = 1, N_2 \\ \boldsymbol{J}'_{ij} & j \in [2, N_2 - 1] \end{cases} \tag{13a}$$

with

$$\boldsymbol{J}'_{ij} = \frac{(b_i - a_i)(-1)^j}{2} \sum_{q=1}^{M} \frac{\hat{w}_q W\left(z_i - y_{qi}\right)}{y_{qi} - z_j} \Big/ {\sum}'_{m=1}^{N_2} \frac{(-1)^m}{y_{qi} - z_m} \tag{13b}$$

where a_i and b_i are the lower and upper limits of integral for the ith discrete point, W is the corresponding integral kernel in **Eq. 5**, and y_{qi} is the qth Chebyshev grid of the ith discrete point in the image space:

$$y_{qi} = a_i + (b_i - a_i) x_q \tag{13c}$$

$\hat{w}_q$ is the corresponding Clenshaw–Curtis weight, and the values at M Chebyshev grids are used to evaluate each integral. In this work, M was set as 45.

For the singularities in the expression of **Eq. 13b** (i.e., $y_{qi} = z_k$), the corresponding elements can be evaluated based on the following expression (Sun et al., 2021):

$$\lim_{y_{qi} \to Z_k} \frac{\frac{\hat{w}_q}{y_{qi} - Z_j}}{{\sum}'_{m=1}^{N_2} \frac{(-1)^m}{y_{qi} - Z_m}} = \begin{cases} 0 & j \neq k \\ 2 & j = k = 1 \\ 2(-1)^{N_2} & j = k = N_2 \\ (-1)^k & 1 < j = k < N_2 \end{cases} \tag{13d}$$

2.5.2 Evaluation of the Mean Electric Potential

Equation 7 can be represented with matrix–vector products with the discrete data:

$$\vec{\psi} = \frac{1}{\varepsilon_0 \varepsilon_r} \left(\boldsymbol{D}_1 \cdot \vec{q} + \boldsymbol{D}_2 \cdot \vec{q} \right) + \psi^w \tag{14a}$$

where $\vec{\psi}$ and $\vec{q}$ are composed of ψ and q at the position of collocation points, respectively.

The two $N \times N$ matrixes $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ read:

$$\begin{cases} \boldsymbol{D}_1 = \boldsymbol{L}' \cdot \boldsymbol{Q}_1 \\ \boldsymbol{D}_2 = \boldsymbol{U}' \cdot \boldsymbol{Q}_2 \end{cases} \tag{14b}$$

where $\boldsymbol{L}'$ and $\boldsymbol{U}'$ are composed of:

$$\begin{cases} \boldsymbol{L}' = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{L} \end{bmatrix} \\ \boldsymbol{U}' = \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{0} \end{bmatrix} \end{cases} \tag{14c}$$

where $\boldsymbol{L}$ and $\boldsymbol{U}$ are the lower and upper triangular matrices with all elements being unity, respectively. The elements in the $(N\text{-}1) \times N$ matrix $\boldsymbol{Q}_1$ read:

$$\boldsymbol{Q}_{1,ij} = \begin{cases} \dfrac{\boldsymbol{Q}'_{1,ij}}{2} & j = 1, N \\ \boldsymbol{Q}'_{1,ij} & j \in [2, N-1] \end{cases} \tag{14d}$$

with

$$\boldsymbol{Q}'_{1,ij} = \frac{(R' - Z_{i+1})(b_i - a_i)(-1)^j}{2R' Z_{i+1}} \sum_{q=1}^{M} \frac{\frac{\hat{w}_q y_{qi}^2}{y_{qi} - Z_j}}{{\sum}'_{m=1}^{N_2} \frac{(-1)^m}{y_{qi} - Z_m}} \tag{14e}$$

The elements in $(N\text{-}1) \times N$ $\boldsymbol{Q}_2$ read:

$$\boldsymbol{Q}_{2,ij} = \begin{cases} \dfrac{\boldsymbol{Q}'_{2,ij}}{2} & j = N \\ \boldsymbol{Q}'_{2,ij} & j \in [1, N-1] \end{cases} \tag{14f}$$

with

$$\boldsymbol{Q}'_{2,ij} = \frac{(b_i - a_i)(-1)^j}{2R'} \sum_{q=1}^{M} \frac{\frac{\hat{w}_q \left(R' - y_{qi}\right) y_{qi}}{y_{qi} - z_j}}{{\sum}'_{m=1}^{N_2} \frac{(-1)^m}{y_{qi} - Z_m}} \tag{14g}$$

where

$$\begin{cases} a_i = z_i \\ b_i = z_{i+1} \end{cases} \tag{14h}$$

y_{qi} can be evaluated with **Eq. 13c** with the a_i and b_i defined in **Eq. 14h**. The singularities in equations **Eq. 14e** and **Eq. 14g** can also be evaluated based on **Eq. 13d**.

Q_1 and Q_2 are the linear operators implementing the piecewise integrations between two adjacent collocation points, while $\boldsymbol{L}'$ and $\boldsymbol{U}'$ are used to sum up the results of these piecewise integrations for obtaining the final integration results of **Eq. 7**.

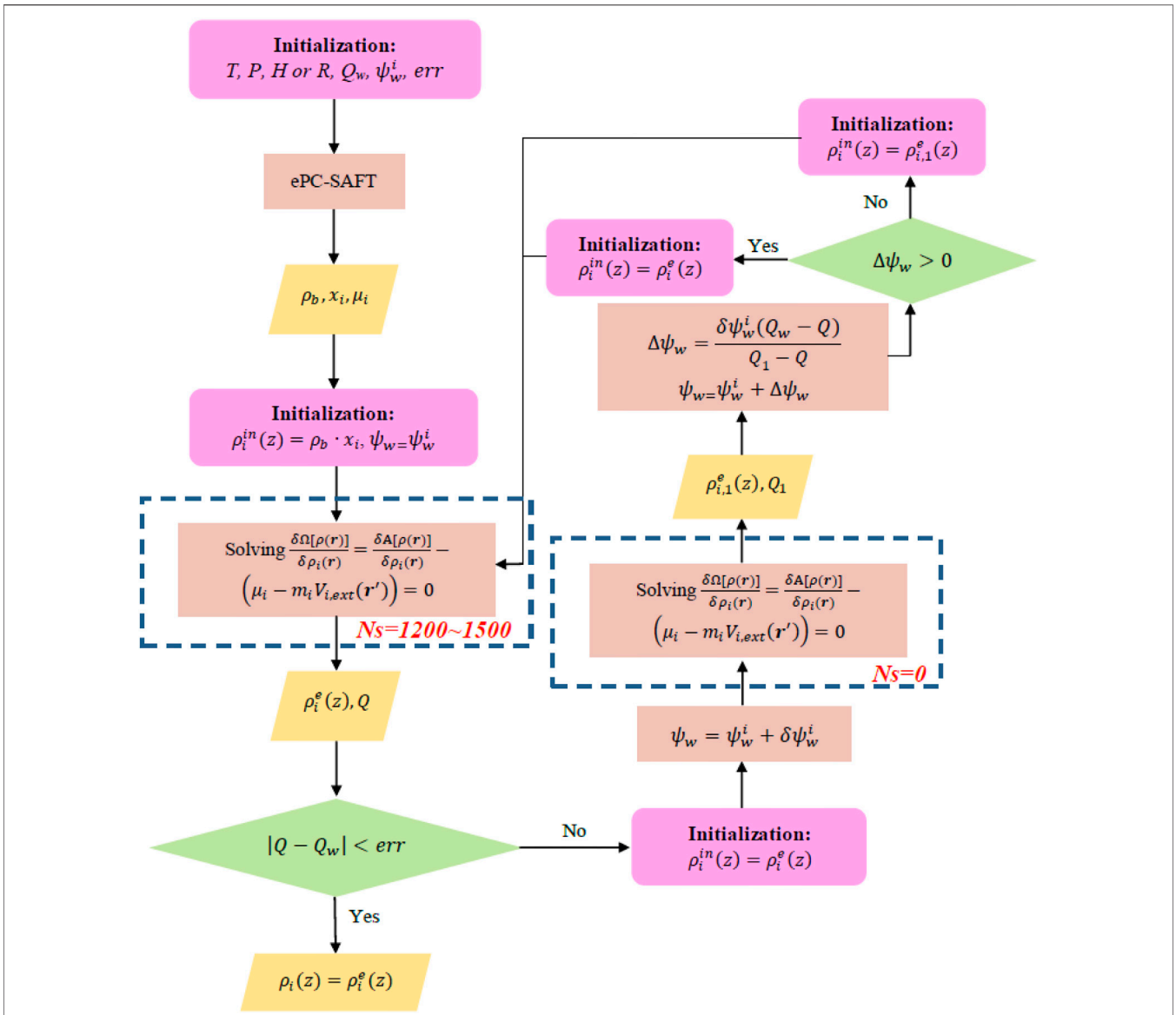
2.6 Solving Euler–Lagrange Equation With Anderson Mixing

In the Chebyshev pseudo-spectral collocation method, the Euler–Lagrange equation (**Eq. 3**) is transformed into a vector equation that can be solved numerically. In our previous work, the Picard iteration was used. In this work, the Anderson mixing is implemented to solve **Eq. 3** iteratively (Anderson, 1965). In the Anderson mixing, the Euler–Lagrange equation can be rewritten as:

$$\begin{aligned} \rho_i(z) &= \rho_{i,b} \, exp\left(-\frac{V_{i,ext}(z)}{kT} + \frac{\mu_i^{res} + \sum_c \frac{\delta A^c}{\delta \rho_i}(z) + q_i \psi(z)}{kT m_i} \right) \\ &= g[\rho_i, z](c = hs, chain, disp, DH) \end{aligned} \tag{15a}$$

where k represents the Boltzmann constant, $\rho_{i,b}$ represents the bulk phase density of component i, and the superscript *res* represents the residual quantities.

Eq. 15a can be solved iteratively by (Anderson, 1965):



SCHEME 1 | Calculating the density profile of ionic liquid (IL)–CO_2 system inside nanopores with specific surface charge Q_w using the general scheme combined with the Anderson mixing.

$$\rho_i^{k+1}(z) = (1-S)\sum_{j=0}^{m^k}\gamma_j^k\rho_i^{k-m^k+j}(z) + S\sum_{j=0}^{m^k}\gamma_j^k g\left[\rho_i^{k-m^k+j}, z\right] \tag{15b}$$

where S is the relaxing factor, $m^k = \min(m, k)$, m was set as 50 in this work, the subscript i represents the ith component, and γ_j^k is determined by:

$$\gamma_j^k = \min_{\gamma_0^k,\cdots\gamma_{m^k}^k} \sqrt{\sum_i\sum_z\sum_{j=0}^{m^k}\gamma_j^{k2}\left\{g\left[\rho_i^{k-m^k+j}, z\right] - \rho_i^{k-m^k+j}(z)\right\}^2} \; s.t. \sum_{j=0}^{m^k}\gamma_j^k = 1 \tag{15c}$$

The constrained optimization can be transformed to unconstrained optimization (Fang and Saad, 2008; Walker and Ni, 2011):

$$\omega_j^k = \min_{\omega_0^k,\cdots\omega_{m^k-1}^k} \sqrt{\sum_i\sum_z\left\{\mathcal{F}_i^k(z) - \sum_{j=0}^{m^k-1}\omega_j^k\left[\mathcal{F}_i^{k-m^k+j+1}(z) - \mathcal{F}_i^{k-m^k+j}(z)\right]\right\}^2} \tag{15d}$$

where

$$\omega_j^k = \begin{cases} \gamma_j^k - \gamma_{j-1}^k & j > 1 \\ \gamma_0^k & j = 0 \end{cases} \quad and \; \mathcal{F}_i^n(z) = g\left[\rho_i^n, z\right] - \rho_i^n(z) \tag{15e}$$

The successive optimization (**Eq. 15d**) can be solved efficiently by the updating QR factorization, and the necessary Matlab code is given in Walker (2011). In order to avoid divergence, *Ns* steps of Picard iterations can be performed at the beginning and then switched to the Anderson mixing procedure.

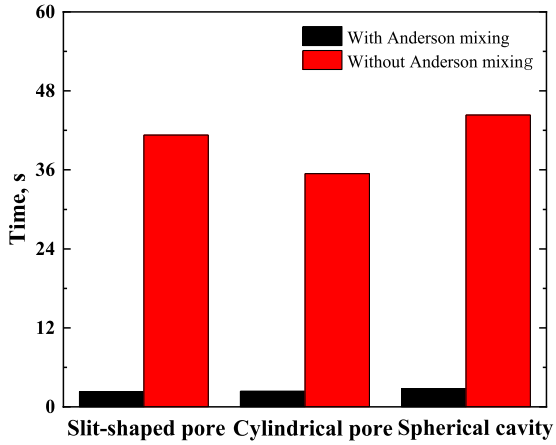


FIGURE 2 | Time required for calculating equilibrium density profile (not including the time for calculating the matrices required in the Chebyshev pseudo-spectral collocation method).

TABLE 1 | Time needed for calculating the matrices required in the Chebyshev pseudo-spectral collocation method.

	Slit-shaped pore	Cylindrical pore	Spherical cavity
Time, s	0.19	39.05	0.11

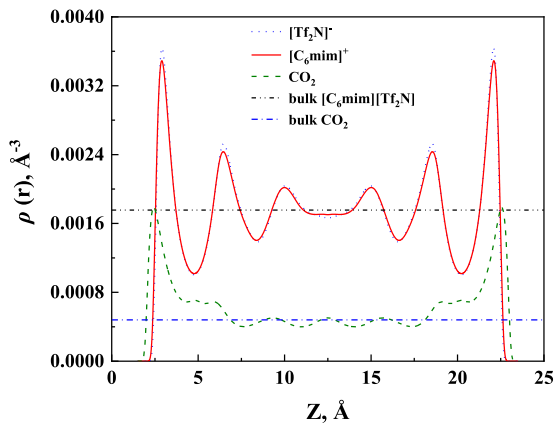


FIGURE 3 | The calculated density profiles of [C_6mim] [Tf_2N]-CO_2 system confined in 25 Å slit silica pore at 10 bar and 323.15 K.

2.7 General Scheme Combined With Anderson Mixing

A general scheme for calculating the density profile of confined systems with ionic contribution was proposed in Sun et al. (2021), as illustrated in **Scheme 1**. In a loop of **Scheme 1**, the Euler–Lagrange equation needs to be solved twice. For the first time, solving the Euler–Lagrange equation in a loop, *Ns* (steps of Picard iteration performed before the Anderson mixing procedure) can be set between 1,200 and 1,500, while for the second time, due to the initial guess of density profile being not far away from the equilibrium density profile, we do not perform Picard iterations prior to the Anderson mixing procedure. (i.e., $Ns = 0$).

In **Scheme 1**, T, p, μ_i, and x_i refer to those in the bulk phase, which were obtained with ePC-SAFT. H (or R) and Q_w are the features of nanopores. ρ^{in} represents the initial guess of the density profile, and ρ^{e} represents the calculated equilibrium density profile with the boundary electric potential ψ_w. The charge absorbed per area of nanopore Q can be computed with:

$$Q = \sum_i q_i \bar{\rho}_i \tag{16a}$$

where the amount of component i adsorbed per surface area $\bar{\rho}_i$ for the slit-like pores can be evaluated by:

$$\bar{\rho}_i = \frac{\int_0^H dz \rho_i(z)}{2} \tag{16b}$$

for the cylindrical pores:

$$\bar{\rho}_i = \frac{\int_0^R dr r \rho_i(r)}{R} \tag{16c}$$

and for the spherical cavities:

$$\bar{\rho}_i = \frac{\int_0^R dr r^2 \rho_i(r)}{R^2} \tag{16d}$$

2.8 Model Ionic Liquid–CO_2 Confined in Nanopores

Following ePC-SAFT for the ILs in the bulk systems (Ji et al., 2012), an ionic liquid molecule is composed of one IL cation and one IL anion. Each individual IL ion was modeled as a nonspherical species with repulsion, dispersive attraction, and Coulomb interactions. The 9-3 Lennard–Jones potential was used to represent the nonelectrostatic interaction between silica and fluid. The nanopore was modeled as an infinitely large slit, infinitely long cylinder, or spherical cavity. The 9-3 Lennard–Jones potential for large slit and infinitely long cylinder have already been presented in the literature (Fitzgerald et al., 2003; Siderius and Gelb, 2011; Lee, 2016), while for a spherical cavity with diameter R, the 9-3 Lennard–Jones potential can be obtained from:

$$U_{s,9-3,i}(r) = \rho_{atom} \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \int_R^{+\infty} dr' 4\varepsilon_{si} \left[\left(\frac{\sigma_{si}}{\sqrt{r'^2 + r^2 - 2rr'cos\theta}} \right)^{12} - \left(\frac{\sigma_{si}}{\sqrt{r'^2 + r^2 - 2rr'cos\theta}} \right)^{6} \right] r'^2 sin\theta$$
$$= \begin{cases} 2\pi\rho_{atom}\sigma_{si}^3\varepsilon_{si} \left\{ \frac{1}{180} \left[\frac{9R - r}{r(R-r)^9} - \frac{9R + r}{r(R+r)^9} \right] \sigma_{si}^9 - \frac{8}{3} \frac{R^3}{\left(R^2 - r^2\right)^3} \sigma_{si}^3 \right\} & r > 0 \\ 2\pi\rho_{atom}\sigma_{si}^3\varepsilon_{si} \left(\frac{8}{9R^9}\sigma_{si}^9 - \frac{8}{3R^3}\sigma_{si}^3 \right) r = 0 \end{cases} \tag{17a}$$

where r is the distance of the fluid molecule from the center of the spherical cavity, ρ_{atom} is the solid atom density, σ_{si} is the

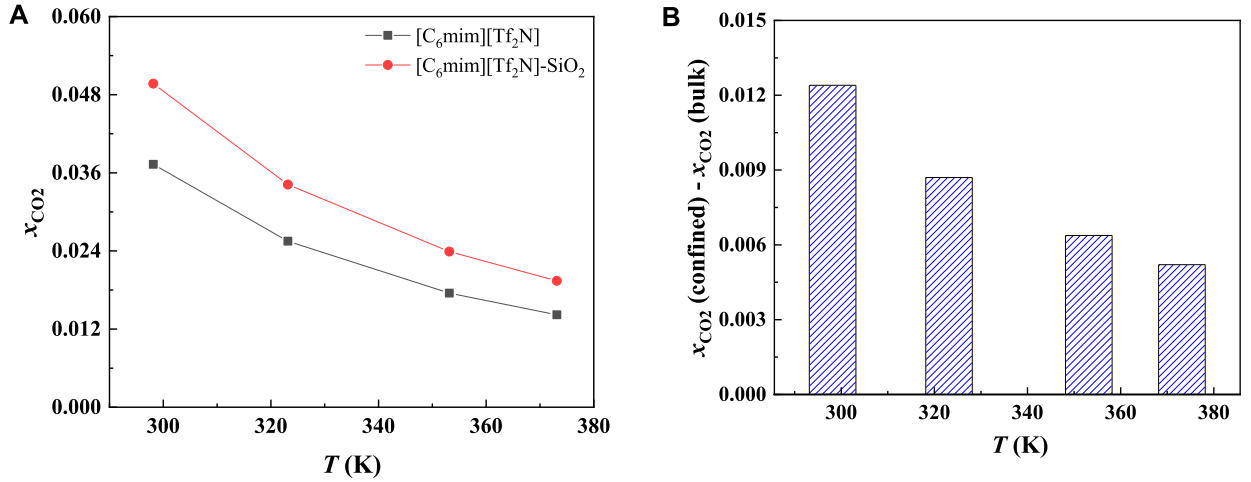


FIGURE 4 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of ILs confined in 25 Å slit-shaped pore at 1 bar and different temperatures.

effective solid-fluid diameter, and ε_{si} is the potential representing the interaction between surface and fluid segment. σ_{si} and ε_{si} can be determined with the Berthelot–Lorentz combining rules:

$$\begin{cases} \sigma_{si} = \dfrac{\sigma_i + \sigma_s}{2} \\ \varepsilon_{si} = \sqrt{\varepsilon_i \varepsilon_s} \end{cases} \tag{17b}$$

where σ_s and ε_s are the size and potential parameters of a solid surface, respectively, and ε_i is the potential parameters of fluid segment i. Here, the used potential parameters for silica are $\sigma_s = 3.0$ Å, $\varepsilon_s = 0.8\, kJ/mol$, and $4\pi\rho_{atom} = 0.5/\text{Å}^3$ (Pinilla et al., 2005). The ePC-SAFT parameters for ILs were taken from the previous work (Ji et al., 2014), while those for CO_2 were taken from Gross and Sadowski (2001).

For the electroneutral silica nanopore, the amount of cation and anion adsorbed per surface area should be equal. The solubility of CO_2 in the confined IL with a neutral surface is defined by:

$$x_{CO_2} = \frac{\bar{\rho}_{CO_2}}{\bar{\rho}_{CO_2} + \bar{\rho}_{IL}} \tag{18}$$

where $\bar{\rho}_{IL} = \bar{\rho}_{IL-cation} = \bar{\rho}_{IL-anion}$, according to the charge neutrality condition.

3 RESULTS AND DISCUSSION

3.1 Efficiency of the General Scheme Combined With Anderson Mixing

Calculating the density profile of [C_6mim] [Tf_2N]-CO_2 confined in electronic neutral silica pore with different structures at 333 K and 16.1 bar was selected as an example here to demonstrate the performance of the general scheme combined with the Anderson mixing, and the calculation efficiency was compared with the general scheme only using the Picard iterations. In the calculation, the width of the slit-shaped pore and the diameters for the cylindrical pore and spherical cavity are 5 nm. We used 120 collocation points to represent the density profile inside the slit-shaped pore (functional derivatives only need to be calculated in half of the collocation points due to the symmetry of the density profile), while 60 collocation points were used for the cylindrical pore and spherical cavity. The relaxing parameters used in all these three cases are 0.001.

The calculations were performed on a computer with an AMD core Ryzen 7 PRO 4750U and X64 processor. The version of the compiler is Matlab 2018b. **Figure 2** compares the time required for calculations.

As illustrated in **Figure 2**, using the Anderson mixing can improve the calculation efficiency significantly. The time needed for calculating the matrices required in the Chebyshev pseudo-spectral collocation method is listed in **Table 1**.

As listed in **Table 1**, the time required for calculating the matrices for the slit-shaped pore and spherical cavity can be ignored. However, the time for the cylindrical pore is pronounced. The reason is that too much time is used for evaluating the Legendre complete elliptic integral of the first and second kinds. This may be a problem in massive ePC-SAFT-DFT calculations, for example, adjusting the model parameters from experimental data, in which the equilibrium density profile needs to be solved multiple times.

However, for systems at the same temperature and pore diameter (cylindrical pore), the elements of these matrices are the same (Sun et al., 2021). Therefore, for the sake of saving computation time, when modeling systems confined in the

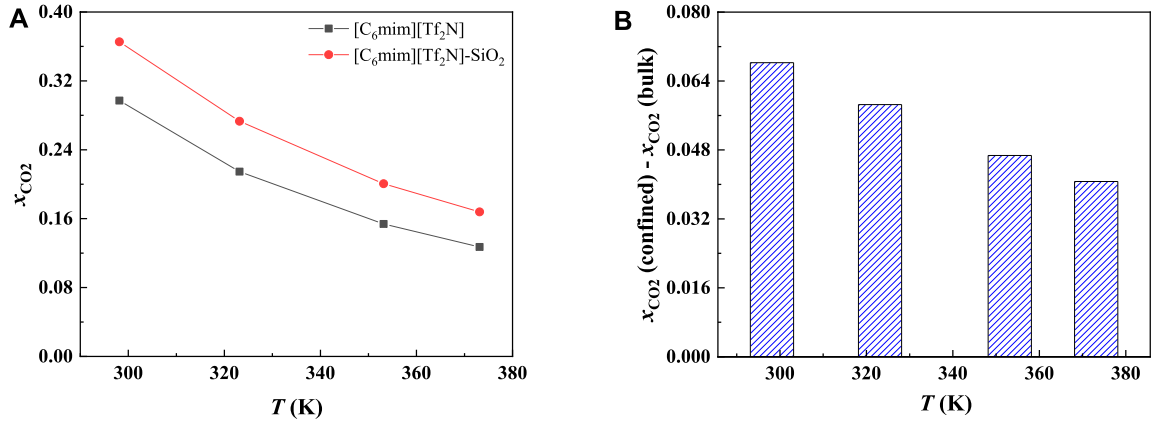


FIGURE 5 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of ILs confined in 25 Å slit-shaped pore at 10 bar and different temperatures.

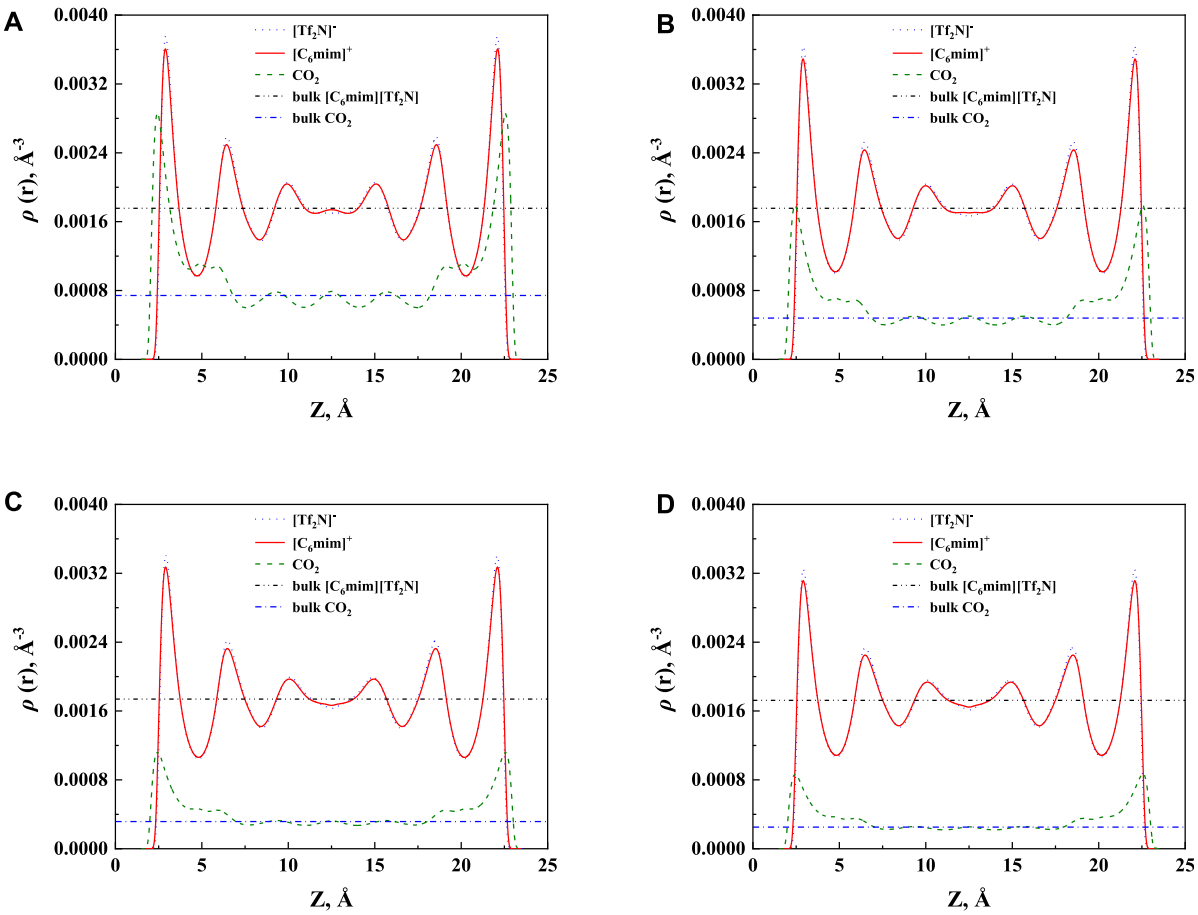


FIGURE 6 | The calculated density profiles of [C_6mim] [Tf_2N]-CO_2 confined in 25 Å slit silica pore at 10 bar and **(A)** 298.15 K, **(B)** 323.15 K, **(C)** 353.15 K, and **(D)** 373.15 K.

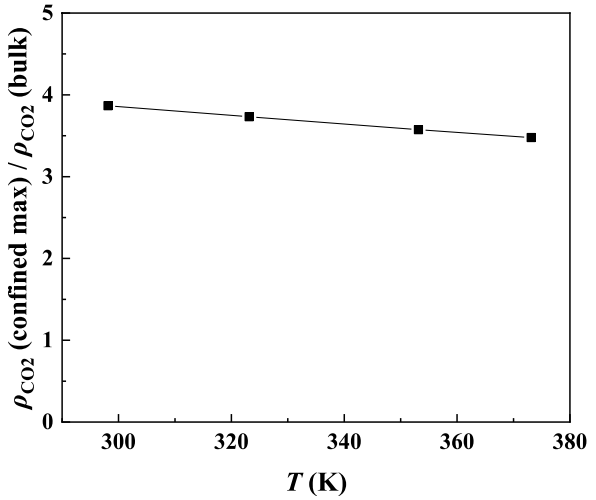


FIGURE 7 | Ratios of the maximum density of CO_2 in the confined [C_6mim] [Tf_2N] to its bulk density in [C_6mim] [Tf_2N].

cylindrical pore, these matrices can be used repeatedly for the systems at the same temperature and pore diameter.

3.2 Model CO_2 Solubilities of Confined Ionic Liquids

Based on the efficient algorithm discussed above, ePC-SAFT-DFT can be used in a wide range of IL systems to obtain results efficiently. In this work, several [$C_n$mim]-based IL-CO_2 systems confined in silica nanopore were selected to investigate the effects of temperature, pressure, IL ions, as well as the size and shape of the pore. As the nanopores of different silica materials have been roughly assumed as slit-like (Li et al., 2005; Yang and Yue, 2007), cylindrical (Kresge et al., 1992; Beck et al., 1992; Maddox et al., 1997), and spherical (Nandiyanto et al., 2009) in the previous theoretical work, in this work, these three pore-shape models were adopted.

3.2.1 General View of CO_2 Confined In Silica Nanopores

The calculated density profile of [C_6mim] [Tf_2N]-CO_2 confined in 25 Å slit silica pore at 10 bar and 323.15 K is presented in **Figure 3**.

As pointed out by Ho et al. (2013), two competitive mechanisms affect the solubility in a nanopore: One is the adsorption near the pore surface, and the other is the absorption inside the IL. These can be seen in the density profile of CO_2. According to the results illustrated in **Figure 3**, the peak in the density profile near the surface of nanopores corresponds to the first mechanism, while in the middle of the nanopore, the density profile of CO_2 tends to the bulk density, which is the reflection of the second mechanism. The absorption enhancement in the confined ILs is mainly contributed by the first mechanism. In this work, the additional solubility x^a is used to identify the absorption enhancement:

$$x^a = x_{confined} - x_{bulk} \tag{19}$$

3.2.2 The Influence of Temperature

In general, with the temperature increase, the CO_2 solubility in the bulk ILs will decrease. According to our calculation results, CO_2 solubility in [C_6mim] [Tf_2N] confined in slit-shaped pore also decreases with the increase in temperature. This is consistent with the observation by Mirzaei et al. (2017). Typical examples are shown in **Figures 3A** and **4A**. However, with the increase in temperature, the solubility of CO_2 in [C_6mim] [Tf_2N] confined in SiO_2 decreases greater than that of the bulk IL. For example, at 1 bar, when the temperature increase from 298.15 K to 373.15 K, the solubility of CO_2 in the bulk [C_6mim] [Tf_2N] decreases by

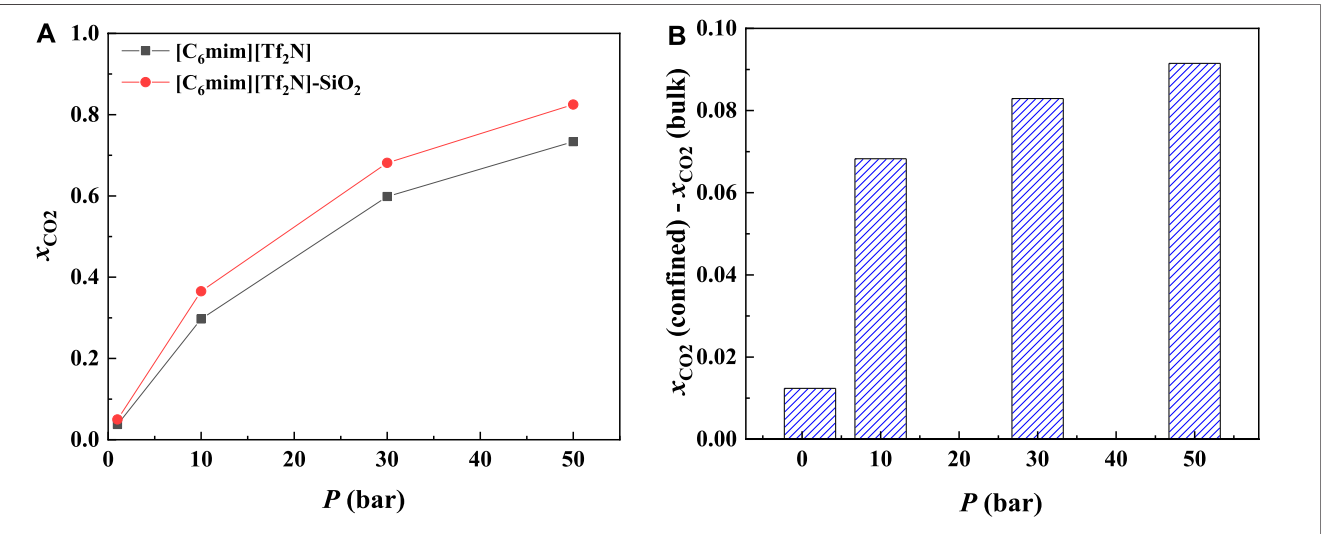


FIGURE 8 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of confined ILs in 25 Å slit-shaped pore at 298.15 K and different pressures.

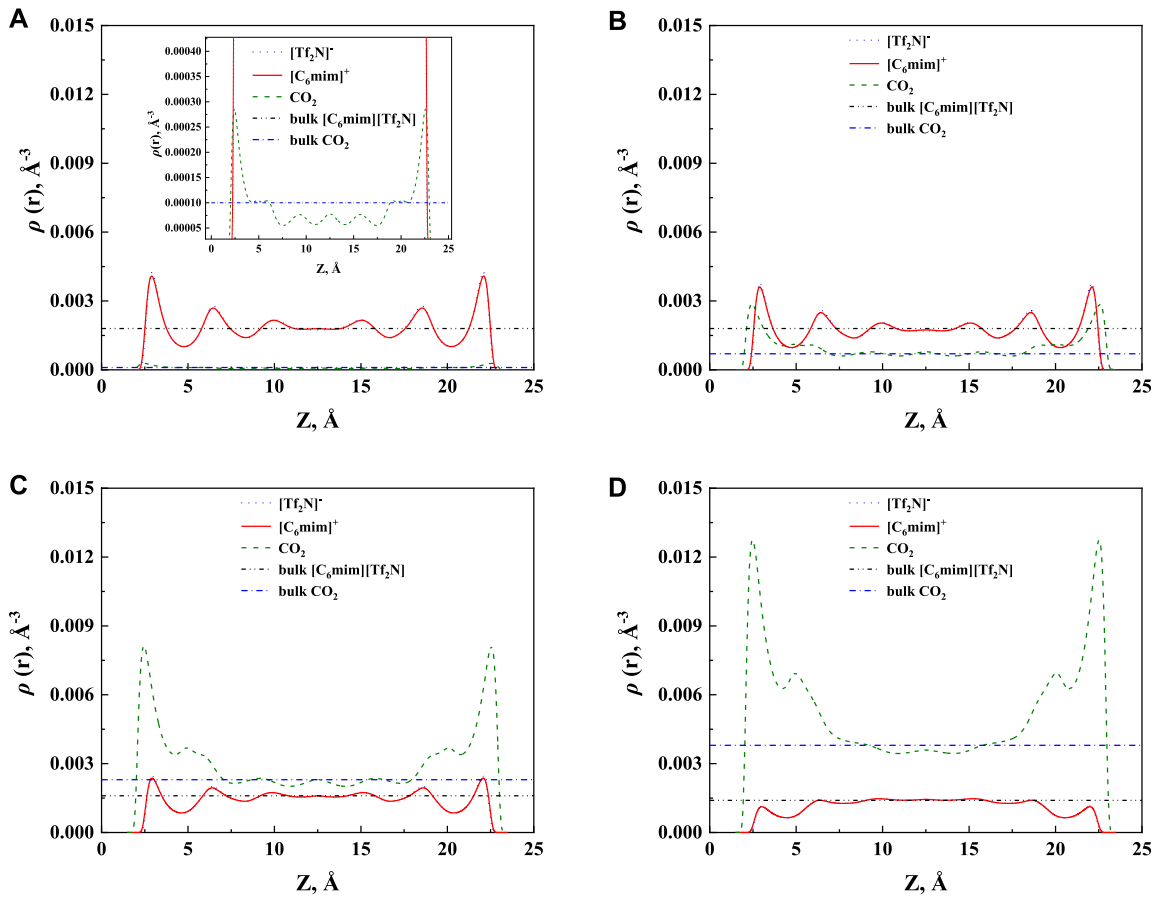


FIGURE 9 | Calculated density profiles of [C_6mim] [Tf_2N]-CO_2 system confined in 25 Å slit silica pore at 298.15 K and **(A)** 1 bar, **(B)** 10 bar, **(C)** 30 bar, and **(D)** 50 bar.

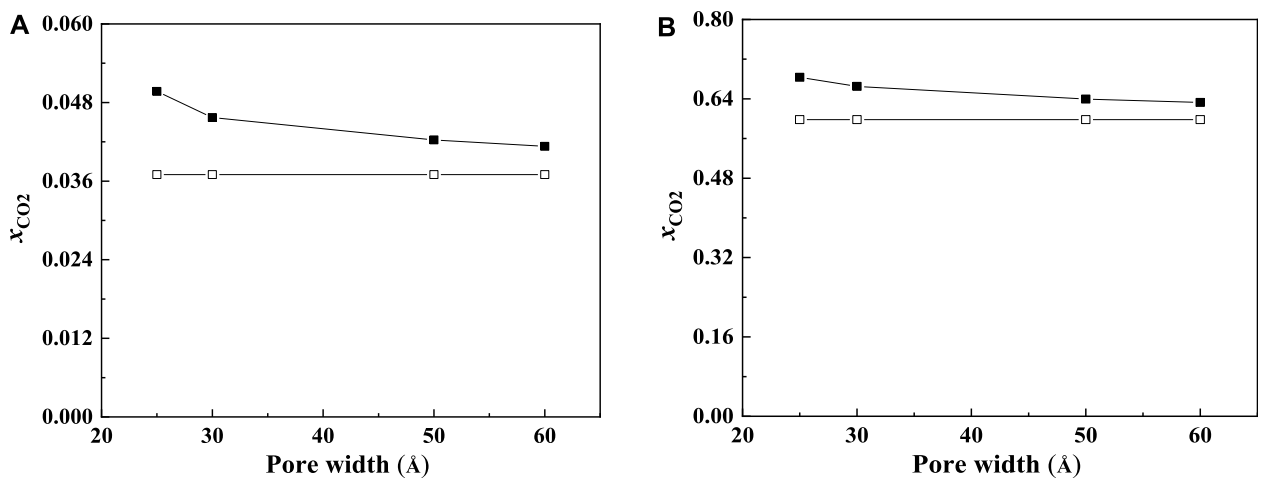


FIGURE 10 | Calculated CO_2 solubility of [C_6mim] [Tf_2N] at 298.15 K and 1 bar **(A)** and 30 bar **(B)** in slit-shaped pores with different widths (solid symbols, data for confined ILs; open symbols, data for bulk ILs.).

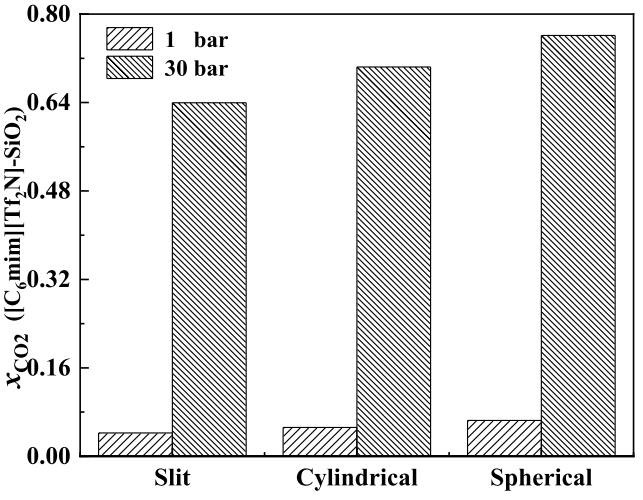


FIGURE 11 | Calculated CO_2 solubility of [C_6mim] [Tf_2N] confined in different-shaped pores at 298.15 K and 1 and 30 bar (the width of slit-shaped pore and the diameter of cylindrical pore and spherical cavity are all 5 nm).

about 0.02 mol CO_2/mol IL. Under the same condition, the calculated solubilities of CO_2 in the confined [C_6mim] [Tf_2N] reduce by about 0.03 mol CO_2/mol IL. It indicates that confined ILs may make it easier to desorb CO_2.

In order to study the effect of temperature on the absorption enhancement effect, we analyzed the relation between the additional solubility and temperatures. As illustrated in **Figures 4B** and **5B**, the additional solubility decreases with increasing temperature. In order to interpret this, the calculated density profiles of [C_6mim][Tf_2N]-CO_2 confined in 25 Å slit silica pore at 10 bar and different temperatures are illustrated in **Figure 6**. To quantitatively describe the effect of the first mechanism, the ratios of the CO_2 density at the first peak near the pore surface (which is also the maximum density inside the nanopore) to its bulk density in ILs were calculated as illustrated in **Figure 7**. According to the results shown in **Figure 7**, the ratio decreases with the increase in the temperature when the pressure keeps constant, and consequently, the additional solubility decreases. The same phenomenon can be observed at other pressures and in other pore structures, as listed in **Supplementary Table S1**.

3.2.3 The Influence of Pressure

Most commercial ILs are physical absorbents for CO_2. In general, the CO_2 solubility in ILs will increase significantly with increasing pressure. **Figure 8A** illustrates the CO_2 solubility at 298.15 K and different pressures in the 25 Å slit-shaped pore. From **Figure 8B**, The CO_2 solubilities in confined ILs also increase with the increase in the pressure. In addition, the increase in CO_2 solubility in the confined [C_6mim] [Tf_2N] is more significant than that of the bulk [C_6mim] [Tf_2N] under the same condition. For instance, when the pressure increases from 1 to 50 bar, the solubility of CO_2 in the confined [C_6mim] [Tf_2N] increases by about 0.78 mol CO_2/mol IL, while that in bulk [C_6mim] [Tf_2N] increases by about 0.70 mol CO_2/mol IL.

The additional solubility also increases with increasing pressure, as illustrated in **Figure 8B**. To have a deep insight into this observation, the density profiles of [C_6mim] [Tf_2N]-CO_2 confined in 25 Å slit silica pore at 298.15 K and different pressures illustrated in **Figure 9** were taken as one example for the detailed analysis. According to the results shown in **Figure 9**, with the increase in pressure, the adsorption of CO_2 near the pore surface increases, while the adsorption of IL ions near the pore surface decreases. This can be interpreted that at high pressure, a large amount of CO_2 inside the nanopore leads to a stronger competitive adsorption capacity of CO_2 in the pore surface than that of IL ions. Consequently, the additional solubility increases with the increase in pressure. The same phenomenon can be observed under other conditions, as listed in **Supplementary Table S1**.

3.2.4 The Influence of Pore Size and Shape

The CO_2 solubilities of confined ILs will also be affected by the pore size and shape. According to the calculation results, the adsorption will be enhanced more significantly in the smaller nanopore, which is consistent with the results of Shi and Luebke (2013). Two typical examples are shown in **Figure 10**. As illustrated in **Figure 10**, the solubilities of CO_2 increase with the decrease in the slit-shaped pore.

The effect of the shape of nanopores was also investigated. According to our calculation results, for nanopores with the same size (i.e., the width for slit-shaped pore and the diameter for cylindrical pore and spherical cavity), ILs confined in the spherical cavity have the highest CO_2 solubilities, while those confined in the slit-shaped pore have the lowest. Typical examples are illustrated in **Figure 11**. This indicates that using supported material with spherical nanopores may lead to better absorption capacity. The calculated density profiles of the three shapes at 298.15 K and 30 bar are presented in **Figure 12**. It shows that the pore shape affects the density profile near the surface of the pore significantly. For IL ions, the density at the first peak follows the trend that ρ (slit-shaped) > ρ (cylindrical) > ρ (spherical), while for CO_2, the opposite trend can be observed. This is because the inward curved surfaces impede the accumulation of large molecules (e.g., IL ions) near the surface. Therefore, under the same situation, the CO_2 solubilities (x) follow the trend that x (spherical) > x (cylindrical) > x (slit-shaped).

3.2.5 The Influence of Cation

In most cases, the IL-cations are chain-like substances. Therefore, in this part, the effect of alkyl-chain length was studied. Following this, ePC-SAFT-DFT was used to model the CO_2 solubilities of [C_4mim][Tf_2N], [C_6mim] [Tf_2N], and [C_8mim] [Tf_2N] inside silica nanopores at different temperatures, pressures, pore widths, and pore shapes. The calculated results are listed in **Supplementary Table S1**.

As illustrated in **Figures 13A** and **14A**, for the ILs in the same homologous series, the CO_2 solubilities in the confined ILs increase with increasing alkyl-chain length in cation, which is

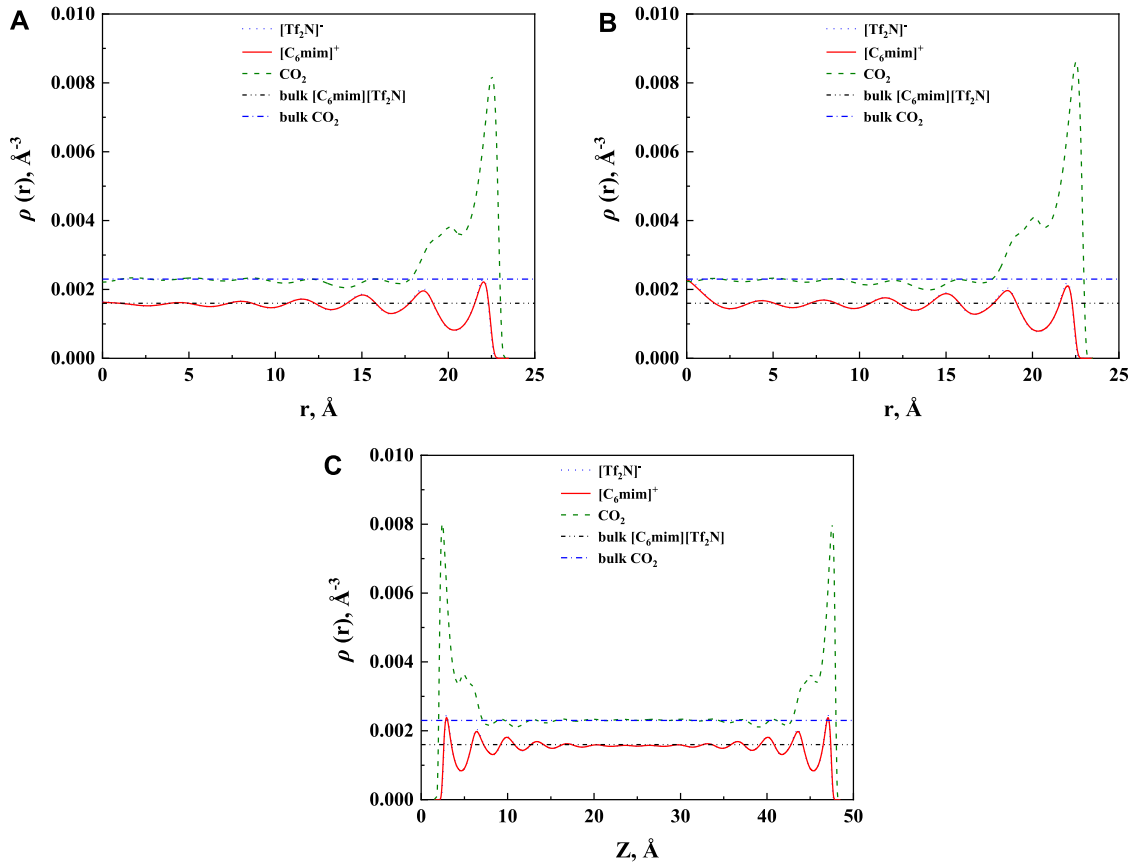


FIGURE 12 | Calculated density profiles [C_6mim] [Tf_2N]-CO_2 confined in 50 Å **(A)** cylindrical, **(B)** spherical, and **(C)** slit-shaped silica pore at 298.15 K and 30 bar.

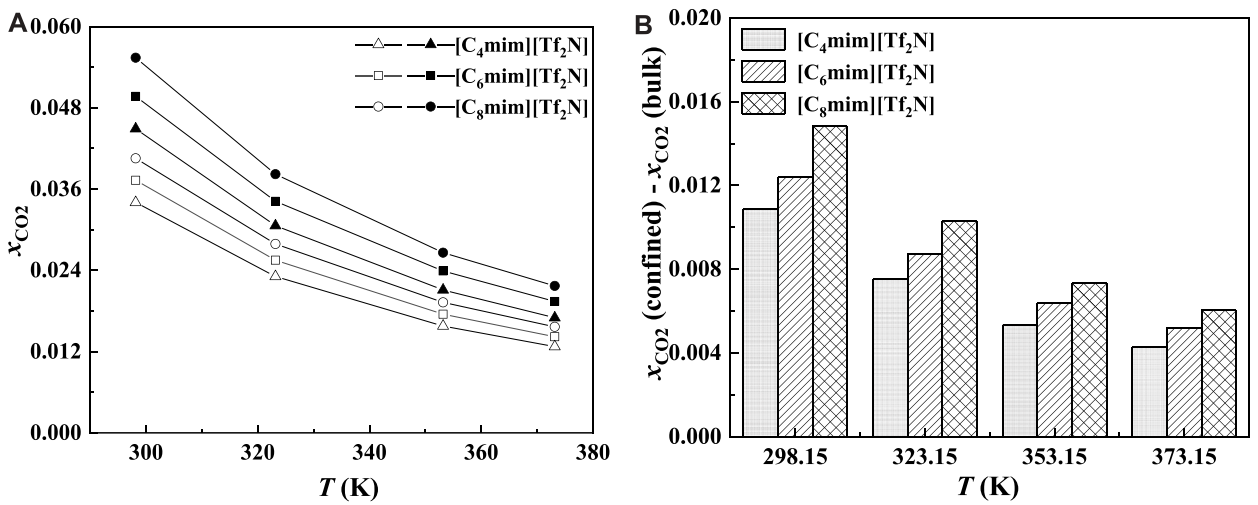


FIGURE 13 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of ILs confined in 25 Å slit-shaped pore at 1 bar and different temperatures. **(A)** Solid symbols, data for confined ILs; open symbols, data for bulk ILs.

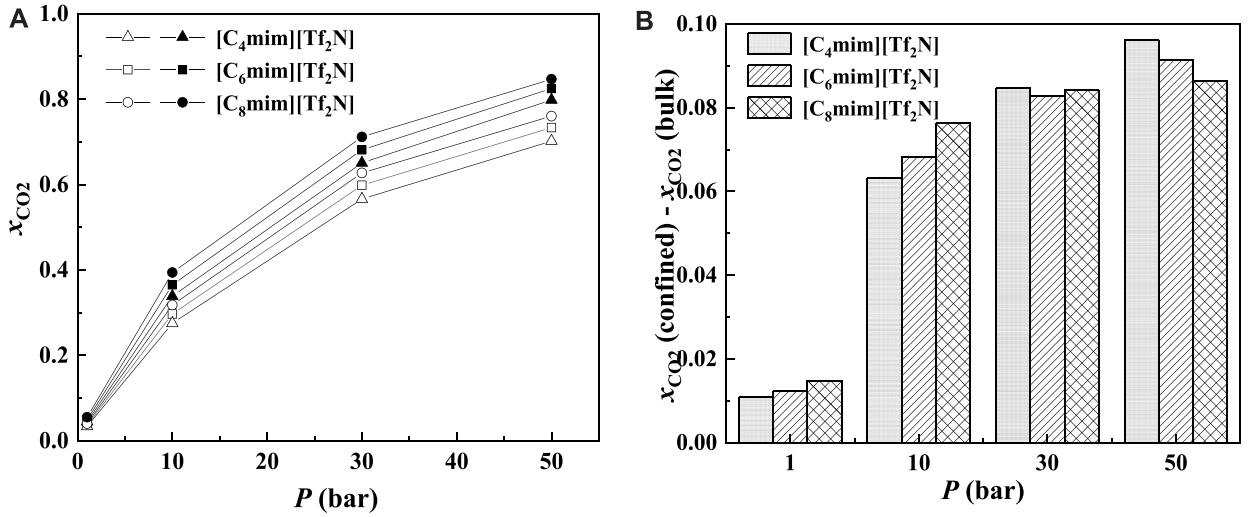


FIGURE 14 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of ILs confined in 25 Å slit-shaped pore at 298.15 K and different pressures. **(A)** Solid symbols, data for confined ILs; open symbols, data for bulk ILs.

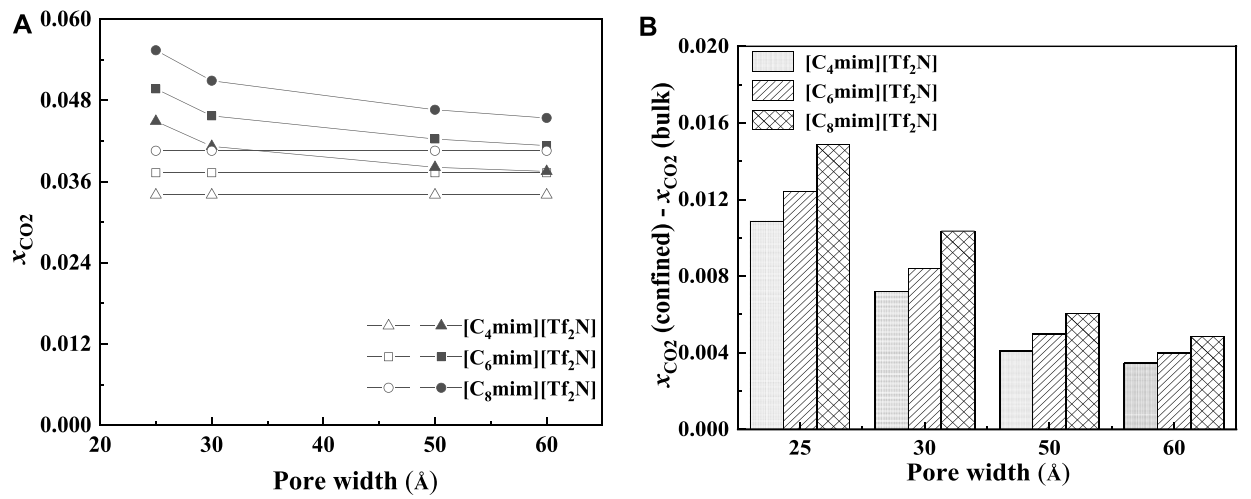


FIGURE 15 | Calculated CO_2 solubility **(A)** and additional solubilities **(B)** of ILs confined in slit-shaped pore with different widths at 298.15 K and 1 bar. **(A)** Solid symbols, data for confined ILs; open symbols, data for bulk ILs.

the same as that in the bulk ILs. According to **Figure 13B** and **Figure 14B**, at low pressures, the additional solubility also increases with the increase in the alkyl-chain length. However, this is not true at high pressures, as illustrated in **Figure 14B**. This indicates that the absorption enhancement of ILs with longer alkyl-chain decreases faster with the increasing pressure than the ILs with shorter alkyl-chain.

As demonstrated in **Figure 15**, with the increase in pore width, the additional solubility decreases, especially for the IL with longer alkyl-chain length in the IL cation. The additional solubilities of confined ILs in different pore shapes are illustrated in **Figure 16**. In the spherical cavity, the additional solubilities increase more significantly with increasing alkyl-chain length in IL cation than that in the slit-shaped and cylindrical pores.

3.2.6 The Influence of Anion

In the bulk phase, the IL anion affects the CO_2 solubilities more significantly than IL cation (Anthony et al., 2005). According to

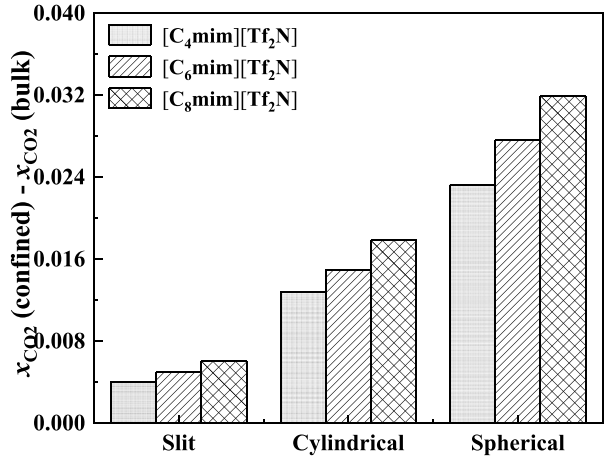


FIGURE 16 | Calculated additional solubilities of ILs confined in different shaped pores at 298.15 K and 1 bar in 50 Å. Solid symbols, data for confined ILs; open symbols, data for bulk ILs.

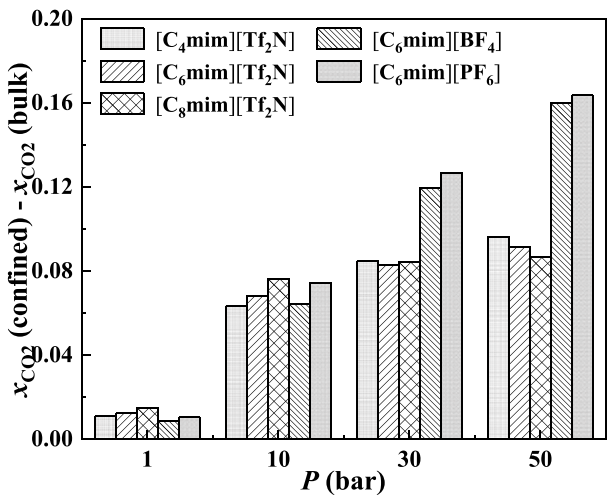


FIGURE 17 | Calculated additional solubilities of ILs confined in 25 Å slit-shaped pore at 298.15 K and different pressures.

Figure 17, the IL anion also has a more significant effect than IL cation in the absorption enhancement. From 1 to 50 bar, the deviation of the calculated additional solubility between [C_4mim] [Tf_2N] and [C_8mim] [Tf_2N] changes from 0.003 to 0.009, while for [C_6mim] [Tf_2N] and [C_6mim] [PF_6], the deviation of the calculated additional solubility changes from 0.002 to 0.071 from 1 to 50 bar.

In **Figure 18**, the calculated additional solubilities of several confined ILs in different pore shapes are presented. The absorption enhancement of anion is more significant in the spherical cavity as shown in **Figure 18**. For example, in the slit-shaped pore at 30 bar, the deviation of the calculated additional solubilities between [C_4mim] [Tf_2N] and [C_8mim] [Tf_2N] is 0.001,

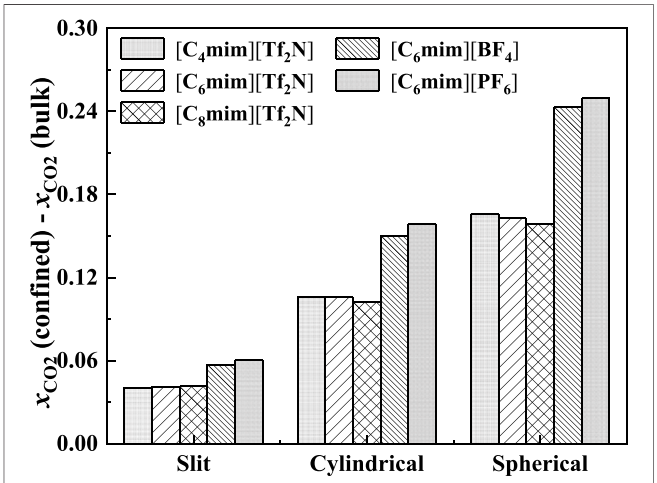


FIGURE 18 | Calculated additional solubilities of ILs in different-shaped pores at 298.15 K and 30 bar (the width of slit-shaped pore and the diameter of cylindrical pore and spherical cavity are all 5 nm).

and that between [C_6mim] [Tf_2N] and [C_6mim] [PF_6] is 0.007. While in the spherical pore, these two values are 0.007 and 0.086. In addition, for [C_6mim] [BF_4] and [C_6mim] [PF_6], the additional solubilities are considerably higher than those of [C_6mim] [Tf_2N] in the spherical cavity at high pressures. This indicates that changing the IL anion of ILs confined in the spherical cavity may be promising to obtain a large absorption enhancement.

Up to here of this section, the effects of temperature, pressure, IL ions, as well as the size and shape of pores on confined CO_2 solubilities were investigated. In order to further improve model performance toward practical applications, the ePC-SAFT-DFT model will be combined with different pore models (pore shape and pore size distribution) with the parameters adjusted from experimental data in our future work based on the newly measured experimental data.

4 CONCLUSION

In this work, the Chebyshev pseudo-spectral collocation method was combined with the Anderson mixing algorithm to further accelerate the ePC-SAFT-DFT calculation. The results show that the computing time can be significantly reduced with the Anderson mixing algorithm. This makes the ePC-SAFT-DFT model more effective in screening promising ILs. However, calculating matrices used in the Chebyshev pseudo-spectral collocation method for the cylindrical pores requires a certain computing time, while in a massive computation, these matrices can be reused to save computation time.

Using the ePC-SAF-DFT model, the CO_2 solubilities of ionic liquid confined *in silica* nanopores were studied. The results show that the CO_2 solubilities of confined ILs are always higher than that in bulk ILs under the same condition, and the absorption enhancement effect will be significantly affected by pressure, pore widths, pore shapes, and IL anion. Based on the simulation results, to obtain high CO_2 solubilities in silica-confined ILs, a better option is to use silica

material with a narrow spherical pore. In addition, the IL anion should be selected specifically considering that the category of IL anion has a significant impact on the absorption enhancement effect.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

GS conceptualized the study. YS formulated the methodology, was in charge of software and investigation, provided the visualization, and wrote the original draft. GS and XLi performed the validation. XLu, XLi, and XJ were in charge of the resources. ZD and XJ reviewed and edited the manuscript. ZD take part in the visualization and wrote the original draft.

FUNDING

This work was supported by the Swedish Research Council (Grant No. 2020-03899), the Joint Research Fund for Overseas Chinese Scholars and Scholars in Hong Kong and Macao Young Scholars (No. 21729601), and the Interreg Nord and Region Norrbotten in Sweden (No. 304-16169-2019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.801551/full#supplementary-material

REFERENCES

Anderson, D. G. (1965). Iterative Procedures for Nonlinear Integral Equations. *J. Acm* 12, 547–560. doi:10.1145/321296.321305

Anthony, J. L., Anderson, J. L., Maginn, E. J., and Brennecke, J. F. (2005). Anion Effects on Gas Solubility in Ionic Liquids. *J. Phys. Chem. B* 109, 6366–6374. doi:10.1021/jp046404l

Baltensperger, R. (2002). Some Results on Linear Rational Trigonometric Interpolation. *Comput. Maths. Appl.* 43, 737–746. doi:10.1016/S0898-1221(01)00317-0

Baltus, R. E., Counce, R. M., Culbertson, B. H., Luo, H., DePaoli, D. W., Dai, S., et al. (2005). Examination of the Potential of Ionic Liquids for Gas Separations. *Sep. Sci. Tech.* 40, 525–541. doi:10.1081/SS-200042513

Banu, L. A., Wang, D., and Baltus, R. E. (2013). Effect of Ionic Liquid Confinement on Gas Separation Characteristics. *Energy Fuels* 27, 4161–4166. doi:10.1021/ef302038e

Barker, J. A., and Henderson, D. (1967). Perturbation Theory and Equation of State for Fluids. II. A Successful Theory of Liquids. *J. Chem. Phys.* 47, 4714–4721. doi:10.1063/1.1701689

Bayliss, A., and Turkel, E. (1992). Mappings and Accuracy for Chebyshev Pseudo-spectral Approximations. *J. Comput. Phys.* 101, 349–359. doi:10.1016/0021-9991(92)90012-N

Beck, J. S., Vartuli, J. C., Roth, W. J., Leonowicz, M. E., Kresge, C. T., Schmitt, K. D., et al. (1992). A New Family of Mesoporous Molecular Sieves Prepared with Liquid crystal Templates. *J. Am. Chem. Soc.* 114, 10834–10843. doi:10.1021/ja00053a020

Berrut, J.-P., and Trefethen, L. N. (2004). Barycentric lagrange Interpolation. *SIAM Rev.* 46, 501–517. doi:10.1137/S0036144502417715

Boot-Handford, M. E., Abanades, J. C., Anthony, E. J., Blunt, M. J., Brandani, S., Mac Dowell, N., et al. (2014). Carbon Capture and Storage Update. *Energy Environ. Sci.* 7, 130–189. doi:10.1039/C3EE42350F

Brennecke, J. F., and Gurkan, B. E. (2010). Ionic Liquids for CO2 Capture and Emission Reduction. *J. Phys. Chem. Lett.* 1, 3459–3464. doi:10.1021/jz1014828

Budhathoki, S., Shah, J. K., and Maginn, E. J. (2017). Molecular Simulation Study of the Performance of Supported Ionic Liquid Phase Materials for the Separation of Carbon Dioxide from Methane and Hydrogen. *Ind. Eng. Chem. Res.* 56, 6775–6784. doi:10.1021/acs.iecr.7b00763

Camacho Vergara, E. L., Kontogeorgis, G. M., and Liang, X. (2019). Gas Adsorption and Interfacial Tension with Classical Density Functional Theory. *Ind. Eng. Chem. Res.* 58, 5650–5664. doi:10.1021/acs.iecr.9b00137

Cameretti, L. F., Sadowski, G., and Mollerup, J. M. (2005). Modeling of Aqueous Electrolyte Solutions with Perturbed-Chain Statistical Associated Fluid Theory. *Ind. Eng. Chem. Res.* 44, 3355–3362. doi:10.1021/ie0488142

Clenshaw, C. W., and Curtis, A. R. (1960). A Method for Numerical Integration on an Automatic Computer. *Numer. Math.* 2, 197–205. doi:10.1007/BF01386223

Fang, H.-r., and Saad, Y. (2009). Two Classes of Multisecant Methods for Nonlinear Acceleration. *Numer. Linear Algebra Appl.* 16, 197–221. doi:10.1002/nla.617

Fitzgerald, J. E., Sudibandriyo, M., Pan, Z., Robinson, R. L., Jr, and Gasem, K. A. M. (2003). Modeling the Adsorption of Pure Gases on Coals with the SLD Model. *Carbon* 41, 2203–2216. doi:10.1016/S0008-6223(03)00202-1

Groh, B., and Mulder, B. (2000). Hard-sphere Solids Near Close Packing: Testing Theories for Crystallization. *Phys. Rev. E* 61, 3811–3822. doi:10.1103/physreve.61.3811

Gross, J. (2009). A Density Functional Theory for Vapor-Liquid Interfaces Using the PCP-SAFT Equation of State. *J. Chem. Phys.* 131 (20), 204705. doi:10.1063/1.3263124

Gross, J., and Sadowski, G. (2001). Perturbed-chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. *Ind. Eng. Chem. Res.* 40, 1244–1260. doi:10.1021/IE0003887

Iarikov, D. D., Hacarlioglu, P., and Oyama, S. T. (2011). Supported Room Temperature Ionic Liquid Membranes for CO2/CH4 Separation. *Chem. Eng. J.* 166, 401–406. doi:10.1016/j.cej.2010.10.060

Ilconich, J., Myers, C., Pennline, H., and Luebke, D. (2007). Experimental Investigation of the Permeability and Selectivity of Supported Ionic Liquid Membranes for CO2/He Separation at Temperatures up to 125°C. *J. Membr. Sci.* 298, 41–47. doi:10.1016/j.memsci.2007.03.056

Ji, X., and Held, C. (2016). Modeling the Density of Ionic Liquids with ePC-SAFT. *Fluid Phase Equilibria* 410, 9–22. doi:10.1016/j.fluid.2015.11.014

Ji, X., Held, C., and Sadowski, G. (2012). Modeling Imidazolium-Based Ionic Liquids with ePC-SAFT. *Fluid Phase Equilibria* 335, 64–73. doi:10.1016/j.fluid.2012.05.029

Ji, X., Held, C., and Sadowski, G. (2014). Modeling Imidazolium-Based Ionic Liquids with ePC-SAFT. Part II. Application to H2S and Synthesis-Gas Components. *Fluid Phase Equilibria* 363, 59–65. doi:10.1016/j.fluid.2013.11.019

Kresge, C. T., Leonowicz, M. E., Roth, W. J., Vartuli, J. C., and Beck, J. S. (1992). Ordered Mesoporous Molecular Sieves Synthesized by a Liquid-crystal Template Mechanism. *Nature* 359, 710–712. doi:10.1038/359710a0

Lee, L. L. (2016). *Molecular Thermodynamics of Nonideal Fluids*. Oxford, United Kingdom: Butterworth-Heinemann.

Li, S., Li, Z., Medina, D., Lew, C., and Yan, Y. (2005). Organic-functionalized Pure-Silica-Zeolite MFI Low-K Films. *Chem. Mater.* 17, 1851–1854. doi:10.1021/cm048171g

Li, Z., and Wu, J. (2006). Density Functional Theory for Planar Electric Double Layers: Closing the gap between Simple and Polyelectrolytes. *J. Phys. Chem. B* 110, 7473–7484. doi:10.1021/jp060127w

MacDowell, N., Florin, N., Buchard, A., Hallett, J., Galindo, A., Jackson, G., et al. (2010). An Overview of CO2 Capture Technologies. *Energ. Environ. Sci.* 3, 1645–1669. doi:10.1039/C004106H

Maddox, M. W., Olivier, J. P., and Gubbins, K. E. (1997). Characterization of MCM-41 Using Molecular Simulation: Heterogeneity Effects. *Langmuir* 13, 1737–1745. doi:10.1021/la961068o

Mairhofer, J., and Gross, J. (2017). Numerical Aspects of Classical Density Functional Theory for One-Dimensional Vapor-Liquid Interfaces. *Fluid Phase Equilibria* 444, 1–12. doi:10.1016/j.fluid.2017.03.023

Mirzaei, M., Badiei, A. R., Mokhtarani, B., and Sharifi, A. (2017). Experimental Study on CO 2 Sorption Capacity of the Neat and Porous Silica Supported Ionic Liquids and the Effect of Water Content of Flue Gas. *J. Mol. Liquids* 232, 462–470. doi:10.1016/j.molliq.2017.02.104

Nandiyanto, A. B. D., Kim, S.-G., Iskandar, F., and Okuyama, K. (2009). Synthesis of Spherical Mesoporous Silica Nanoparticles with Nanometer-Size Controllable Pores and Outer Diameters. *Microporous Mesoporous Mater.* 120, 447–453. doi:10.1016/j.micromeso.2008.12.019

Nold, A., Goddard, B. D., Yatsyshin, P., Savva, N., and Kalliadasis, S. (2017). Pseudospectral Methods for Density Functional Theory in Bounded and Unbounded Domains. *J. Comput. Phys.* 334, 639–664. doi:10.1016/j.jcp.2016.12.023

Pinilla, C., Del Pópolo, M. G., Lynden-Bell, R. M., and Kohanoff, J. (2005). Structure and Dynamics of a Confined Ionic Liquid. Topics of Relevance to Dye-Sensitized Solar CellsTopics of Relevance to Dye-Sensitized Solar Cells. *J. Phys. Chem. B* 109, 17922–17927. doi:10.1021/jp052999o

P. Wasserscheid and W. Thomas (Editors) (2008). *Ionic Liquids in Synthesis* (John Wiley and SonsWiley-VCH).

Qiao, Y., Lian, C., Lu, B., and Wu, J. (2018). Modeling Selective Ion Adsorption into Cylindrical Nanopores. *Chem. Phys. Lett.* 709, 116–124. doi:10.1016/j.cplett.2018.08.047

Ramdin, M., de Loos, T. W., and Vlugt, T. J. H. (2012). State-of-the-Art of CO2 Capture with Ionic Liquids. *Ind. Eng. Chem. Res.* 51, 8149–8177. doi:10.1021/ie3003705

Ren, J., Wu, L., and Li, B.-G. (2012). Preparation and CO2 Sorption/Desorption of N-(3-Aminopropyl)Aminoethyl Tributylphosphonium Amino Acid Salt Ionic Liquids Supported into Porous Silica Particles. *Ind. Eng. Chem. Res.* 51, 7901–7909. doi:10.1021/ie2028415

Romanos, G. E., Schulz, P. S., Bahlmann, M., Wasserscheid, P., Sapalidis, A., Katsaros, F. K., et al. (2014). CO2 Capture by Novel Supported Ionic Liquid Phase Systems Consisting of Silica Nanoparticles Encapsulating Amine-Functionalized Ionic Liquids. *J. Phys. Chem. C* 118, 24437–24451. doi:10.1021/jp5062946

Rosenfeld, Y. (1989). Free-energy Model for the Inhomogeneous Hard-Sphere Fluid Mixture and Density-Functional Theory of Freezing. *Phys. Rev. Lett.* 63, 980–983. doi:10.1103/physrevlett.63.980

Rosenfeld, Y., Schmidt, M., Löwen, H., and Tarazona, P. (1997). Fundamental-measure Free-Energy Density Functional for Hard Spheres: Dimensional Crossover and Freezing. *Phys. Rev. E* 55, 4245–4263. doi:10.1103/PhysRevE.55.4245

Roth, R., Evans, R., Lang, A., and Kahl, G. (2002). Fundamental Measure Theory for Hard-Sphere Mixtures Revisited: the White Bear Version. *J. Phys. Condens. Matter* 14, 12063–12078. doi:10.1088/0953-8984/14/46/313

Roth, R. (2010). Fundamental Measure Theory for Hard-Sphere Mixtures: a Review. *J. Phys. Condens. Matter* 22, 063102. doi:10.1088/0953-8984/22/6/063102

Santos, E., Albo, J., and Irabien, A. (2014). Acetate Based Supported Ionic Liquid Membranes (SILMs) for CO2 Separation: Influence of the Temperature. *J. Membr. Sci.* 452, 277–283. doi:10.1016/j.memsci.2013.10.024

Sauer, E., and Gross, J. (2017). Classical Density Functional Theory for Liquid-Fluid Interfaces and Confined Systems: A Functional for the Perturbed-Chain Polar Statistical Associating Fluid Theory Equation of State. *Ind. Eng. Chem. Res.* 56, 4119–4135. doi:10.1021/acs.iecr.6b04551

Shen, G., Held, C., Lu, X., and Ji, X. (2015). Modeling Thermodynamic Derivative Properties of Ionic Liquids with ePC-SAFT. *Fluid Phase Equilibria* 405, 73–82. doi:10.1016/j.fluid.2015.07.018

Shen, G., Ji, X., and Lu, X. (2013). A Hybrid Perturbed-Chain SAFT Density Functional Theory for Representing Fluid Behavior in Nanopores. *J. Chem. Phys.* 138, 224706. doi:10.1063/1.482507810.1063/1.4808160

Shen, G., Laaksonen, A., Lu, X., and Ji, X. (2018). Developing Electrolyte Perturbed-Chain Statistical Associating Fluid Theory Density Functional Theory for CO2 Separation by Confined Ionic Liquids. *J. Phys. Chem. C* 122, 15464–15473. doi:10.1021/acs.jpcc.8b04120

Shen, G., Sun, Y., Zhang, X., Gao, X., Qian, Y., Lu, X., et al. (2021). Partition and Selectivity of Electrolytes in Cylindrical Nanopores with Heterogeneous Surface Charge. *J. Mol. Liquids* 340, 116839. doi:10.1016/j.molliq.2021.116839

Shen, Y., and Hung, F. R. (2017). A Molecular Simulation Study of Carbon Dioxide Uptake by a Deep Eutectic Solvent Confined in Slit Nanopores. *J. Phys. Chem. C* 121, 24562–24575. doi:10.1021/acs.jpcc.7b07315

Shi, W., and Luebke, D. R. (2013). Enhanced Gas Absorption in the Ionic Liquid 1-N-Hexyl-3-Methylimidazolium Bis(trifluoromethylsulfonyl)amide ([hmim][Tf2N]) Confined in Silica Slit Pores: A Molecular Simulation Study. *Langmuir* 29, 5563–5572. doi:10.1021/la400226g

Siderius, D. W., and Gelb, L. D. (2011). Extension of the steele 10-4-3 Potential for Adsorption Calculations in Cylindrical, Spherical, and Other Pore Geometries. *J. Chem. Phys.* 135, 084703. doi:10.1063/1.3626804

Sun, Y., Lu, X., Shen, G., and Ji, X. (2021). Accelerate the ePC-SAFT-DFT Calculation with the Chebyshev Pseudospectral Collocation Method. *Ind. Eng. Chem. Res.* 60, 9269–9285. doi:10.1021/acs.iecr.1c01077

Sun, Y., Schemann, A., Held, C., Lu, X., Shen, G., and Ji, X. (2019). Modeling Thermodynamic Derivative Properties and Gas Solubility of Ionic Liquids with ePC-SAFT. *Ind. Eng. Chem. Res.* 58, 8401–8417. doi:10.1021/acs.iecr.9b00254

Tripathi, S., and Chapman, W. G. (2005). Microstructure of Inhomogeneous Polyatomic Mixtures from a Density Functional Formalism for Atomic Mixtures. *J. Chem. Phys.* 122, 094506. doi:10.1063/1.1853371

Walker, H. F. (2011). *Anderson Acceleration: Algorithms and Implementations*. WPI Math. Sciences Dept. Report MS-6-15-50. Available at: https://users.wpi.edu~walker/Papers/anderson_accn_algs_imps.pdf.

Walker, H. F., and Ni, P. (2011). Anderson Acceleration for Fixed-point Iterations. *SIAM J. Numer. Anal.* 49, 1715–1735. doi:10.1137/10078356X

Xu, X., Cao, D., and Wang, W. (2008). Orientation of Rod Molecules in Selective Slits: a Density Functional Theory. *J. Phys. Condens. Matter* 20, 425221. doi:10.1088/0953-8984/20/42/425221

Yang, X., and Yue, X. (2007). Adsorption and Structure of Lennard-Jones Model Fluid in Slit-like Amorphous Silica Nanopores. *Colloids Surf. A: Physicochemical Eng. Aspects* 301, 166–173. doi:10.1016/j.colsurfa.2006.12.057

Yatsyshin, P., Savva, N., and Kalliadasis, S. (2012). Spectral Methods for the Equations of Classical Density-Functional Theory: Relaxation Dynamics of Microscopic Films. *J. Chem. Phys.* 136, 124113. doi:10.1063/1.3697471

Yu, Y.-X., and Wu, J. (2002a). Density Functional Theory for Inhomogeneous Mixtures of Polymeric Fluids. *J. Chem. Phys.* 117, 2368–2376. doi:10.1063/1.1491240

Yu, Y.-X., and Wu, J. (2002b). Structures of Hard-Sphere Fluids from a Modified Fundamental-Measure Theory. *J. Chem. Phys.* 117, 10156–10164. doi:10.1063/1.1520530

Zhang, J., Zhang, Q., Li, X., Zhang, Q., Li, X., Liu, S., et al. (2010). Nanocomposites of Ionic Liquids Confined in Mesoporous Silica Gels: Preparation, Characterization and Performance. *Phys. Chem. Chem. Phys.* 12, 1971–1981. doi:10.1039/b920556j

Zhang, J., Zhang, S., Dong, K., Zhang, Y., Shen, Y., and Lv, X. (2006). Supported Absorption of CO2 by Tetrabutylphosphonium Amino Acid Ionic Liquids. *Chem. Eur. J.* 12, 4021–4026. doi:10.1002/chem.200501015

Zhang, X., Zhang, X., Dong, H., Zhao, Z., Zhang, S., and Huang, Y. (2012). Carbon Capture with Ionic Liquids: Overview and Progress. *Energ. Environ. Sci.* 5, 6668–6681. doi:10.1039/C2EE21152A

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fchem.2021.827085

Energy Redistribution Following CO_2 Formation on Cold Amorphous Solid Water

*Meenu Upadhyay and Markus Meuwly**

Department of Chemistry, University of Basel, Basel, Switzerland

The formation of molecules in and on amorphous solid water (ASW) as it occurs in interstellar space releases appreciable amounts of energy that need to be dissipated to the environment. Here, energy transfer between CO_2 formed within and on the surface of amorphous solid water (ASW) and the surrounding water is studied. Following $CO(^1\Sigma^+) + O(^1D)$ recombination the average translational and internal energy of the water molecules increases on the ~10 ps time scale by 15–25% depending on whether the reaction takes place on the surface or in an internal cavity of ASW. Due to tight coupling between CO_2 and the surrounding water molecules the internal energy exhibits a peak at early times which is present for recombination on the surface but absent for the process inside ASW. Energy transfer to the water molecules is characterized by a rapid ~10 ps and a considerably slower ~1 ns component. Within 50 ps a mostly uniform temperature increase of the ASW across the entire surface is found. The results suggest that energy transfer between a molecule formed on and within ASW is efficient and helps to stabilize the reaction products generated.

Keywords: reactive molecular dynamics, amorphous solid water, interstellar chemistry, energy redistribution, CO_2 formation

OPEN ACCESS

Edited by:
Heribert Reis,
National Hellenic Research Foundation, Greece

Reviewed by:
Masashi Tsuge,
Hokkaido University, Japan
Jiao He,
Max Planck Society, Germany

****Correspondence:***
Markus Meuwly
m.meuwly@unibas.ch

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *01 December 2021*
Accepted: *08 December 2021*
Published: *08 February 2022*

Citation:
Upadhyay M and Meuwly M (2022) Energy Redistribution Following CO_2 Formation on Cold Amorphous Solid Water. Front. Chem. 9:827085. doi: 10.3389/fchem.2021.827085

1 INTRODUCTION

The motion of adsorbates in and on amorphous solid water (ASW) is essential for chemistry at astrophysical conditions. Typically, bulk water is present in the form of ASW which is the main component of interstellar ices. (Hagen et al., 1981). The structure of ASW is usually probed by spectroscopic measurements (Hagen et al., 1981; Jenniskens and Blake, 1994) although interference-based methods have also been employed. (Bossa et al., 2012). ASWs are porous structures characterized by surface roughness and internal cavities of different sizes which can retain molecular or atomic guests. (Bar-Nun et al., 1987). Under laboratory conditions the water ices have been reported to be porous (He et al., 2016; Kouchi et al., 2020) or non-porous (Oba et al., 2009; He et al., 2016; Kouchi et al., 2020) ASW whereas the morphology of ices in the interstellar medium is more debated. (Keane et al., 2001; Kouchi et al., 2021).

The high porosity of ASW (Bossa et al., 2014; Bossa et al., 2015; Cazaux et al., 2015) makes it a good catalyst for gas-surface reactions involving oxygen (Ioppolo et al., 2011a; Romanzin et al., 2011; Chaabouni et al., 2012; Minissale et al., 2013a; Dulieu et al., 2017; Pezzella et al., 2018; Pezzella and Meuwly, 2019; Christianson and Garrod, 2021), hydrogen (Hama and Watanabe, 2013), carbonaceous (Minissale et al., 2013b; Minissale et al., 2016a; Qasim et al., 2020; Molpeceres et al., 2021) and nitrogen-containing (Minissale et al., 2014) species and helps maintaining those species on (Minissale et al., 2019) or inside ASW. (Minissale et al., 2016b; Tsuge et al., 2020). This

increases the probability for the reaction partners to diffuse to locations for collisions and association reactions to occur. As the diffusivity of individual atoms and small molecules has been established from both, experiments and simulations, (Minissale et al., 2013b; Lee and Meuwly, 2014; Pezzella et al., 2018), this is a likely scenario for formation of molecules on and within ASW.

Earlier thermoluminescence experiments suggested that the $O(^3P)+CO(^1\Sigma^+)$ reaction with both reaction partners in their electronic ground state yields excited CO_2^* which, after emission of a photon, leads to formation of CO_2. (Fournier et al., 1979). Such a process has also been proposed to occur on interstellar grains (Ruffle and Herbst, 2001) and has been confirmed experimentally (Minissale et al., 2013b) with an estimated entrance barrier of 0.014–0.103 eV for the process on ASW, compared with a value of 0.3 eV from high-level electronic structure calculations. (Veliz et al., 2021). The surrounding water matrix provides the necessary coupling (Roser et al., 2001) to facilitate relaxation of the $^3A'$ or $^3A''$ states of CO_2 to the $^1A'$ ground state (correlating with linear $^1\Sigma_g^+$). The presence of an entrance barrier for the $O(^3P)+CO(^1\Sigma^+)$ reaction has one led to consider the alternative CO + OH pathway for CO_2 formation. (Watanabe and Kouchi, 2002; Ioppolo et al., 2011b). This was, however, reconsidered to yield the HOCO intermediate in such environments in more recent experiments. (Qasim et al., 2019). Also, the reaction products of the CO + OH reaction have been found to depend on the experimental conditions. (Oba et al., 2010; Noble et al., 2011). As the CO + OH reaction also appears to have a barrier in the entrance channel, (Noble et al., 2011), attention has recently shifted to the HOCO + H reaction for CO_2 formation. (Qasim et al., 2019).

For adsorbed species to react on ASW they need to be able to diffuse. This has been demonstrated from MD simulations with diffusion coefficients and desorption energies consistent with experiments. (Lee and Meuwly, 2014; Ghesquière et al., 2015). Atomic oxygen (Pezzella et al., 2018) on ASW experiences diffusional barriers between E_{dif} = 0.2 kcal/mol and 2 kcal/mol (100–1000 K) compared with values of $E_{dif} = 990^{+530}_{-360}$ K determined from experiments. (Minissale et al., 2016b). For CO, MD simulations reported (Pezzella and Meuwly, 2019) desorption energies between 3.1 and 4.0 kcal/mol (1560–2012 K or 130–170 meV), compared with 120 meV from experiments. (Karssemeijer et al., 2013). It was also found that the CO desorption energy from ASW depends on CO coverage with ranges from E_{des} = 1700 K for low to 1000 K for high coverage (He et al., 2016) which is consistent with the simulations. (Pezzella and Meuwly, 2019). On non-porous and crystalline water surfaces submonolayer desorption energies for CO are 1307 and 1330 K (~ 115 meV), respectively. (Noble et al., 2012). Experimental diffusional barriers range from 350 ± 50 K (Kouchi et al., 2020) to 490 ± 12 K. (He et al., 2018).

As such association reactions are in general exothermic, the energy released needs to be transferred to environmental degrees of freedom for the reaction products to stabilize. This is the quest of the present work which investigates the time scale and degrees of freedom to receive the energy liberated for the $O(^1D)+CO(^1\Sigma^+)$ reaction to form ground state $CO_2(^1\Sigma_g^+)$. The chemical precursors for formation of CO_2 are believed to be carbon monoxide and atomic oxygen and the CO + O reaction has been proposed as a non-energetic pathway, close to conditions in interstellar environments, for CO_2 formation 20 years ago from experiments involving a water-ice cap on top of CO and O deposited on a copper surface. (Roser et al., 2001). Formation of $CO_2(^1\Sigma_g^+)$ from ground state $CO(^1\Sigma^+)$ and electronically excited $O(^1D)$ is barrierless. The excited atomic oxygen species can, for example, be generated from photolysis of H^2O (Stief et al., 1975) which has a radiative lifetime of 110 min (Garstang, 1951). An alternative pathway proceeds$_1$ via electron-induced neutral dissociation of water into H^2 + O(D). (Schmidt et al., 2019). In the presence of CO formation of CO^2 in cryogenic CO/H^2O films was observed. (Schmidt et al., 2019).

After recombination $O(^1D)+CO(^1\Sigma^+) \rightarrow CO_2(^1\Sigma_g^+)$ the product is in a highly vibrationally excited state. For it to stabilize, excess internal energy needs to be channeled into the environment which is the ASW. The present work characterizes and quantifies energy relaxation of the $CO_2(^1\Sigma_g^+)$ product into internal and translational degrees of freedom of the surrounding water matrix. First, the methods used are described. Then, results are presented and discussed. Finally, conclusions are drawn.

2 COMPUTATIONAL METHODS

All molecular dynamics (MD) simulations were carried out using the CHARMM suite of programs (Brooks et al., 2009) with provisions for bond forming reactions through multi state adiabatic reactive MD (MS-ARMD). (Nagy and Yosa Reyes, 2014). The simulation system, **Figure 1**, consisted of an equilibrated cubic box of amorphous solid water with dimension 31 × 31 × 31 $Å^3$ containing 1000 water molecules. As all bonds and angles are flexible, the simulations were run with a time step of Δt = 0.1 fs and the non-bonded cutoff was at 13 Å. Simulations were started from an existing, equilibrated ASW structure (Pezzella et al., 2018; Pezzella and Meuwly, 2019; Upadhyay et al., 2021) by adding CO_A and O_B inside (**Figure 1A**) or on top of (**Figure 1B**) ASW.

In the following, the coordinates are the CO stretch r, the separation R between the center of mass of CO_A and O_B and θ is the O_ACO_B angle. In addition, the C–O_B separation will be considered where appropriate. Initial conditions were generated for a grid of angles θ and separations R and simulations were carried out to obtain initial coordinates and velocities for each of the grid points. With constrained CO and O position, first 750 steps of steepest descent and 100 steps Adopted Basis Newton-Raphson minimization were carried out, followed by 50 ps heating dynamics to 50 K. Then, 100 ps equilibration dynamics was carried out. From each of the runs coordinates and velocities were saved regularly to obtain initial conditions for each combination of angle and distance. Production simulations 500 ps or 6 ns in length were then run from saved coordinates and velocities in the *NVE* ensemble. Data (energies, coordinates and velocities) were saved every 1000 steps for subsequent analysis.

Water was described by a reparametrized, (Burnham et al., 1997; Plattner and Meuwly, 2008), flexible KKY (Kumagai, Kawamura, Yokokawa) model. (Kumagai et al., 1994). The typical water modes that couple in the ~ 2000 cm^{-1} region

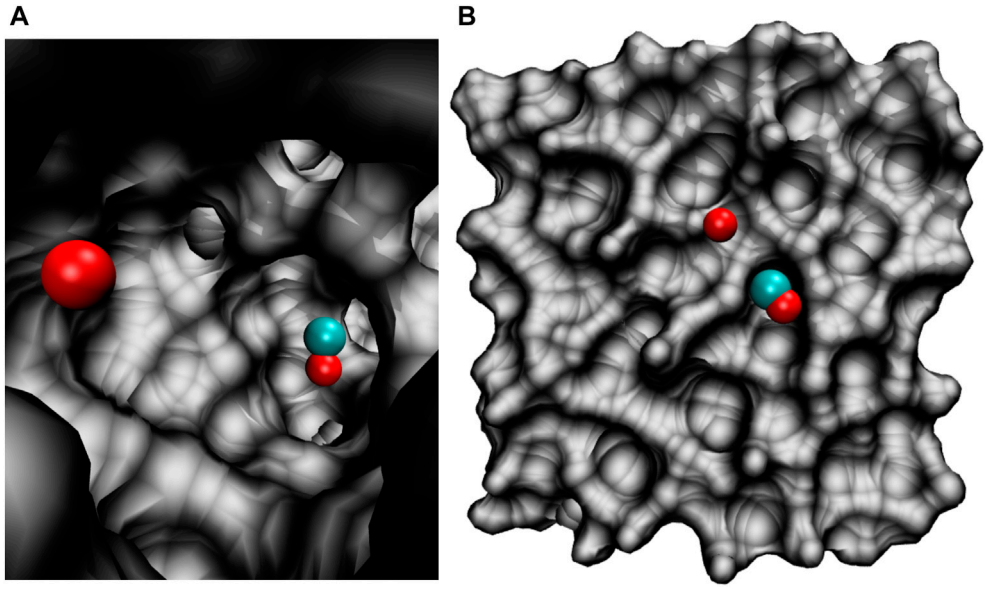


FIGURE 1 | The simulation system for studying the $O(^1D)+CO(^1\Sigma^+) \rightarrow CO_2(^1\Sigma_g^+)$ recombination reaction. **(A)**: CO and O trapped inside a cavity of ASW; **(B)**: CO and O on the top of the ASW surface.

relevant in the present work are the water bend (1,600 cm^{-1}) and the framework rotation (600 cm^{-1}) as was also found for the vibrational relaxation of cyanide in water. (Lee and Meuwly, 2011). To describe $CO_A + O_B$ recombination to form CO_2 the Morse-Morse-Harmonic (MMH) parametrization was employed. (Upadhyay et al., 2021). This model treats the two CO bonds with a Morse potential and the OCO bend as a harmonic function. MMH is a computationally efficient model (fitted to MRCI/aug-cc-pVTZ data), which yields results for recombination probabilities on ASW comparable to a more elaborate reproducing kernel Hilbert space (RKHS) representation with an exothermicity of –7.27 eV (Veliz et al., 2021; Upadhyay et al., 2021).

For CO_2, the partial charges were $q_O = -0.3e$ and $q_C = 0.6e$ with standard van der Waals parameters from CHARMM. These charges are consistent with those obtained from B3LYP/6-31G (d,p) calculations snapshots from the MD simulations with CO_2 adsorbed to a small water cluster $(H_2O)_{10}$ which yield $q_C = 0.73e$ and $q_O = -0.35e$. This compares with charges of $q_C = 0.22e$ and $q_O = -0.21e$ for the CO molecule and $q_O = -0.1e$ for an oxygen atom adsorbed to $(H_2O)_{10}$. To assess the dependence of the results on the partial charges used, additional reactive MD simulations using the MMH parametrization were carried out with $q_O = -0.1e$ and $q_C = 0.2e$ (i.e., $q_{CO} = 0.1e$) and with $q_O = -0.2e$ and $q_C = 0.4e$ (i.e., $q_{CO} = 0.2e$). In all cases, recombination was found to speed up compared with $q_{CO} = 0.3e$ and $q_O = -0.3e$ due to the increased mobility of the CO molecule and the O atom on the ASW when reduced partial charges are used.

The main focus of the present work is to study energy redistribution within the system following recombination of $CO_A + O_B$ to form CO_2. For this, the average total, translational and internal energy of the water molecules is analyzed for recombination on top of and inside ASW. Both, the time scale and amount of energy dissipated into translational and internal degrees of freedom was determined. The translational energy for the water molecules at each timestep was determined by first calculating the magnitude of the linear momentum of each water molecule from the stored velocities. From this the translational energy contribution is calculated for each water molecule considered and the total translational energy is accumulated. The internal energy for each water molecule is determined from the difference of the total kinetic energy and the translational energy.

3 RESULTS AND DISCUSSION

In the following, the energy distribution in the water matrix of the ASW is separately discussed on the ~100 ps and on the nanosecond time scale. Next, the energy flow away from the recombination site is analyzed and, finally, the energy redistribution to neighboring water molecules surrounding the recombination site is considered.

3.1 Recombination on the 100 ps Time Scale

A typical trajectory for $CO_A + O_B$ recombination inside the ASW cavity is shown in **Figure 2** (left column). Initially, the C–O_B separation is ~6 Å (**Figure 2A**). Within 150 ps recombination takes place and angular distortions lead to exploration of angles $\theta_{OCO} \sim 90^\circ$ (**Figure 2B**). Relaxation of the angle occurs within the following 50 ps and the CO_2 molecule remains in an internally excited state on much longer time scales, see **Figure 2C**. (Upadhyay et al., 2021). Concomitantly, the average internal energy of the surrounding water molecules increases by about

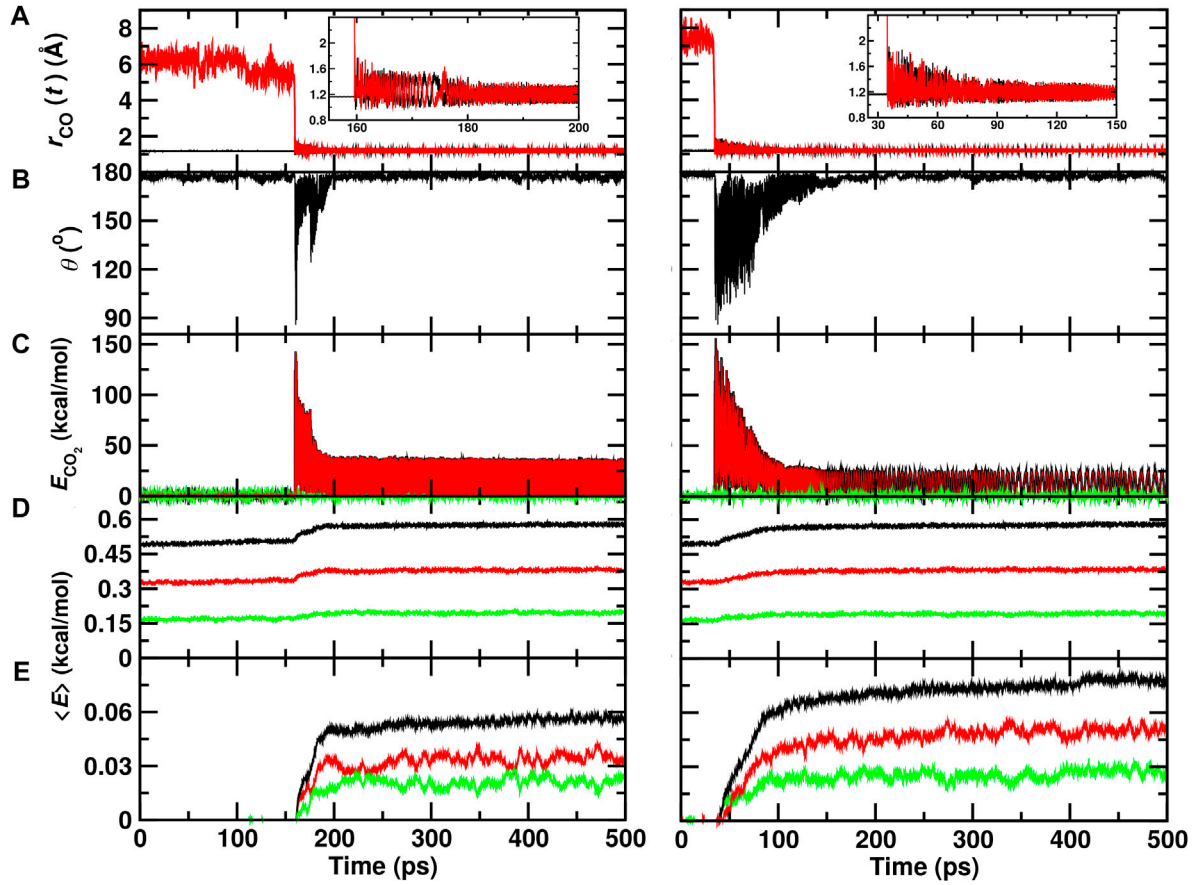


FIGURE 2 | Recombination of O (^{1}D)+CO($^1\Sigma^+$) to form ground state CO_2 ($^1\Sigma_g$) in the ASW cavity (left column) and on top of the ASW surface (right column). Initially, R = 6.0 Å and θ = 180°. **(A)**: O_B–C_{CO} separation (red) and CO_A separation (black); **(B)**: the O-C-O angle θ; **(C–E)**: the average total (black), internal (red), and translational (green) energies for the CO_2 molecule **(C)**, the average per water molecule **(D)**, and the magnitude of the average per water molecule relative to the energy before recombination **(E)**.

10%, see black, red and green traces in **Figures 2D,E**. The translational (phononic) modes (green) acquire approximately 1/3 of the additional energy whereas the internal energy (red) increases by the remaining 2/3.

Figure 2 (right column) reports a recombination trajectory on top of ASW. In this case, recombination takes place after ~ 35 ps and wide angular excursions extend out to 100 ps The amount of energy picked up by the water matrix is larger compared to recombination inside ASW (**Figures 2D,E**). The average total energy per water molecule increases by close to 20% and the amount that goes into internal degrees of freedom is considerably larger. For the translational modes, the energy after recombination is comparable to that for recombination within the cavity.

From a set of 70 recombination trajectories for the reaction within the cavity and on top of the ASW surface, the averaged energy contents in translational, internal and all degrees of freedom of the water molecules were determined (**Figure 3**). For this analysis, the time of reaction was set to zero (t = 0) to align all reactive trajectories and all energies are reported relative to the averages before recombination. The translational contribution for recombination within and on top of ASW re-equilibrates on the ~ 25 ps time scale after which no change in the phononic degrees of freedom is observed. Contrary to that, the internal degrees of freedom (red traces) show temporal evolution on two time scales: a rapid phase on the picosecond time scale, followed by a slow, long increase in the internal energies. This is also reflected in the averaged total energy (black).

As for the single trajectories, the amount of energy released from the recombination reaction into the translational degrees of freedom is similar for the reaction inside the cavity and on top of the ASW surface. For the internal degrees of freedom, however, recombination on top of the ASW surface leads on an average increase per water molecule by 0.075 kcal/mol within 400 ps (**Figure 3B**) compared with 0.06 kcal/mol for the process inside the cavity. Also, there is a characteristic decrease in the internal contribution for recombination on the surface after 15 ps which is even present when averaging over 70 independent runs. This feature is not found for recombination within ASW.

To estimate approximate time scales for the different processes involved, the average energies were fitted to an empirical

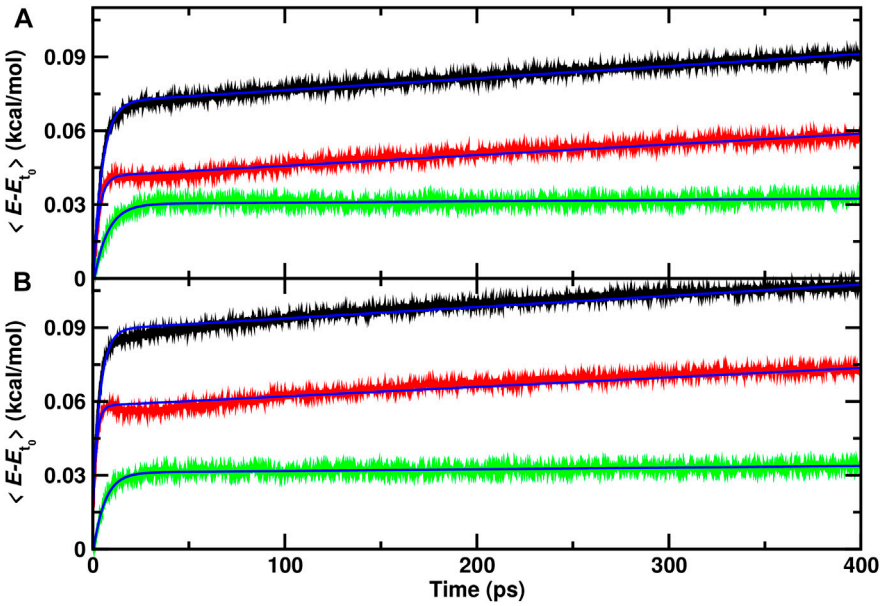


FIGURE 3 | Average total (black), internal (red) and translational (green) energies for water over 70 independent runs relative to the average before recombination. The time of reaction for all trajectories is shifted to t =0 and defined by the first instance at which $r_{C-O_B} < 1.6$ Å. **(A)**: recombination within the ASW cavity. **(B)**: recombination for $CO_A + O_B$ on the top of the ASW surface. The blue solid line is a fit to an empirical expression $\epsilon = a_0 e^{-t/a_1} + a_2 t + a_3$, see text.

expression $\epsilon = a_0 e^{-t/a_1} + a_2 t + a_3$ where ϵ is any of the energies considered. Such a functional form was chosen after inspection of the data in **Figure 3** and accounts for the rapid initial increase in the three energies together with the slow variation of the internal energy on longer times. This parametrization is not able to model the dip around 15 ps for recombination on to of the surface, though. The time scales a_1 for total, internal, and translational energies are (4.8, 2.9, 7.1) ps for recombination inside the cavity and speed up to (3.9, 1.9, 6.1) ps for the process on the ASW surface. It is of interest to note that the rapid time scale for the

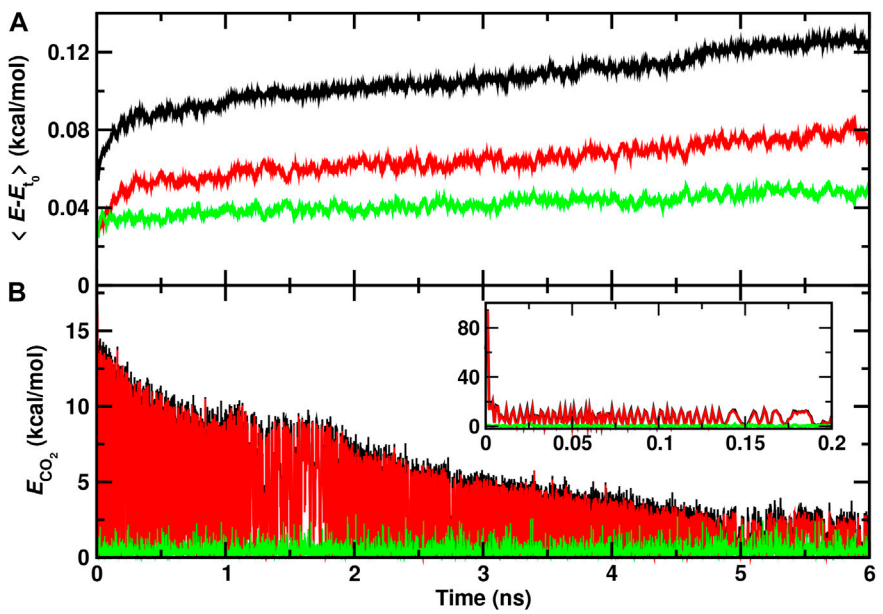


FIGURE 4 | Total (black), internal (red), and translational (green) energies for water molecules **(A)** and for the recombined CO_2 molecule **(B)** form a 6 ns rebinding trajectory on the top of the ASW surface. The time of reaction is shifted to t = 0 and defined by the first instance at which $r_{C-O_B} < 1.6$ Å. CO_2 continues to relax and the energy in the ASW further increases beyond the maximum simulation time of 6 ns after CO_2 recombination.

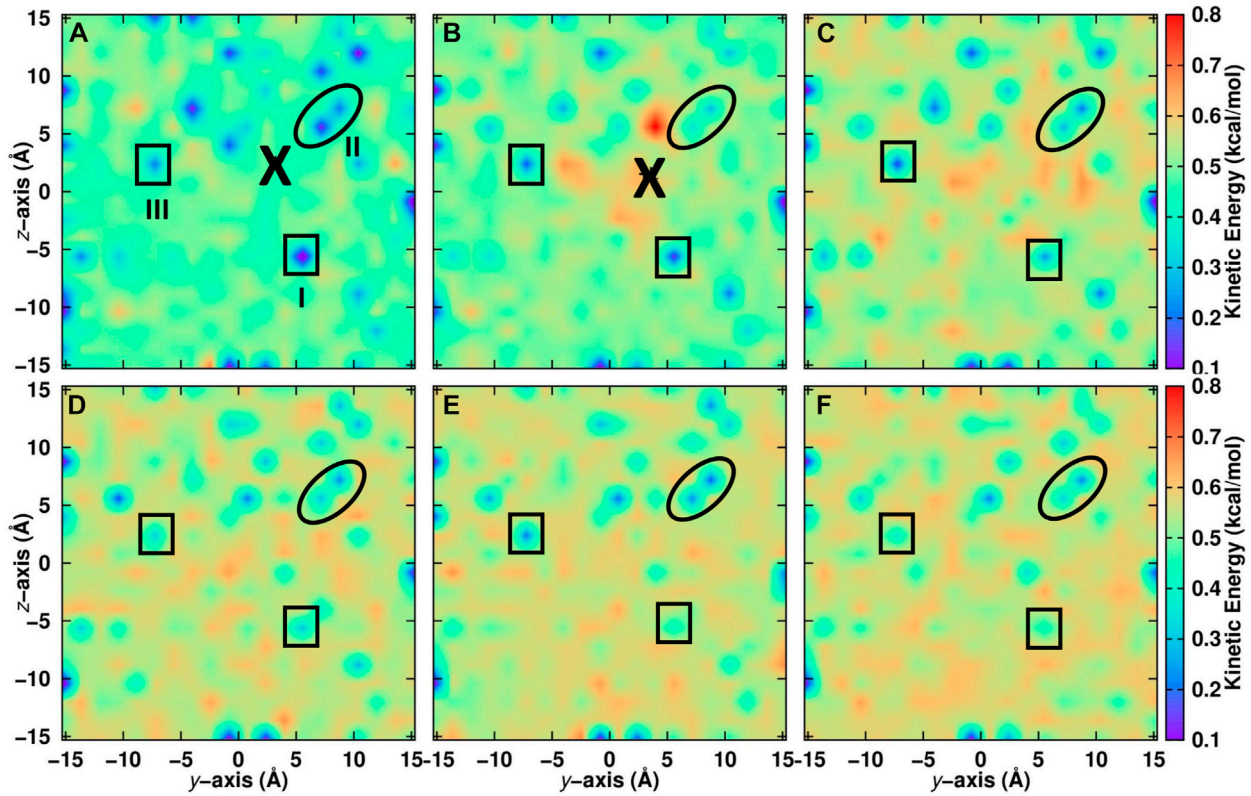


FIGURE 5 | Kinetic energy of water molecules projected onto the (y, z) – plane averaged over seven independent simulations on the top layer of the ASW surface. Recombination of CO and O takes place at the location labelled with "X". Before recombination **(A)**: $-5 \le t \le 0$ ps) and after recombination **(B)**: $0 \le t \le 5$ ps, **(C)**: $5 \le t \le 10$ ps, **(D)**: $10 \le t \le 50$ ps, **(E)**: $50 \le t \le 100$ ps and **(F)**: $100 \le t \le 200$ ps). The average position of CO is indicated by a large black cross. Noteworthy regions are labelled I to III and surrounded by solid lines. Region I is cool at early times and gradually warms up. Region II remains cool for most of the simulation time and region III alternates between cool and warm. For results within 10 Å of the surface, see **Supplementary Figure S4**.

internal energy is considerably faster than the kinetics of the translational degrees of freedom for both types of recombinations. The parameter a_2 which describes the slow increase of internal energy has a value of $a_2 = 4.3 \times 10^{-2}$ (kcal/mol)/ns for recombination in the cavity and $a_2 = 3.8 \times 10^{-2}$ (kcal/mol)/ns for the reaction on the surface, and is vanishingly small for the translational energy.

Average internal energies from representative independent runs for recombination inside the cavity and on top of the ASW surface are shown in **Supplementary Figures S1, S2**. For recombination inside the cavity (**Supplementary Figure S1**) the results confirm that the energy content in the internal degrees of freedom increases considerably faster than for the translation. Also, it is found that the amount of energy transferred to translation after recombination is smaller than that partitioned into internal degrees of freedom. For recombination on the ASW surface the same observations are made. In addition, the pronounced maximum after ~ 5 ps is present in all examples shown in **Supplementary Figure S2**. To provide a molecularly resolved interpretation of this feature the HOH angle time series $\theta(t)$ was analyzed for a trajectory in which CO + O recombination occurred after 35 ps, see **Supplementary Figure S3B**. At the time of reaction the water bending angle decreases from its average equilibrium value by $\langle \Delta\theta \rangle \sim 1^{\circ}$ over the next 70 ps after which it relaxes back to the original value. The signature in the internal energy extends over ~ 30 ps, see **Supplementary Figure S2**. Hence, it is possible that changes in the average water geometry following CO + O recombination are responsible for the overshooting and subsequent relaxation of the internal energy for the reaction on the surface. The HOH angle for a simulation within ASW in **Supplementary Figure S3A** also shows a slight adjustment of the valence angle after recombination. Contrary to the situation on the surface, the average angle does not relax to the value before recombination, though.

3.2 Recombination Dynamics on Longer Time Scales

It is also of interest to analyze the energy redistribution on the multi-nanosecond time scale. **Figure 4A** demonstrates that the average total kinetic energy per water molecule continuously increases even on the nanosecond time scale. Most of this increase is due to the internal degrees of freedom although the translational component also shows a continuous slow increase on the nanosecond time scale.

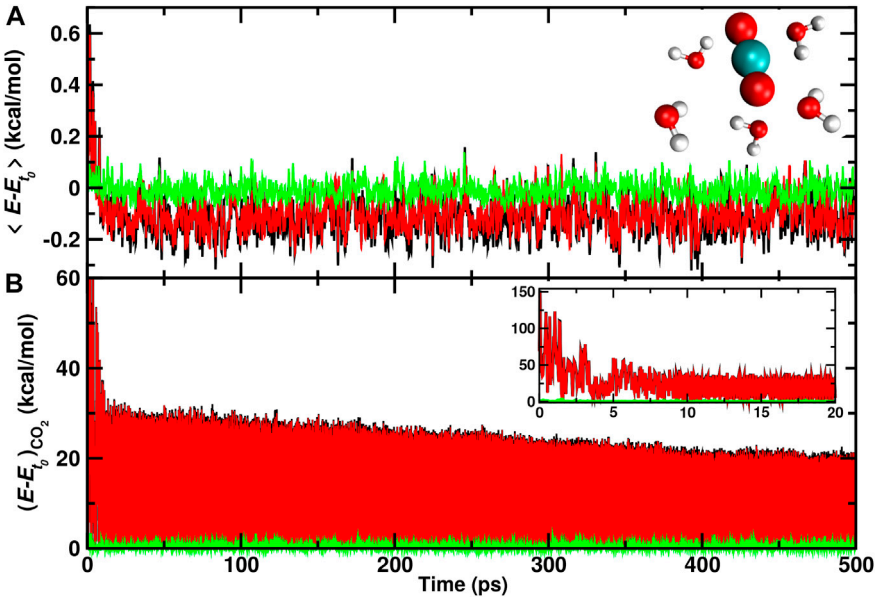


FIGURE 6 | Average total (black), internal (red), and translational (green) energies for 5 water molecules **(A)** closest to the CO_2, and the CO_2 molecule **(B)** formed from a recombination trajectory on the ASW surface. The time of reaction is shifted to $t = 0$ and defined by the first instance at which $r_{C-O_B} < 1.6$ Å. The initial time scale for energy redistribution from the relaxing CO_2 molecule to internal and translational degrees of freedom of the ASW occurs on the 10 ps time scale with slow gradual relaxation on the ~ 100 ps to ns time scale. On the 500 ps time scale the 5 water molecules slightly cool compared with the kinetic energy before recombination.

The relaxation of the CO_2 internal energy is reported in **Figure 4B**. Within the first few picoseconds (inset) the internal energy is quenched to ~ 10 kcal/mol after which two relaxations are observed. A first phase during 1 nanosecond following recombination and a second, slower phase extending out to 6 ns and beyond. By the end of the simulation the average internal energy of the CO_2 molecule has decreased to ~ 2.5 kcal/mol on average. Hence, it is expected that energy transfer to the surrounding water continues but slows down considerably on the 10 ns time scale and longer.

3.3 Energy Migration Around the Recombination Site

For a positionally resolved picture of energy flow the simulation system was separated in voxels with dimension 31 × 1 × 1 $Å^3$. The kinetic energy of all water molecules within one such voxel was averaged along the trajectory and projected onto the (y, z) – plane. Which water molecules belong to a particular voxel was decided based on the water-oxygen atom coordinates. **Figure 5A** reports the distribution of total kinetic energy distribution before recombination. The recombination site is at $(y = 2, z = 2)$ Å and marked as a large cross. Within the first 5 ps after recombination the kinetic energy of water molecules within ~ 10 Å of the recombination site increases considerably, by up to a factor of 4. Following this, energy redistributes continuously across the entire surface on the 200 ps time scale, see panels C to F.

Certain regions that are initially "cold" (blue)—e.g., the region labelled "I" at $(y = 5, z = -5)$ Å in **Figure 5**—warm up as energy transfer from CO_2 to the water molecules occurs. Conversely, other regions remain "cool", such as region "II" around $(y = 5, z = 5)$ Å for which the color code remains blue until 200 ps. Yet for other regions, such as "III", the total kinetic energy oscillates between cooler and warmer. It is also instructive to include only the first few ASW layers in this analysis which was done in **Supplementary Figure S4**. Here, the voxels have sizes 1 × 1 × 1 A^3. For one, the cool regions are more extended before recombination. After recombination energy transfer occurs in a similar fashion as for the full system. However, the warm regions are less extended. This suggests that energy transfer also occurs to a considerable extent *into* the bulk rather than across the surface of the ASW even for recombination on top of ASW.

3.4 Energy Flow to Nearby Water Molecules

Finally, individual water molecules in immediate proximity of the recombination site are analyzed. For one trajectory with recombination on the ASW surface the average total, internal, and translational energies for the 5 water molecules closest to the recombination site are reported in **Figure 6**. During the first 10 ps after recombination the average total kinetic energy increases by up to 0.6 kcal/mol per water molecule. Conversely, the translational energy contribution fluctuates around zero which indicates that the local structure of ASW remains intact and most of the energy flows into internal degrees of freedom.

After this initial increase, cooling of these few nearby water molecules takes place with a long-time average of −0.1 kcal/mol

per water molecule in the internal degrees of freedom. On the 500 ps no noticeable change in the translational energy content is observed. For the CO_2 molecule (**Figure 6B**) the translational energy remains small throughout the trajectory whereas the internal energy decreases rapidly within the first 5 ps following recombination. Subsequently, slow gradual cooling on the 100 ps to nanosecond time scale takes place as was already found earlier, **Figure 4**.

4 DISCUSSION AND CONCLUSION

The present work reports on the energy redistribution across ASW following $O(^1D)+CO(^1\Sigma^+)$ recombination to form CO_2 $(^1\Sigma_g^+)$ on the surface and in a cavity. It is found that energy distribution occurs in two phases, one on the picosecond and one on the nanosecond time scale for both locations. Although the time dependence of the processes is similar for the two different recombination sites (inside vs. on top), the dynamics differs in a number of ways. Firstly, recombination on the surface leads to excess internal energy on the picosecond time scale which subsequently relaxes and additional energy transfer into water modes occurs on longer time scales. Secondly, recombination within the cavity considered here leads to smaller magnitude (~ 15%) of energy transferred per water molecule compared with the process on the surface (~ 25%). A possible reason for this is that within a sufficiently large cavity the recombined CO_2 molecule exchanges energy with the surrounding through direct collision whereas on the surface CO_2 is always in contact with the ASW. In other words, the coupling between CO_2 and water differs for recombination within ASW and on top of it. Finally, heating of the water molecules occurs on the 10 ps time scale following the recombination reaction. Consistent with earlier work, (Fredon et al., 2021; Upadhyay et al., 2021), no CO_2 desorption is found from the simulations carried out here.

It is of interest to note that - ultimately - energy redistribution in such systems follows quantum mechanical principles. The present results suggest that the local energy generated from CO + O recombination is probably sufficient to excite internal modes of individual water molecules surrounding the recombination site. Hence, after CO + O recombination the ASW will be in a state characterized by a few internally and vibrationally excited water molecules embedded into a matrix of water molecules in the ground state. Earlier work on a related problem—the vibrational relaxation of a quantum oscillator coupled to oscillators of a biomolecule (Stock, 2009)—found that using classical mechanics leads to qualitatively correct results compared with a full quantum treatment. For the relaxation times a moderate factor of two for the difference between classical and rigorous quantum simulations was reported. Hence, for the present problem it is also expected that similar conclusions apply and that the nonequilibrium relaxation dynamics of individual vibrationally excited water molecules surrounded by vibrationally cold water molecules can be captured qualitatively from using classical dynamics.

In summary, the present work demonstrates that $O(^1D)+CO(^1\Sigma^+)$ recombination to form CO_2 $(^1\Sigma_g^+)$ leads to excitation of both, phononic and internal modes of the water molecules that constitute the ASW. The time scales for this are on the pico- and nano-second and lead to warming the water matrix. Water molecules in direct proximity of the recombination site may become vibrationally excited and the time scale for their relaxation back to the ground state will depend on the coupling to the immediate environment. Full relaxation of the CO_2 molecule is expected to require several 10–100 nanoseconds.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation. Data on the simulations and the PESs is available at: https://github.com/MMunibas/co2.asw and https://github.com/MMunibas/CO2-PESs.

AUTHOR CONTRIBUTIONS

MM designed research, MU carried out simulations, MU and MM analyzed results, MU and MM wrote manuscript.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Swiss National Science Foundation through grant 200 021-117 810 and to the NCCR-MUST.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.827085/full#supplementary-material

REFERENCES

Bar-Nun, A., Dror, J., Kochavi, E., and Laufer, D. (1987). Amorphous Water Ice and its Ability to Trap Gases. *Phys. Rev. B* 35, 2427–2435. doi:10.1103/physrevb.35.2427

Bossa, J.-B., Isokoski, K., de Valois, M. S., and Linnartz, H. (2012). Thermal Collapse of Porous Interstellar Ice. *A&A* 545, A82. doi:10.1051/0004-6361/201219340

Bossa, J.-B., Isokoski, K., Paardekooper, D. M., Bonnin, M., van der Linden, E. P., Triemstra, T., et al. (2014). Porosity Measurements of Interstellar Ice Mixtures Using Optical Laser Interference and Extended Effective Medium Approximations. *A&A* 561, A136. doi:10.1051/0004-6361/201322549

Bossa, J.-B., Maté, B., Fransen, C., Cazaux, S., Pilling, S., Rocha, W. R. M., et al. (2015). Porosity and Band-Strength Measurements of Multi-phase Composite Ices. *ApJ* 814, 47. doi:10.1088/0004-637x/814/1/47

Brooks, B. R., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., et al. (2009). CHARMM: the Biomolecular Simulation Program. *J. Comput. Chem.* 30, 1545–1614. doi:10.1002/jcc.21287

Burnham, C. J., Li, J.-C., and Leslie, M. (1997). Molecular Dynamics Calculations for Ice Ih. *J. Phys. Chem. B* 101, 6192–6195. doi:10.1021/jp9632596

Cazaux, S., Bossa, J.-B., Linnartz, H., and Tielens, A. G. G. M. (2015). Pore Evolution in Interstellar Ice Analogues. *A&A* 573, A16. doi:10.1051/0004-6361/201424466

Chaabouni, H., Minissale, M., Manicò, G., Congiu, E., Noble, J. A., Baouche, S., et al. (2012). Water Formation through O_2 + D Pathway on Cold Silicate and Amorphous Water Ice Surfaces of Interstellar Interest. *J. Chem. Phys.* 137, 234706–706. doi:10.1063/1.4771663

Christianson, D. A., and Garrod, R. T. (2021). Chemical Kinetics Simulations of Ice Chemistry on Porous versus Non-porous Dust Grains. *Front. Astro. Space Sci.* 8, 21. doi:10.3389/fspas.2021.643297

Dulieu, F., Minissale, M., and Bockelée-Morvan, D. (2017). Production of O_2 through Dismutation of H_2O_2 during Water Ice Desorption: a Key to Understanding Comet O_2 Abundances. *A&A* 597, A56. doi:10.1051/0004-6361/201628445

Fournier, J., Deson, J., Vermeil, C., and Pimentel, G. C. (1979). Fluorescence and Thermoluminescence of N_2O, CO, and CO_2 in an Argon Matrix at Low Temperature. *J. Chem. Phys.* 70, 5726–5730. doi:10.1063/1.437399

Fredon, A., Radchenko, A. K., and Cuppen, H. M. (2021). Quantification of the Role of Chemical Desorption in Molecular Clouds. *Acc. Chem. Res.* 54, 745–753. doi:10.1021/acs.accounts.0c00636

Garstang, R. H. (1951). Energy Levels and Transition Probabilities in P2 and P4 Configurations. *Monthly Notices R. Astronomical Soc.* 111, 115–124. doi:10.1093/mnras/111.1.115

Ghesquière, P., Mineva, T., Talbi, D., Theulé, P., Noble, J. A., and Chiavassa, T. (2015). Diffusion of Molecules in the Bulk of a Low Density Amorphous Ice from Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* 17, 11455–11468. doi:10.1039/c5cp00558b

Hagen, W., Tielens, A. G. G. M., and Greenberg, J. M. (1981). The Infrared Spectra of Amorphous Solid Water and Ice Ic between 10 and 140 K. *Chem. Phys.* 56, 367–379. doi:10.1016/0301-0104(81)80158-9

Hama, T., and Watanabe, N. (2013). Surface Processes on Interstellar Amorphous Solid Water: Adsorption, Diffusion, Tunneling Reactions, and Nuclear-Spin Conversion. *Chem. Rev.* 113, 8783–8839. doi:10.1021/cr4000978

He, J., Acharyya, K., and Vidali, G. (2016). Binding Energy of Molecules on Water Ice: Laboratory Measurements and Modeling. *ApJ* 825, 89. doi:10.3847/0004-637x/825/2/89

He, J., Emtiaz, S., and Vidali, G. (2018). Measurements of Diffusion of Volatiles in Amorphous Solid Water: Application to Interstellar Medium Environments. *ApJ* 863, 156. doi:10.3847/1538-4357/aad227

Ioppolo, S., Cuppen, H. M., and Linnartz, H. (2011). Surface Formation Routes of Interstellar Molecules: Hydrogenation Reactions in Simple Ices. *Rend. Fis. Acc. Lincei* 22, 211–224. doi:10.1007/s12210-011-0135-3

Ioppolo, S., Van Boheemen, Y., Cuppen, H. M., Van Dishoeck, E. F., and Linnartz, H. (2011). Surface Formation of CO_2 Ice at Low Temperatures. *Mon. Not. R. Astron. Soc.* 413, 2281–2287. doi:10.1111/j.1365-2966.2011.18306.x

Jenniskens, P., and Blake, D. F. (1994). Structural Transitions in Amorphous Water Ice and Astrophysical Implications. *Science* 265, 753–756. doi:10.1126/science.11539186

Karssemeijer, L., Ioppolo, S., van Hemert, M., van der Avoird, A., Allodi, M., Blake, G., et al. (2013). Dynamics of CO in Amorphous Water-Ice Environments. *Astrophys. J.* 781, 1–15. doi:10.1088/0004-637x/781/1/16

Keane, J. V., Tielens, A. G. G. M., Boogert, A. C. A., Schutte, W. A., and Whittet, D. C. B. (2001). Ice Absorption Features in the 5-8μm Region toward Embedded Protostars. *A&A* 376, 254–270. doi:10.1051/0004-6361:20010936

Kouchi, A., Furuya, K., Hama, T., Chigai, T., Kozasa, T., and Watanabe, N. (2020). Direct Measurements of Activation Energies for Surface Diffusion of CO and CO_2 on Amorphous Solid Water Using *In Situ* Transmission Electron Microscopy. *ApJL* 891, L22. doi:10.3847/2041-8213/ab78a2

Kouchi, A., Tsuge, M., Hama, T., Oba, Y., Okuzumi, S., Sirono, S.-i., et al. (2021). Transmission Electron Microscopy Study of the Morphology of Ices Composed of H_2O, CO_2, and CO on Refractory Grains. *ApJ* 918, 45. doi:10.3847/1538-4357/ac0ae6

Kumagai, N., Kawamura, K., and Yokokawa, T. (1994). An Interatomic Potential Model for H_2O: Applications to Water and Ice Polymorphs. *Mol. Simulation* 12, 177–186. doi:10.1080/08927029408023028

Lee, M. W., and Meuwly, M. (2014). Diffusion of Atomic Oxygen Relevant to Water Formation in Amorphous Interstellar Ices. *Faraday Discuss.* 168, 205–222. doi:10.1039/c3fd00160a

Lee, M. W., and Meuwly, M. (2011). On the Role of Nonbonded Interactions in Vibrational Energy Relaxation of Cyanide in Water. *J. Phys. Chem. A.* 115, 5053–5061. doi:10.1021/jp202503m

Minissale, M., Congiu, E., Baouche, S., Chaabouni, H., Moudens, A., Dulieu, F., et al. (2013). Quantum Tunneling of Oxygen Atoms on Very Cold Surfaces. *Phys. Rev. Lett.* 111, 053201. doi:10.1103/PhysRevLett.111.053201

Minissale, M., Congiu, E., and Dulieu, F. (2016). Direct Measurement of Desorption and Diffusion Energies of O and N Atoms Physisorbed on Amorphous Surfaces. *A&A* 585, A146. doi:10.1051/0004-6361/201526702

Minissale, M., Congiu, E., Manicò, G., Pirronello, V., and Dulieu, F. (2013). CO_2 formation on Interstellar Dust Grains: a Detailed Study of the Barrier of the CO + O Channel. *A&A* 559, A49. doi:10.1051/0004-6361/201321453

Minissale, M., Fedoseev, G., Congiu, E., Ioppolo, S., Dulieu, F., and Linnartz, H. (2014). Solid State Chemistry of Nitrogen Oxides - Part I: Surface Consumption of NO. *Phys. Chem. Chem. Phys.* 16, 8257–8269. doi:10.1039/c3cp54917h

Minissale, M., Moudens, A., Baouche, S., Chaabouni, H., and Dulieu, F. (2016). Hydrogenation of CO-bearing Species on Grains: Unexpected Chemical Desorption of CO. *Mon. Not. R. Astron. Soc.* 458, 2953–2961. doi:10.1093/mnras/stw373

Minissale, M., Nguyen, T., and Dulieu, F. (2019). Experimental Study of the Penetration of Oxygen and Deuterium Atoms into Porous Water Ice. *A&A* 622, A148. doi:10.1051/0004-6361/201833830

Molpeceres, G., Kästner, J., Fedoseev, G., Qasim, D., Schömig, R., Linnartz, H., et al. (2021). Carbon Atom Reactivity with Amorphous Solid Water: H_2O-Catalyzed Formation of H_2CO. *J. Phys. Chem. Lett.* 12, 10854–10860. doi:10.1021/acs.jpclett.1c02760

Nagy, T., and Yosa Reyes, J. (2014). Multisurface Adiabatic Reactive Molecular Dynamics. *J. Chem. Theor. Comput.* 10, 1366–1375. doi:10.1021/ct400953f

Noble, J. A., Dulieu, F., Congiu, E., and Fraser, H. J. (2011). CO_2 Formation in Quiescent Clouds: An Experimental Study of the CO + OH Pathway. *ApJ* 735, 121. doi:10.1088/0004-637x/735/2/121

Noble, J., Congiu, E., Dulieu, F., and Fraser, H. (2012). Thermal Desorption Characteristics of CO, O_2 and CO_2 on Non-porous Water, Crystalline Water and Silicate Surfaces at Submonolayer and Multilayer Coverages. *Mon. Not. R. Astron. Soc.* 421, 768–779.

Oba, Y., Miyauchi, N., Hidaka, H., Chigai, T., Watanabe, N., and Kouchi, A. (2009). Formation of Compact Amorphous H_2O Ice by Codeposition of Hydrogen Atoms with Oxygen Molecules on Grain Surfaces. *ApJ* 701, 464–470. doi:10.1088/0004-637x/701/1/464

Oba, Y., Watanabe, N., Kouchi, A., Hama, T., and Pirronello, V. (2010). Experimental Study of CO_2 Formation by Surface Reactions of Non-energetic OH Radicals with CO Molecules. *ApJ* 712, L174–L178. doi:10.1088/2041-8205/712/2/l174

Pezzella, M., and Meuwly, M. (2019). O_2 Formation in Cold Environments. *Phys. Chem. Chem. Phys.* 21, 6247–6255. doi:10.1039/c8cp07474g

Pezzella, M., Unke, O. T., and Meuwly, M. (2018). Molecular Oxygen Formation in Interstellar Ices Does Not Require Tunneling. *J. Phys. Chem. Lett.* 9, 1822–1826. doi:10.1021/acs.jpclett.8b00328

Plattner, N., and Meuwly, M. (2008). Atomistic Simulations of CO Vibrations in Ices Relevant to Astrochemistry. *ChemPhysChem* 9, 1271–1277. doi:10.1002/cphc.200800030

Qasim, D., Fedoseev, G., Chuang, K.-J., He, J., Ioppolo, S., van Dishoeck, E. F., et al. (2020). An Experimental Study of the Surface Formation of Methane in Interstellar Molecular Clouds. *Nat. Astron.* 4, 781–785. doi:10.1038/s41550-020-1054-y

Qasim, D., Lamberts, T., He, J., Chuang, K.-J., Fedoseev, G., Ioppolo, S., et al. (2019). Extension of the HCOOH and CO_2 Solid-State Reaction Network during the CO Freeze-Out Stage: Inclusion of H_2CO. *A&A* 626, A118. doi:10.1051/0004-6361/201935068

Romanzin, C., Ioppolo, S., Cuppen, H. M., van Dishoeck, E. F., and Linnartz, H. (2011). Water Formation by Surface O_3 Hydrogenation. *J. Chem. Phys.* 134, 084504–504. doi:10.1063/1.3532087

Roser, J. E., Vidali, G., Manicò, G., and Pirronello, V. (2001). Formation of Carbon Dioxide by Surface Reactions on Ices in the Interstellar Medium. *Astrophys. J. Lett.* 555, L61–L64. doi:10.1086/321732

Ruffle, D. P., and Herbst, E. (2001). New Models of Interstellar Gas-Grain Chemistry - III. Solid CO_2. *Mon. Not. R. Astron. Soc.* 324, 1054–1062. doi:10.1046/j.1365-8711.2001.04394.x

Schmidt, F., Swiderek, P., and Bredehöft, J. H. (2019). Formation of Formic Acid, Formaldehyde, and Carbon Dioxide by Electron-Induced Chemistry in Ices of Water and Carbon Monoxide. *ACS Earth Space Chem.* 3, 1974–1986. doi:10.1021/acsearthspacechem.9b00168

Stief, L. J., Payne, W. A., and Klemm, R. B. (1975). A Flasch Photolysis-Resonance Fluorescence Study of the Formation of $O(^1D)$ in the Photolysis of Water and the Reaction of $O(^1D)$ with H_2, Ar, and He –. *J. Chem. Phys.* 62, 4000–4008. doi:10.1063/1.430323

Stock, G. (2009). Classical Simulation of Quantum Energy Flow in Biomolecules. *Phys. Rev. Lett.* 102, 118301. doi:10.1103/physrevlett.102.118301

Tsuge, M., Hidaka, H., Kouchi, A., and Watanabe, N. (2020). Diffusive Hydrogenation Reactions of CO Embedded in Amorphous Solid Water at Elevated Temperatures ~70 K. *ApJ* 900, 187. doi:10.3847/1538-4357/abab9b

Upadhyay, M., Pezzella, M., and Meuwly, M. (2021). Genesis of Polyatomic Molecules in Dark Clouds: CO_2 Formation on Cold Amorphous Solid Water. *J. Phys. Chem. Lett.* 12, 6781–6787. doi:10.1021/acs.jpclett.1c01810

Veliz, J. C. S. V., Koner, D., Schwilk, M., Bemish, R. J., and Meuwly, M. (2021). The $C(^3P) + O_2(^3\Sigma_g^-) \leftrightarrow CO_2 \leftrightarrow CO(^1\Sigma^+) + O(^1D)/O(^3P)$ Reaction: Thermal and Vibrational Relaxation Rates from 15 K to 20000K. *Phys. Chem. Chem. Phys.* 23, 251–263. doi:10.1039/d1cp01101d

Watanabe, N., and Kouchi, A. (2002). Measurements of Conversion Rates of CO to CO_2 in Ultraviolet-induced Reaction of $D_2O(H_2O)$/CO Amorphous Ice. *ApJ* 567, 651–655. doi:10.1086/338491

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 25 February 2022
doi: 10.3389/fchem.2022.836994

Caging Polycations: Effect of Increasing Confinement on the Modes of Interaction of Spermidine^{3+} With DNA Double Helices

Tudor Vasiliu[1], *Francesca Mocci*[2]*, *Aatto Laaksonen*[1,2,3,4,5]*, *Leon De Villiers Engelbrecht*[2] *and Sergiy Perepelytsya*[6]*

[1]Centre of Advanced Research in Bionanoconjugates and Biopolymers "Petru Poni" Institute of Macromolecular Chemistry, Iasi, Romania, [2]Dipartimento di Scienze Chimiche e Geologiche, Cagliari University, Cagliari, Italy, [3]Division of Energy Science, Energy Engineering, Luleå University of Technology, Luleå, Sweden, [4]Division of Physical Chemistry, Department of Materials and Environmental Chemistry, Arrhenius Laboratory, Stockholm University, Stockholm, Sweden, [5]State Key Laboratory of Materials-Oriented and Chemical Engineering, Nanjing Tech University, Nanjing, China, [6]Bogolyubov Institute for Theoretical Physics of the NAS of Ukraine, Kyiv, Ukraine

OPEN ACCESS

Edited by:
Heribert Reis,
National Hellenic Research Foundation, Greece

Reviewed by:
Pascal Auffinger,
UPR9002 - CNRS - Strasbourg, France
Antonio Monari,
Université de Paris, France

***Correspondence:**
Francesca Mocci
fmocci@unica.it
Aatto Laaksonen
aatto@mmk.su.se
Sergiy Perepelytsya
perepelytsya@bitp.kiev.ua

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: 16 December 2021
Accepted: 04 February 2022
Published: 25 February 2022

Citation:
Vasiliu T, Mocci F, Laaksonen A, Engelbrecht LDV and Perepelytsya S (2022) Caging Polycations: Effect of Increasing Confinement on the Modes of Interaction of Spermidine^{3+} With DNA Double Helices. Front. Chem. 10:836994. doi: 10.3389/fchem.2022.836994

Polyamines have important roles in the modulation of the cellular function and are ubiquitous in cells. The polyamines putrescine^{2+}, spermidine^{3+}, and spermine^{4+} represent the most abundant organic counterions of the negatively charged DNA in the cellular nucleus. These polyamines are known to stabilize the DNA structure and, depending on their concentration and additional salt composition, to induce DNA aggregation, which is often referred to as condensation. However, the modes of interactions of these elongated polycations with DNA and how they promote condensation are still not clear. In the present work, atomistic molecular dynamics (MD) computer simulations of two DNA fragments surrounded by spermidine^{3+} (Spd^{3+}) cations were performed to study the structuring of Spd^{3+} "caged" between DNA molecules. Microsecond time scale simulations, in which the parallel DNA fragments were constrained at three different separations, but allowed to rotate axially and move naturally, provided information on the conformations and relative orientations of surrounding Spm^{3+} cations as a function of DNA-DNA separation. Novel geometric criteria allowed for the classification of DNA-Spd^{3+} interaction modes, with special attention given to Spd^{3+} conformational changes in the space between the two DNA molecules (caged Spd^{3+}). This work shows how changes in the accessible space, or confinement, around DNA affect DNA-Spd^{3+} interactions, information fundamental to understanding the interactions between DNA and its counterions in environments where DNA is compacted, e.g. in the cellular nucleus.

Keywords: polyamine, condensation, DNA, counterion, molecular dynamics

1 INTRODUCTION

The most common structural organization of the DNA macromolecule in cells consists of two polynucleotide chains wound in a double helix (Watson and Crick, 1953). The nucleotides occurring in DNA are adenine (A), thymine (T), guanine (G) and cytosine (C), the sequence of which codes genetic and structural information. In aqueous solution at physiological pH, each nucleotide is negatively charged due to deprotonation of the phosphate groups. The phosphates, together with the sugar rings, constitute the DNA backbone, which is the region most exposed to the solvent in DNA helices. The double helix is stabilized by positively-charged ions in the solvent media (counterions), typically metal ions (Na^+, K^+, Mg^{2+}) and charged organic molecules like polyamines (PAs) (Franklin and Gosling, 1953; Ames and Dubin, 1960; Blagoi et al., 1991; Maleev et al., 1993; Mocci and Laaksonen, 2012; Mocci et al., 2021). The counterions bind to different regions of the double helix (minor and major grooves, phosphate groups), and these interactions are essential for the organization of the macromolecule in high order structures (Mocci and Laaksonen, 2012). In the cells of living organisms, DNA is organized in a highly compact form, wrapping around histone proteins in the nucleosome core particles (NCPs), which further assemble, forming chromatin fibers (Saenger, 1984; Schlick, 2002). The PAs are involved in the neutralization of DNA in chromatin and are essential for NCPs formation (van Dam et al., 2002; Korolev et al., 2012). Multivalent counterions are present also in viruses, neutralizing the negatively-charged DNA and RNA macromolecules, and allowing them to pack densely inside the small volume of the viral capsid (Ames and Dubin, 1960; Roos et al., 2007; Carrivain et al., 2012; Mounce et al., 2017; Firpo and Mounce, 2020). The interactions of positively-charged PAs with DNA have a significant biological effect, and are also involved in some emerging biotechnological applications (Korolev et al., 2001; D'Agostino, 2018; Perepelytsya et al., 2019; Mocci et al., 2021; Vasiliu et al., 2021). The study of the role of PAs in the structural organization of DNA in living organisms is of paramount importance for the understanding of key biological functions.

While the binding of monoatomic metal ions to DNA have been extensively studied, and there exists a vast amount of data showing the character of their interaction with nucleic acids (NAs) (Saenger, 1984; Blagoi et al., 1991; Maleev et al., 1993; Young et al., 1997; McConnell and Beveridge, 2000; Mocci and Saba, 2003; Mocci et al., 2004; Ponornarev et al., 2004; Várnai and Zakrzewska, 2004; Marincola et al., 2009; Mocci and Laaksonen, 2012; Lavery et al., 2014; Pasi et al., 2015; Atzori et al., 2016; Dans et al., 2016; Perepelytsya, 2018; Zdorevskyi and Perepelytsya, 2020), the binding of molecular counterions to NAs have been less studied. PAs are known to affect the dynamics and structure of the DNA double helix, inducing condensation (Chattoraj et al., 1978; Gosule and Schellman, 1978; Bloomfield, 1996; Kornyshev et al., 2007; Estévez-Torres and Baigl, 2011). Experimental data show that counterions with charge greater than 2 induce compaction of DNA; this effect depends on the type and concentration of the counterions (Estévez-Torres and Baigl, 2011). The PAs spermidine^{3+} and spermine^{4+} induce the condensation of DNA when the PA concentration is sufficient to completely neutralize the NA. At low concentration of spermine^{4+}, or some of its isomers, the effects of enhancement and inhibition of gene activity were established for the case of low and high concentrations the PAs, respectively (Kanemura et al., 2018; Kitagawa et al., 2021). The problem of understanding the molecular mechanisms of PA-induced DNA condensation belongs to the frontiers between chemistry, biology and physics.

To describe the DNA condensation induced by multivalent metal ions and PAs, different theoretical models have been proposed [see the reviews (Bloomfield, 1996; Kornyshev et al., 2007)]. The collapse of the DNA macromolecular chain was considered as a coil-globule transition in a statistical mechanics approaches (Post and Zimm, 1979) that was developed further for different cases of the DNA state [see the review (Bloomfield, 1997)]. The attraction between DNA double helices, eventually resulting in condensation, was shown to arise due to the interaction between the polyanionic macromolecules with the mobile counterions, which in the case of ion charge ≥3 form a structured system between two macromolecules resembling Wigner crystal (Rouzina and Bloomfield, 1996). Such a model describes the character of DNA-DNA attraction in the case of small multivalent ions, but the attraction that appears in the case of some bivalent metal ions (Mn^{2+} or Ca^{2+}) and elongated PA molecules are not clear (Kornyshev et al., 2007). The localization of counterions in the grooves of the double helix are taken into consideration in the electrostatic "zipper" model (Kornyshev and Leikin, 1999). In this model, the enhanced attraction between different helices appears due to the juxtaposition of negatively charged phosphate groups of the double helix backbone with the positively charged counterions in the DNA grooves. In spite of great efforts by scientists in developing these theoretical models, the microscopic mechanism of DNA condensation induced by flexible highly charged PAs (spermidine^{3+} and spermine^{4+}) and their aggregates is not yet clear. Many aspects of these processes are still to be determined, such as how the interactions are dependent on the concentration of the PA, or by the particular nucleotides sequence, or by the distance between DNA molecules (or portion of the same long DNA fragment). In this context, molecular dynamics (MD) simulation methods can be a powerful tool to obtain detailed information.

The first MD simulations of DNA with PAs showed that these counterions strongly bind to the double helix, substituting monovalent metal ions (Korolev et al., 2001, 2004; van Dam et al., 2002). As shown also by other simulations, the modes of interaction between PAs and DNA double helix are governed by noncovalent interactions and are extremely variable, affected by the PAs charge and length (Korolev et al., 2001, 2004; van Dam et al., 2002; Bignon et al., 2017; Perepelytsya et al., 2019), and yet to be properly classified. Interacting with DNA, the PAs induce changes in the double helix structure, in particular narrowing of the minor groove (Korolev et al., 2004). While the interactions of PAs with the DNA were for a long time considered non-sequence specific, recent experiments (Patel and Anchordoquy, 2006; Kabir and Suresh Kumar, 2013) and MD simulations (Perepelytsya et al., 2019; Mocci et al., 2021) have shown that putrescine^{2+},

spermidine^{3+}, and spermine^{4+} prefer to be localized in the DNA minor groove of the AT-rich regions. The preferential localization of PAs in the minor groove of the double helix is modulated by the sequence of nucleotides determining the natural narrowing of the minor groove. The aggregation of DNA induced by PAs was found to be stabilized by PA bridges formed in different regions between the double helices (Dai et al., 2008). MD simulation studies of DNA arrays with spermine^{4+} also revealed the sequence-specific effects of PA interactions with DNA and the formation of PA cross-links between two double helices (Yoo and Aksimentiev, 2016; Yoo et al., 2016). The spermidine^{3+} cross-linking between two DNA 22-mers in water solution was observed in a recent MD simulation (Mocci et al., 2021).

The goal of the present work is to study how the spatial organization and dynamics of spermidine^{3+} (Spd^{3+}) molecules at the interface between two DNA fragments depends on the separation between the DNA helices. To tackle this problem, atomistic MD simulations were performed for three model systems, each consisting of two 22-base-pairs-long DNA double helices restrained at different interhelix separation distances. A nucleosomal 146-base-pairs-long DNA fragment, with coordinates taken from an X-ray structure of an NCP, was also simulated, constraining it to the experimental geometry and helix separation. The DNA in the NCP system constitutes an ideal model of DNA compaction at a level found in cellular environments. It is important to note that in the NCP both the separation between the DNA helices and the high density of the positive charge of the histones are expected to influence the Spd^{3+}/DNA interactions. To disentangle the effects of the DNA separation from those due to the presence of the positively charged core of the NCP, and to allow a comparison with the other parallel double helix models used in this investigation, we chose to omit the protein core. Its effect will be addressed in future studies. The details of the MD simulations are described in **Section 2**. The structural organization, and the dynamics, of Spd^{3+} around the DNA double helices are analyzed and discussed in **Section 3**.

The MD simulations show strong attraction between the DNA helices induced by Spd^{3+} and highlight how the increasing confinement of the PA molecules, obtained by decreasing the distance between DNA helices, affect their structural and dynamical behavior.

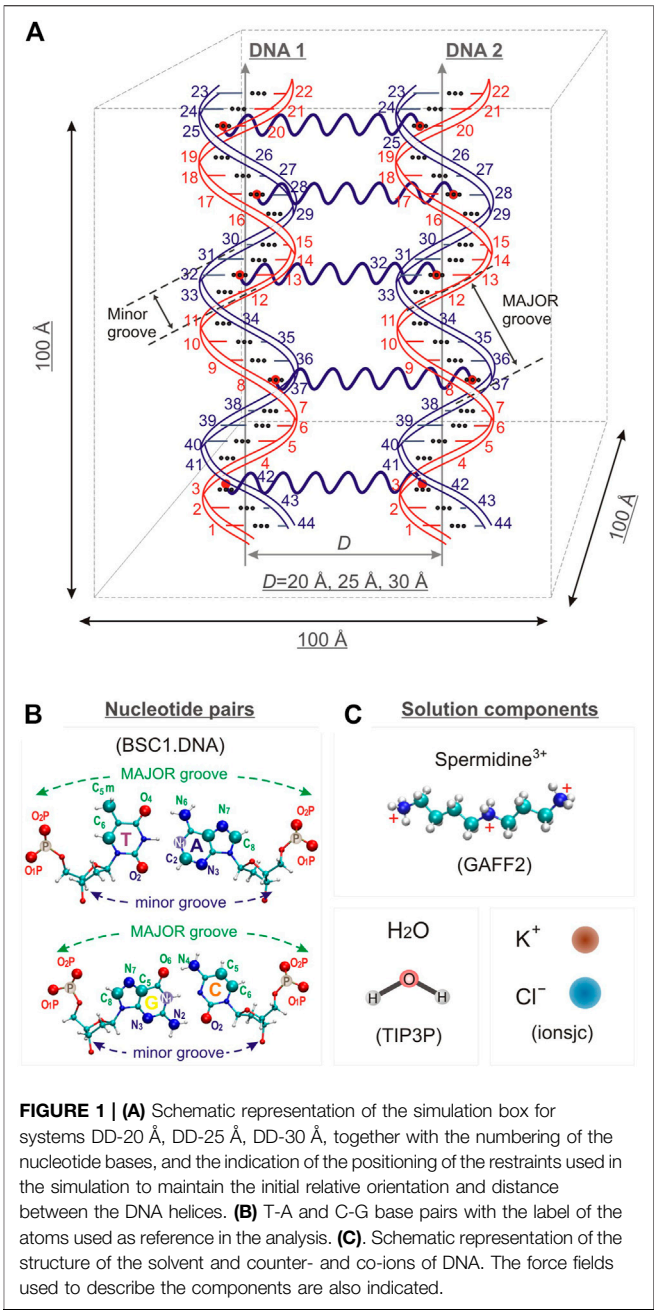


FIGURE 1 | (A) Schematic representation of the simulation box for systems DD-20 Å, DD-25 Å, DD-30 Å, together with the numbering of the nucleotide bases, and the indication of the positioning of the restraints used in the simulation to maintain the initial relative orientation and distance between the DNA helices. **(B)** T-A and C-G base pairs with the label of the atoms used as reference in the analysis. **(C)**. Schematic representation of the structure of the solvent and counter- and co-ions of DNA. The force fields used to describe the components are also indicated.

2 METHODS

Three model systems, each containing two identical DNA molecules consisting of 22 base pairs (bps) were constructed by placing the DNA molecules parallel to each other at distances between their axis of 20, 25 and 30 Å, respectively; Spd^{3+} and water molecules were then added. The systems will be referred to as DD-20Å, DD-25Å, and DD-30Å. In addition, to compare the results from the model systems described above with an experimental parallel DNA helix, we performed a simulation of a nucleosomal DNA sequence, 146 bps long, with configuration taken from an experimental NCP X-ray structure, where it is known that the wrapping of DNA around the histones leads the two segments of the duplex to align in parallel (for approximately 66 bps). Schematic representations of the simulation boxes with two DNA duplexes and the system with nucleosomal DNA are shown in **Figures 1**, **2**, respectively. The composition of each system is detailed in **Table 1**.

The initial structure of Spd^{3+} was built with the Avogadro software (Hanwell et al., 2012; Avogadro, 2018), and the corresponding general AMBER force field (GAFF2) parameters (Wang et al., 2004) were generated with the Antechamber software comprised in the AmberTools 18 package (Case et al., 2018). The atomic partial charges of the Spd^{3+} molecule

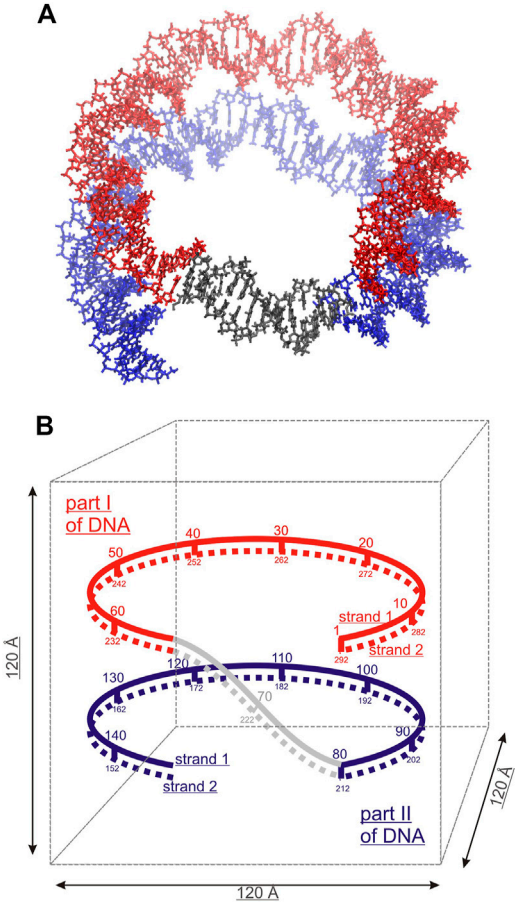


FIGURE 2 | Simulated system of nucleosomal DNA. **(A)** Licorice representation of the nucleosomal DNA: red and blue colors indicate the overlapping parts of the DNA double helix, non-overlapping DNA colored grey. **(B)** Schematic structure of the nucleosomal DNA in the simulation box. The numbering of the nucleotide pairs is indicated.

were calculated using the RESP methodology and are presented in **Supplementary Figure S1**.

The nucleic acid builder (NAB) tool contained in the AmberTools 18 package was used to create the 22 bps DNA in the Arnott *B*-DNA canonical structure with the sequence d(CGCGAATTCGCGCGAATTCGCG), containing two motifs of the Drew-Dickerson dodecamer (Drew et al., 1981) sequence, with two A-tracts underlined. This sequence, containing both A-tracts and CG-rich regions, constitute an important model system for the study of interaction of PAs with the DNA double helix (Perepelytsya et al., 2019; Mocci et al., 2021). The structure of DNA in an NCP was retrieved from the experimental structure of Tsunaka et al. (Tsunaka et al., 2005) as deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (ID: 2cv5). For a comparison with the other simulated systems, the protein core of the NCP was omitted in the simulations, and the DNA structure was constrained to the experimental structure throughout the entire simulation.

The BSC1. DNA (Ivani et al., 2016) AMBER force field was used for DNA, the TIP3P model was used for water (Jorgensen et al., 1983), while the ionsjc parameters optimized for this water model were used for K^+ and Cl^- ions (Joung and Cheatham, 2008).

All simulations were performed using the GROMACS 2020 software package (Abraham et al., 2015). The simulations were done at constant temperature and pressure. The temperature was set to 298 K, and was controlled using the Nosé-Hoover thermostat (Nosé, 1984; Hoover, 1985). The pressure was set to 1 bar and was controlled using the Parrinello-Rahman barostat (Parrinello and Rahman, 1981). The length of all bonds between hydrogen and other atoms was constrained using the LINCS algorithm (Hess et al., 1997). The smooth particle mesh Ewald method (Darden et al., 1993) was used to calculate the long-range electrostatic interactions. The cut-offs for the switching and the long-range interactions were set to 10 Å, and the Fourier spacing was set to 1.2 Å. The length of each production simulation trajectory was 500 ns.

To restrain the distance between DNA molecules in the DD-20Å, DD-25Å and DD-30Å systems, we used the center-of-mass (COM) pulling method with an umbrella potential (Abraham et al., 2015). In detail, we placed five harmonic springs between the N_1 atoms of parallel bps in the two DNA molecules, as depicted in **Figure 1A**. The main advantage of using this restraining method, instead of the classic position restraints, is that the DNA molecules can now move freely (move around the simulation box, rotate around their own helix axes, compress, undergo sequence dependent structural modifications, elongate or bend). We chose this restraining method to eliminate any artifacts in the mode of interaction with the Spd^{3+} that could arise due to the rigid nature of the position restraints. These restraints allowed us to keep the DNA fragments parallel to each other and inhibit the rotation between their axis. We kept the fragments parallel for two reasons: firstly, the parallel conformation occurs naturally in

TABLE 1 | Simulation box details.

System name	Number of DNA molecules	Number of bps	Distance between DNA segments[a]	Number of Spd^{3+} molecules	Number of water molecules	Number of ions (K^+/Cl^-)
DD-20Å	2	22	20 Å	30	31,308	84/90
DD-25Å	2	22	25 Å	30	31,308	84/90
DD-30Å	2	22	30 Å	30	31,308	84/90
Nucleosome	1	146	25–28 Å	100	50,583	156/166

[a]*In the case of nucleosomal DNA, the reported distance range refers to the separation between the part I and part II, as defined in* ***Figure 2****.*

nucleosomes and DNA fibers; secondly the parallel conformation provides a greater area with a controlled distance between the DNA fragments, which translates into a higher number of caged Spd^{3+}, which in turn increases the sampling of the conformational space, reducing the need to repeat or greatly extend the simulations. For the nucleosomal DNA simulation we used the classic position restraints, restraining all heavy atoms, because it was important to keep the DNA fragment in the specific experimental "nucleosome conformation". Due to its size and specific conformation, using COM pulling restraints on this DNA fragment would have implied using at least 15 springs, and the fine tuning of the parameters describing all these springs, to have them maintain the nucleosome conformation, proved to be an unfeasible and inefficient task. The snapshots of the nucleosomal DNA are shown in **Supplementary Figure S5**.

The analysis of Spd^{3+} distribution was done in terms of radial distribution functions (RDFs) calculated for the PA heavy atoms and selected atoms of the DNA double helix:

$$g(r) = \lim_{\Delta r \to 0} \frac{p(r)}{4\pi r^2 \Delta r N_p / V} \quad (1)$$

where $p(r)$ is the average number of particles that is found at the distance r within a shell with thickness Δr; N_p is the number of pairs of selected atoms for which the RDF is calculated, and V is the system volume. In our calculations, the shell thickness Δr has been taken equal to 0.1 Å. The average number of particles within a given distance r (coordination number) can be determined by the direct integration of the RDF.

RDFs for Spd^{3+} in the minor and major grooves (RDF_{MIN} and RDF_{MAJ}) were calculated between all heavy atoms of Spd^{3+} and the atoms N_3, N_2, O_2, and C_8, N_7, C_5, O_6, N_4, C_5, C_6 of the nucleotide bases, for the minor and major grooves respectively. The distribution of Spd^{3+} with respect to the phosphate groups was characterized by the RDF_{PH} computed for the Spd^{3+} heavy atoms, with the oxygen atoms O_1P and O_2P of the phosphate groups taken as reference atoms (**Figure 1B**). To analyze the orientation of the DNA double helices with respect to each other, the RDFs (hereafter RDF_{DD}s) of the P atoms of DNA1 with respect to the P atoms of DNA2, and vice versa, were calculated. In the case of nucleosomal DNA, where the parallel segments of DNA belong to the same molecule, we indicate with DNA1 and DNA2 the parallel segments with the nucleotide pairs 1–67 and 80–146, respectively (**Figure 2B**).

The visual inspection of the simulated systems and the analysis of the simulation trajectories were performed using the VMD program and plug-ins implemented in the software package (William et al., 1996). In particular, to characterize the structure of the DNA double helix, the parameters of the minor and major grooves were calculated using VMD plug-in do_x2dna (Lu and Olson, 2003; Kumar and Grubmüller, 2015). The groove widths were calculated according to the definition of El Hassan and Calladine (El Hassan and Calladine, 1998). The RDFs were calculated using the VMD plug-in (Levine et al., 2011) implemented in VMD. The quantitative analysis of the distribution and residence time of Spd^{3+} was performed with VMD using in-house scripts.

3 RESULTS

3.1 Interaction Between the DNA Helices

To characterize the mutual orientation of the DNA helices, the RDFs between the P atoms of one DNA molecule with respect to the P atoms of the other DNA molecule, were calculated for all simulated systems. These RDFs will be indicated as RDF_{DD}s in the text and are shown in **Figure 3**. To identify the nucleotides among which intermolecular contacts are most probable, the RDF calculated for each P atom was integrated up to 6.4 Å, to obtain the coordination number (CN) (**Figure 3**).

The RDF_{DD}s represented in **Figure 3** are obtained averaging the RDF_{DD}s for the two parallel helices (or DNA portions in case of nucleosomal DNA). The RDF_{DD}s calculated separately for each DNA molecule are shown in **Supplementary Figure S6**. The first maximum of RDF_{DD}s of the systems DD-20Å and DD-25Å corresponds to close contacts between the parallel double helixes. This type of interaction, detailed in the following, involve the NH_2^+ or NH_3^+ group(s) of one or more Spd^{3+} molecules bridging the O atoms of the phosphate groups of the two double helices (O-HNH-O contacts). In the case of the system DD-20Å, where the distance between the two helices is the smallest among the simulated systems, the RDFs are characterized by a maximum splitting into two sharp peaks at about 5 Å and 6 Å, and a minimum at about 8 Å (**Figure 3A**). The broad band from 8 Å to 14 Å is due to other phosphate groups of the DNA chains; this band is not informative for the analysis of DNA-DNA interaction and is not considered further. The RDF_{DD}s for the system DD-25Å have a rather regular form with a prominent maximum at about 5.5 Å and a minimum at ca. 6.5 Å (**Figure 3B**). As in the case of DD-20Å system, the first maximum arises from the amino group(s)-mediated DNA-DNA contacts, of the O-HNH-O type. This implies that even in the case of systems where the DNA molecules have restraints imposing 25 Å separation between the center of selected base pairs, the presence of Spd^{3+} in the region between the two DNA molecules induces an attraction between them, which, in turn, induces helix bending to achieve the amino group-mediated close contact. The intensity of the first peak is, however, greatly reduced compared to the DD-20Å system. In the case of the system DD-30Å, there are no prominent maxima at short distances, since the macromolecules are held at too large a separation for O-HNH-O DNA-DNA contacts to be formed (**Figure 3C**). Similar to what was observed for the DD-30Å system, the RDF_{DD}s of the nucleosome DNA (**Figure 3D**), are characterized by very low values at short distances, and a gradual increase with increasing distance, without any relevant maximum. It is important to note that the atoms of nucleosome DNA fragment are fixed in our simulation and

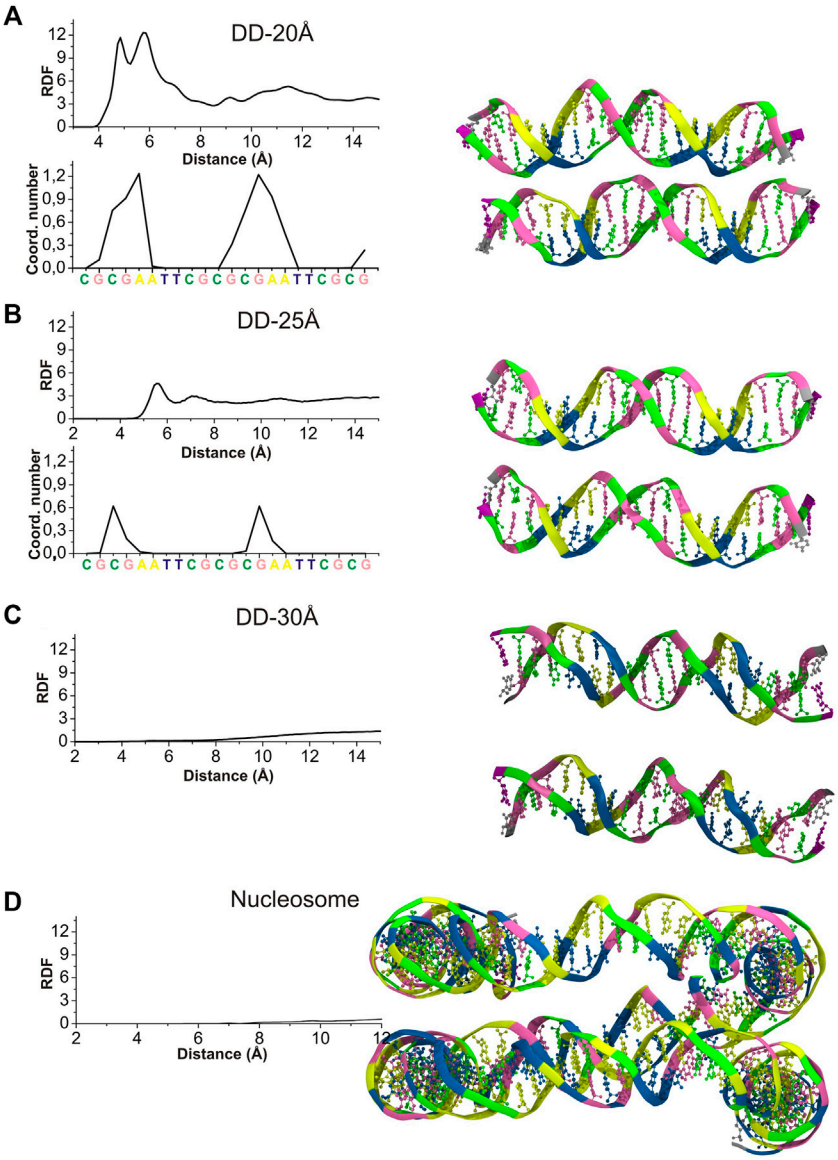


FIGURE 3 | Left: RDF_{DD}s (continuous lines) calculated between the P atoms of phosphate groups belonging to different DNA molecules (or parallel segments for nucleosome DNA), and the values of integrals of the RDFs calculated for each individual phosphate group up to 6.4 Å, revealing the sites were the DNA-DNA contacts are most probable. **(A)** System DD-20 Å **(B)** System DD-25 Å. **(C)** System DD-30 Å **(D)** Nucleosome DNA system. Right: The snapshots from the MD simulation of each system are represented at the right side of the corresponding RDF. Coloring scheme: A, yellow; T, blue; C, green; G, pink. Spd^{3+}, water and ions have been omitted for clarity.

its RDF_{DD} describe the distribution that corresponds to the X-ray experimental structure.

To determine whether inter-helix contacts can occur with the same probability at any position of the helices, or if on the contrary some regions are favored, the RDF_{DD}s were calculated separately for each nucleotide, and the CNs were calculated using an integration limit of 6.5 Å. The obtained CNs are shown in the insets of **Figures 3A,B** for DD-20Å and DD-25Å, while for DD-30Å and the nucleosome DNA fragment the helices are too far apart, and no contacts are observed. In the case of the DD-20Å system, the nucleotide-specific CN varies greatly along the double helix, with the highest CNs observed in the regions preceding the A-tracts, including the initial portion of the A-tracts, whereas in other regions there are essentially no contacts between the DNA molecules. The CN peaks are localized in the regions where the minor grooves of DNA1 and DNA2 are facing each other, as shown in the snapshot in **Figure 3A**, forming a "sandwich structure" of minor groove—Spd^{3+} molecules—minor groove, that will be described in the following sections devoted to the study of the interactions

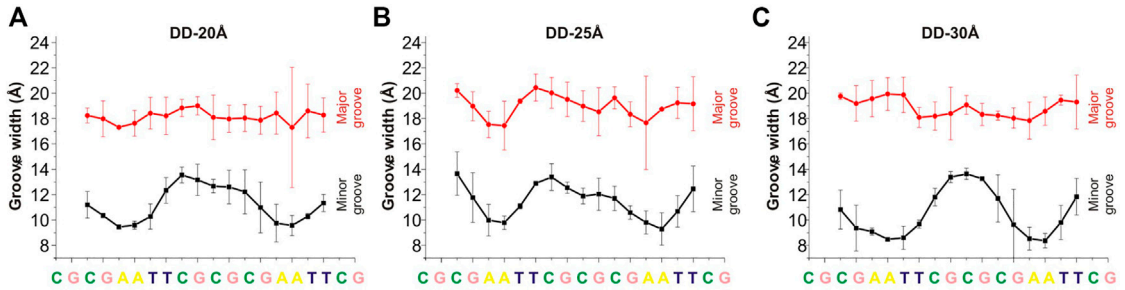


FIGURE 4 | Width of the minor and major grooves for each base pair, in **(A)** DD-20 Å, **(B)** DD-25 Å, **(C)** DD-30 Å. The values of the groove width are averaged over two DNA duplexes. The widths of the grooves for DNA1 and DNA2 are shown in **Supplementary Figure S7**.

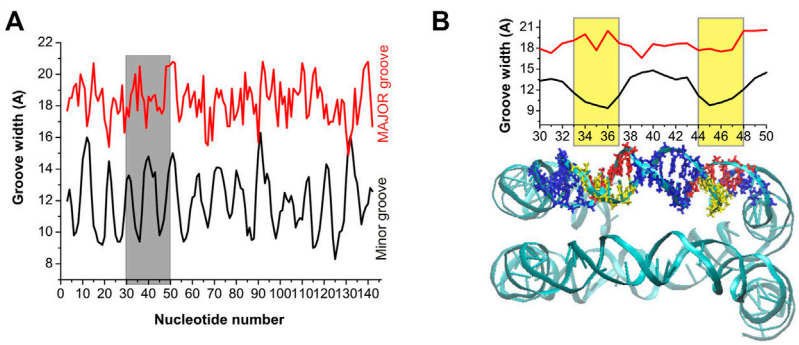


FIGURE 5 | (A) Minor and major groove width across the nucleosomal DNA. **(B)** Expansion of the grey band portion of Panel **(A)**, and a representation of the nucleosomal DNA in which the base pairs included in the width calculation are highlighted with a stick representation. The yellow bands, in the groove width plot, mark the region of DNA with the narrowest minor groove, and the corresponding base pairs are highlighted in yellow and red in the DNA snapshot.

with between DNAs and Spd^{3+}. In the DD-25Å system, due to the increased separation, the DNA molecules do not adopt the minor groove-to-minor groove orientation, which leads us to believe that the "sandwich structure" requires a certain separation distance. The contacts between the two helices occur approximately in the same regions as for DD-20Å, with much smaller CNs. In system DD-30Å there are no close contacts between the two DNA molecules, although it must be mentioned that, while fine-tuning the rigidity of the springs to maintain the separation, we noticed that when the two DNA came into contact after more than 500 ns of simulations (due to springs being too lax), we obtained the "sandwich structure". However, in the present paper we have focused our attention on the portion of the MD trajectories in which the separation between the helices is kept to 20, 25 or 30 Å, since we are interested in how this separation influences the interactions with Spd^{3+}. Although the nucleosome system has a typical separation between parallel DNA portions that is intermediate between DD-25Å and DD-30Å, the presence of the minor groove to minor groove orientation is still evident. It must be noted that, while in the DD systems this reciprocal orientation of the helices only appeared after the Spd^{3+} molecules had pulled the two DNA helices closer together, the same orientation is observed in the simulated nucleosomal DNA system, the structure of which was constrained to maintain the crystallographic structure, which did not contain Spd^{3+}. Most likely, the positively charged amino groups in the lateral chains of histone protein amino acids are responsible for the occurrence of this orientation in the crystallographic structures of this and other NCPs.

3.2 DNA Groove Width

An important structural parameter of the DNA double helix structure is the width of the groove, which is dependent on the base pairs sequence and is highly relevant and interconnected with the interactions with the surrounding molecules: a smaller groove width is connected to stronger interactions with PAs (Perepelytsya et al., 2019). In **Figures 4**, **5** are reported the average values of minor and major groove widths along the double helices, and it can be seen that the minor groove width varies within the range 8–15 Å, while the major groove within 15–22 Å. The fluctuations of the groove width are quantified with the error bars, with exception of the nucleosome DNA system, where a DNA static structure was simulated, and therefore no fluctuation is possible.

In the case of the DD-20Å, DD-25Å and DD-30Å systems, the minor groove width is narrower in the A-tracts regions. Such dependence of the minor groove width on nucleotide sequence is well known and has also been observed previously

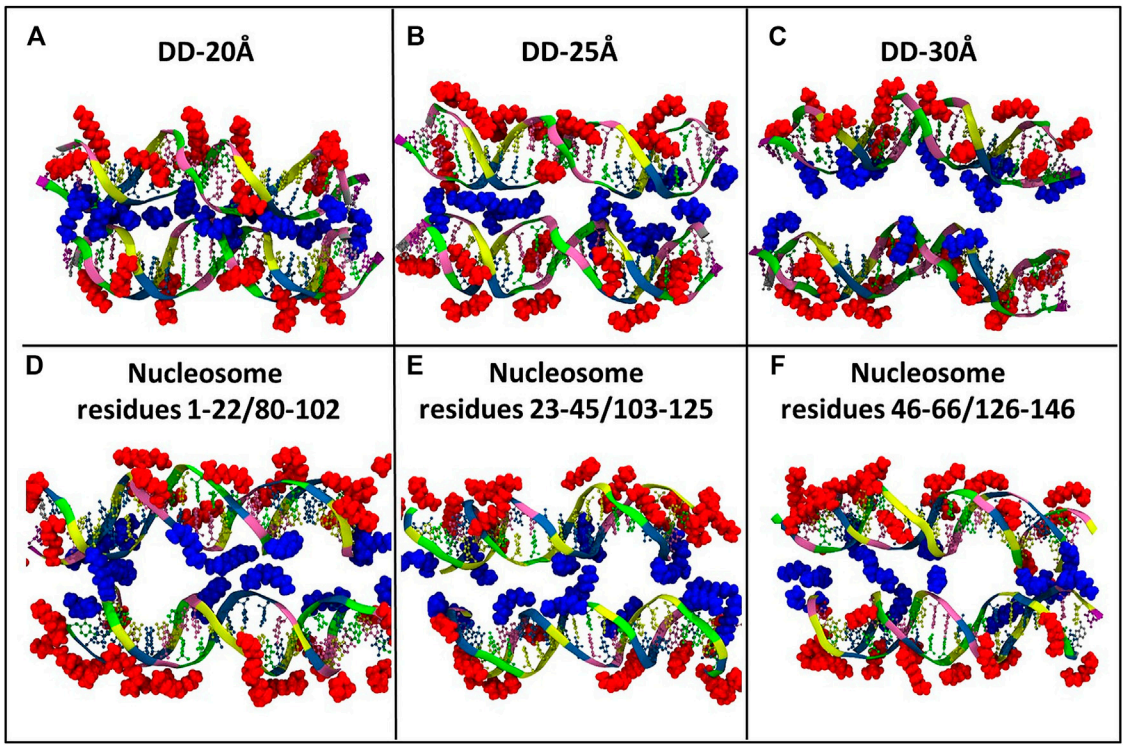


FIGURE 6 | Representative snapshots of the distribution of Spd^{3+} around the two DNA molecules of each system **(A)** DD-20Å, **(B)** DD-25Å, **(C)** DD-30Å, and around the nucleosomal DNA **(D–F)**. To facilitate the visual comparison, the section of the nucleosome system that contains the DNA1 and DNA2 segments was split in three parts with lengths similar to the DD systems. Water and ions are omitted for clarity. Spd^{3+} molecules are colored blue if located in the space between the two DNA molecules (caged), and red if located elsewhere (uncaged).

for systems with PAs (Perepelytsya et al., 2019; Mocci et al., 2021). The dependence of the major groove width on the nucleotide sequence is not as well defined as that of the minor groove.

In the case of the nucleosomal DNA, the minor groove width displays a regular variation along the double helix (**Figure 5A**). The local minima appear with a periodicity of about 10 bps, and visual inspection of a selected portion of the structure and of the corresponding minor groove width (see **Figure 5B**) reveals that the minor groove narrowing is observed in the regions corresponding to a close distance between the parallel portions of the DNA. Interestingly, in these regions the minor grooves of the parallel DNA segments face each other, similar to what was observed for system DD-20Å in the region corresponding to the shortest DNA-DNA distances. It should be noted that since in our simulation the structure of the nucleosomal DNA was constrained to its initial geometry, the width of the groove at each base pair cannot vary during the simulation; however, some fluctuation in the groove widths should necessarily occur in the unrestrained NCP DNA, although probably smaller than in the case of the DD-20Å, DD-25Å and DD-30Å systems (**Figure 4**), due to the constraints imposed in the NCP by the proteins in the core. Therefore, we do not expect great variation in the periodicity of the minor groove width in the real NCP DNA compared to

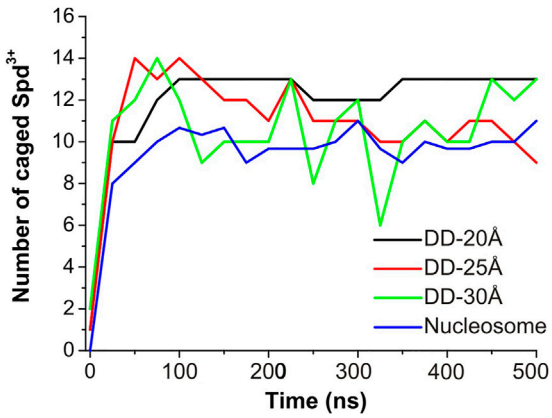


FIGURE 7 | Evolution of the number of Spd^{3+} caged between parallel DNA molecules (or molecular portions). In the case of the nucleosomal system, the total number of Spd^{3+} has been normalized to match the other systems.

what reported in **Figure 5**. The variation of the major groove width along the helix is not as regular as for the minor groove, however some correlation between the minima of the major

FIGURE 8 | Evolution of the movement of each Spd^{3+} molecule with respect to the caged—uncaged position. The Spd^{3+} are numbered from 1 to 30 for the DD systems and 1 to 100 for the nucleosome system, and the blue corresponds to the caged position.

groove and the maxima of the minor groove widths is present (**Figure 5A**).

3.3 DNA-Spd^{3+} Interactions

3.3.1 Distribution and Dynamics in the Interfacial Region Between DNA Molecules

Visual inspection of the simulation trajectories of the systems DD-20Å, DD-25Å and DD-30Å reveals that all the Spd^{3+} molecules interact with the DNA surface, both in the space between the DNA molecules and in the outer regions. In order to determine the influence of the distance between the DNA molecules on the interaction between DNA and Spd^{3+}, we analyzed the simulation snapshots classifying Spd^{3+} molecules as "caged" or "uncaged" as depicted in **Figure 6**. We considered a Spd^{3+} molecule to be caged if the distance from the central N atom of Spd^{3+} to the center of at least one base pair (i.e. the N_1 atom in **Figure 1**) of each DNA molecule is smaller than 15 Å for the DD-20 Å system, 19 Å for the DD-25 Å system, 23 Å for the 30 Å system and 22 Å for the nucleosome system. As it will be shown in the following, the classification of "caged" in our analysis does not necessarily imply that the molecules cannot exit from the region at the interface between the two DNA molecules.

To evaluate the effect of DNA-DNA separation on Spd^{3+} - DNA interactions, we calculated at each saved point of the trajectories the number of caged and uncaged Spd^{3+}; see **Figure 7**. It can be seen that in the DD-20Å system, the number of caged molecules reaches a plateau value of 12–13 after t = 100 ns. In the DD-25Å system we notice a maximum of 14 caged Spd^{3+} in the first 100 ns, followed by a fast decrease that results in a fluctuation between 9–11 molecules in the second part of the simulation (250–500 ns). Notably, the fluctuations are much greater than in DD-20A system, where the number of caged molecules varies by a maximum of one unit after the plateau was reached. It should also be considered that the region between the two DNA fragments increases its volume with increasing distance between the DNA helices, and thus the local concentration of Spd^{3+} molecules between the helices decreases significantly. In the DD-30Å system where the distance between the two DNA molecules is further increased, the number of caged Spd^{3+} fluctuates even more over the duration of the simulation, between a maximum of 14 at t = 75 ns and a minimum of 6 at t = 325 ns. Considering that the volume of the region between the two DNA molecules is larger compared to that in the system DD-25Å, the density of Spd^{3+} continues to decrease, and the instantaneous number of caged molecules is much more

variable than in the other cases. The variation in the number of caged Spd^{3+} in the nucleosome DNA system closely resembles that in the DD-25Å system; since the DNA size in the nucleosome system is larger, the number of Spd^{3+} has been normalized to allow comparison with DD-25Å, by dividing the instantaneous number of Spd^{3+} by 3 (since the DNA length of each of the nucleosome DNA parallel portions is thrice that of the other systems).

Next, the dynamics of Spd^{3+} in the caged state was analyzed to better understand how the distance between DNAs effects this property. To this end, we checked the position of each Spd^{3+} (caged or uncaged), sampling configurations at 25 ns intervals over the entire simulation. The results are presented in **Figure 8**. It can be seen that for the DD-20Å system, once a Spd^{3+} enters the region between the two DNA helices, it effectively remains caged there, i.e. seldom exits from this region. In fact, 11 out of 13 Spd^{3+} stay in the caged position from t = 100 ns (depicted as continuous blue lines in **Figure 8**), while 15 molecules never reach the caged position. In the DD-25Å system, it can be seen that 7 out of a maximum of 14 molecules stay in the caged position starting from t = 100 ns to the end, while the other caged molecules switch from the caged to the uncaged position several times. Also, there are 13 molecules that never enter the caged region. In the case of the DD-30Å, the Spd^{3+} molecules move between caged and uncaged positions throughout the simulation. Moreover, no molecules remain exclusively in the caged or uncaged state as found in the DD-20Å and DD-25Å simulations. When analyzing the dynamics of Spd^{3+} in the nucleosome system, the similarity to the DD-25Å system becomes even more apparent. The Spd^{3+} dynamics observed in DD-25Å are observed also in the nucleosomal DNA system: part of the Spd^{3+} molecules remain in the caged position for the entire simulation (18 out of 100), others are always in the uncaged position (57 out of 100), while the remainder switch between caged and uncaged states multiple times throughout the entire simulation (25 out of 100). If we consider the ratio between the single-state (i.e. always caged or uncaged) and mixed-state Spd^{3+}, a decrease in the ratio of single/mixed-state molecules with increasing distance between the DNA fragments is found: 6.5 for DD-20Å, 2 for DD-25Å and 0 for DD-30Å. Although the nucleosome system has an inter-segment distance of 25-28Å, which places it between the DD-25Å and DD-30Å systems, the ratio of single/mixed-state Spd^{3+} does not follow the same trend, having a value of 3, thus higher than in the DD-25Å system.

On the overall, the Spd^{3+} interactions with DNA are clearly affected by the distance between the DNA molecules, with the increasing confinement between two negatively charged double helices leading to an increase in the concentration of Spd^{3+} in the confined region, and a reduction of the exchange rate of the Spd^{3+} between the caged and uncaged state.

3.3.2 Interaction of Spd^{3+} With the DNA Grooves and Phosphates Groups

To quantify how the overall Spd^{3+} distribution around DNA is affected by the distance between the DNA molecules, the RDFs between Spd^{3+} heavy atoms (i.e., not hydrogen) and DNA were calculated for different regions of the DNA double helices: minor groove (RDF_{MIN}), major groove (RDF_{MAJ}), and phosphate groups (RDF_{PH}). The RDFs averaged over all the nucleotide bases of the DNA molecules are reported and discussed in **Supplementary Figures S8, S9**, and the corresponding CNs calculated from the integration of the RDFs are reported in **Supplementary Table S1**. The CNs decrease with increasing DNA-DNA separation from 20 to 25 Å, and then only marginally upon increasing the separation to 30 Å; the corresponding CNs for the nucleosomal DNA system are significantly lower due to the lower Spd^{3+} concentration, which results from our choice to keep the Spd^{3+}/phosphate group ratio constant for all the systems studied.

The RDFs in **Supplementary Figures S8, S9** provide only general information about the distribution of Spd^{3+} around DNA. In order to analyze the DNA sequence specificity of Spd^{3+} binding, the CN was calculated separately for each base pair by integrating the corresponding RDF up to 4.55 Å. The dependence of coordination numbers on the nucleotide sequence, averaged over the two DNA helices, are shown in **Figure 9**. The coordination numbers for each DNA duplex in the system (DNA1 and DNA2) are shown in **Supplementary Figure S10**. While in previously reported simulations (Perepelytsya et al., 2019) we observed a clear preferential binding of PAs in the minor groove of A-tracts, the sequence specificity of the binding is much less clear in the present simulations (**Figures 9A–C**). This can be due to several factors: on the one hand, the concentration used in this study is higher, making the sequence specificity less dominant; on the other hand, the presence of multiple charged DNA chains close to each other in a crowded environment strongly affects the electrostatic potential and consequently the interactions. Also, it is possible that the reduced mobility of the Spd^{3+} molecule in the caged state implies longer time to obtain a complete sampling. In the major groove the A-tracts appear to be the less favored binding sites, as also previously observed, due to the steric hindrance of the methyl group of thymine (Perepelytsya et al., 2019). In the grooves of the nucleosome DNA (**Supplementary Figure S10**) there are large regions which do not interact with Spd^{3+}; this is probably due to the fact that the nucleosomal DNA is not allowed to move in this simulation, while dynamical structural rearrangements are necessary for the PAs to enter the grooves, or to find proper coordination modes with the partially negative atoms in these regions.

3.3.3 Modes of Interaction of Spd^{3+} With DNA

In addition to the changes in the general mode of interaction of Spd^{3+} with DNA as a function of DNA-DNA distance discussed above, visual analysis of the simulation trajectories reveals changes in the detailed interactions of caged Spd^{3+} molecules. **Figure 10** shows selected simulation snapshots of Spd^{3+} molecules in caged and uncaged position. It can be seen that in both the DD-20Å and DD-25Å systems, the uncaged Spd^{3+} behave in one of two ways: a) they remain in close proximity of the same DNA residue throughout nearly the entire simulation, or, b) they move across the surface of the DNA, exploring a larger surface. In the DD-30Å system, all of the Spd^{3+} have a mixed-state distribution, i.e., no Spd^{3+} remains in the caged or uncaged state for the entire simulation. In the DD-20Å system, the caged

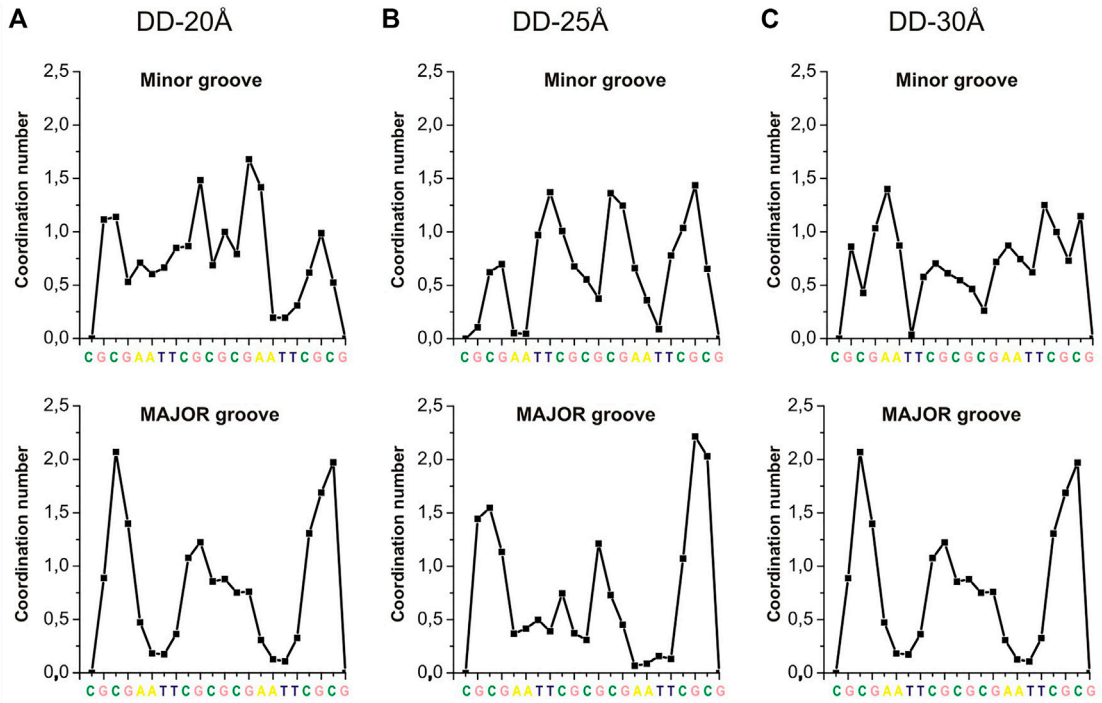


FIGURE 9 | CNs of Spd^{3+} (calculated separately for each DNA base, integrating the corresponding RDFs up to 4.55 Å) in the minor and major grooves of: **(A)** DD-20 Å system; **(B)** DD-25 Å system; **(C)** DD-30Å system.

Spd^{3+} typically remain trapped between the DNA molecules with virtually no movement. In the DD-25Å system, we find that although there are some Spd^{3+} that remain locked in position, other Spd^{3+} molecules can move from one DNA molecule to the other and back, with an intermediate "bridge like" structure (see discussion below). In the DD-30Å system, all Spd^{3+} are moving either along the same DNA molecule or from one DNA molecule to the other forming the same "bridge like" structures. In the nucleosome system, the Spd^{3+} behave in the same way as the DD-25Å system, with some molecules remaining caged in the same position throughout the entire simulation, and other molecules moving along one of the DNA fragments, or from one DNA fragment to the other, as exemplified by the Spd^{3+} on the right in **Figure 10D**.

As stated above, in the system DD-20Å the phosphate groups from different DNA molecules form amino group-mediated contact (O-HNH-O). Visual inspection of the trajectories revealed that several modes of binding of Spd^{3+} to DNA facilitate the close contact between the two DNA helices. **Figure 11** depicts various representative interaction modes. In **Figure 11A**, it can be seen that at the points of contact between the DNAs in system DD-20Å, several Spd^{3+} molecules adopt a "parallel-perpendicular" (pp) orientation, in which the Spd^{3+} are parallel to the backbone of one DNA molecule and perpendicular to the backbone of the second DNA molecule. A detail representation of a Spd^{3+} molecule in the pp orientation simultaneously interacting with 3 DNA strand and 4 phosphate groups, is shown in **Figure 11C**. It can be seen how this conformation enables the close contacts between the DNA molecules, responsible for the first maximum of the RDF_{DD} and RDF_{PH} discussed above. **Figure 11B**, depicts another conformation in which Spd^{3+} adopts a C-shape, forming hydrogen bonds with the oxygens of two phosphate groups from different DNA fragments. The amino groups coordinating a given phosphate group often belong to different Spd^{3+} molecules, forming a complex network, as shown in **Figure 11D**. Here is depicted a contact point between the two DNA molecules and 3 Spd^{3+}. It can be seen that the 3 Spd^{3+} molecules interact, each through multiple contacts, with 7 PO_3 groups. The high local concentration of both positive and negative charged groups, explains why, in the DD-20Å system, the Spd^{3+} caged between the two DNA molecules remains trapped for the entire duration of the simulation as seen in **Figure 8**.

From **Figure 11** we can also see that the Spd^{3+} molecules can adopt different conformations when interacting with DNA; more precisely, rotation around the 7 torsional bonds of this small molecule leads to a wide range of conformations, ranging from a linear conformation to rather compact folded conformation. In order to estimate the probability of different Spd^{3+} conformational states and to determine if the caging alters the distribution among them, we calculated the end-to-end distance (EE) for each Spd^{3+} molecule in the caged and uncaged regions. The EE calculations were performed on configurations obtained by sampling the trajectories every 25 ns, and the EE distributions are represented as histograms in **Supplementary Figure S15**. In the following, we identify the folded and elongated states as those with EE < 7 Å and EE > 9 Å, respectively. Considering the DD systems, there are only small differences in the EE distributions of the

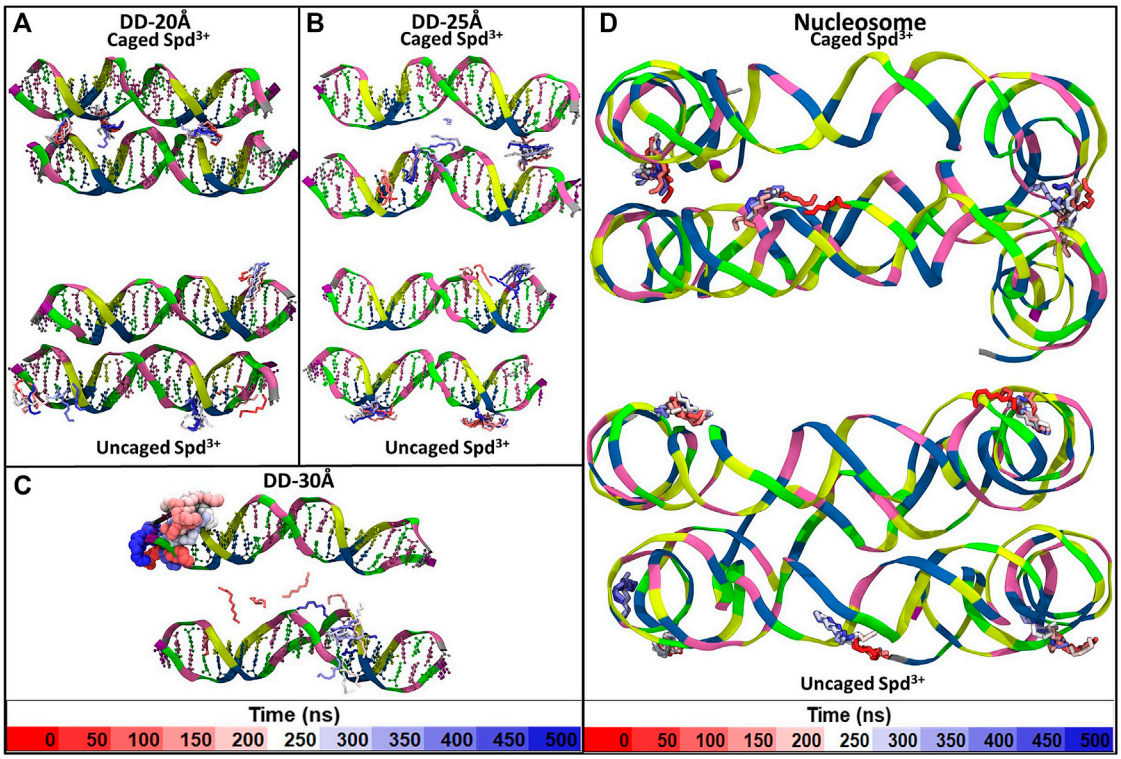


FIGURE 10 | Example of the movement of individual Spd^{3+} over the duration of the simulation of the **(A)** DD-20Å system, **(B)** DD-25Å system, **(C)** DD-30Å system and **(D)** Nucleosome system. For Spd^{3+} we used the licorice representation with H atoms omitted. Each Spd^{3+} changes color from red to white to blue as the simulation time increases, and each change in color corresponds to 50 ns, as shown by the color-coded time scale in the legend. Larger size representations of the Figures are presented in the **Supplementary Figures S2–S5**.

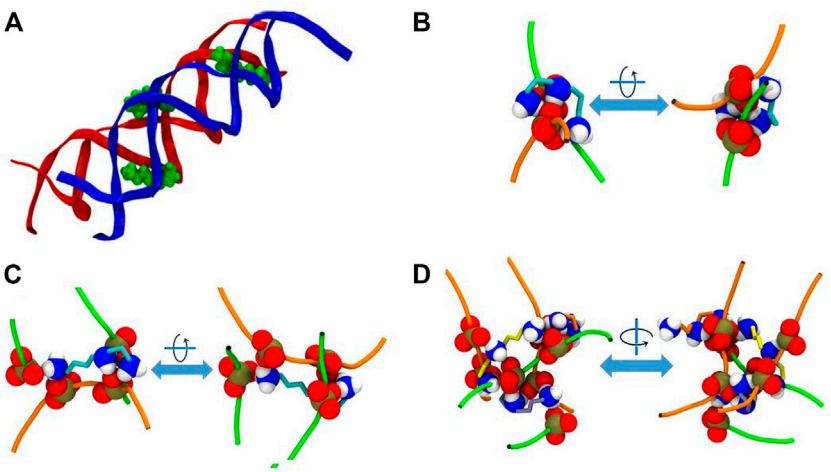


FIGURE 11 | Representative interaction modes of Spd^{3+} with DNA helices. **(A)** Snapshot depicting the parallel-perpendicular (pp) conformation of the Spd^{3+} in the DD-20 Å system. The DNA backbone is represented in red and blue, Spd^{3} green VdW spheres. **(B)** Detail depicting a C-shaped Spd^{3+} that mediates the contact between two phosphate groups, one from each DNA molecules. **(C)** Detail depicting a Spd^{3+} molecule in the pp conformation that mediates the contact between four phosphate groups, two from each DNA molecules. **(D)** Detail depicting a contact point between the two DNA molecules. Three Spd^{3+} mediate this contact point by interacting with 7 PO_3 groups. In **(B–D)** DNA backbones are depicted as an orange and green tube, one color for each molecule, while the NH_3, NH_2 and PO_3 groups are represented as VdW spheres and the C atoms of Spd^{3+} are represented with licorice in teal, orange, and yellow color. Larger representation of the snapshots can be found in the **Supplementary Figures S11–S14**.

uncaged Spd^{3+} among the three systems: 4–8% are in a folded state, and 46–50% in an elongated state. A different picture emerges when analyzing the caged Spd^{3+} molecules: for the DD-25Å and DD-30Å systems, the EE distributions are similar to those of the uncaged Spd^{3+}, with nearly 10% being in a folded state and 44–48% in an elongated state. Conversely, in the case of the DD-20Å system, there is a large increase of the folded state population, which represents 19% of the total, and a decrease to 36% of Spd^{3+} in the elongated state. Therefore, we can conclude that decreasing the distance between DNA molecules induces an increase in the folded states of the caged Spd^{3+} and a decrease in the elongated states, the variations being much larger when the DNA molecules are very close to each other. Considering the nucleosome system, it can be seen that it follows the same trend as the other systems for the uncaged molecules, whereas the caged molecules have a lower probability to be in a folded state (5%) and higher probability to be in an extended state (56%) compared to DD systems. These differences compared to the other systems are most likely due to the constraints applied to the DNA atoms, which prevent the local DNA rearrangement necessary to establish an optimal interaction with the compacted forms of Spd^{3+}; this finding suggests that the constraints employed might significantly affect the distribution among different Spd^{3+} conformational states.

4 CONCLUSION

Four model systems, each containing two DNA double helices with different DNA-DNA separation, in the presence of Spd^{3+} and KCl, have been studied using MD computer simulations, with the aim to understand how the separation between DNA double helices influences the interaction with polyamines.

In all the simulated systems, Spd^{3+} molecules bind in all the different regions of the double helix: minor and major grooves and phosphate groups. The presence of a second DNA double helix influences strongly the interactions with Spd^{3+}. At small DNA-DNA separation (<25 Å between the helix axes), an increase of Spd^{3+} concentration is observed in the region between the parallel DNA helices, compared to concentration when the helices are further apart (30 Å separation between the axes). The separation between the double helices also affects the modes of interactions of the Spd^{3+} molecule with DNA, indicating that some binding modes accessible when DNA molecules are separated (e.g., those involved in the preferential binding to the minor groove of A-tract in diluted aqueous solution) might not be very accessible in highly compact system (e.g., in the cell nucleus) while other binding modes, involving folded Spd^{3+} configurations could be favored in compact DNA aggregates.

In the most condensed form observed in our simulations, the DNA helices adopt a reciprocal orientation with the DNA-DNA contacts mostly occurring between the minor grooves of the parallel helices. This type of orientation is observed also in the experimental structure of nucleosomal DNA. It is useful to note that to reach this type of arrangement in the simulations, the DNA molecules should be free to rotate around their helix axis, and/or to shift along the same.

The dynamics of the Spd^{3+} molecules are also strongly affected by the DNA-DNA separation: at very small separations (20 Å), the Spd^{3+} located between the DNA molecules remain effectively stuck in their binding sites; increasing the inter-helical separation to 25 Å, the PAs still maintain relatively long residence times in the region between the helices, but they move from one binding site to another. Further increasing the separation between the DNA helices to 30 Å leads to a further increase in Spd^{3+} mobility, thus reducing the residence time in the inter-helical space.

The presented data are of relevance for understanding how the interaction of PAs with DNA in compact systems may differ from those in diluted solution, and to understand the mechanisms of compaction of DNA in biological systems. We wish to add that the effect of DNA compaction on the competition between Spd^{3+} and other counterions found in the solvating shell of nucleic acids in biological systems, e.g., K^+ (Auffinger and Westhof, 2000; Auffinger and Hashem, 2007; Mocci and Laaksonen, 2012), is an important related topic that will be analyzed in a future study.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

TV performed the simulations, TV, SP, LV, and FM analysed the simulation, all the authors planned the research, discussed the data, contributed to the writing of the manuscript.

FUNDING

SP acknowledges the support from the National Academy of Sciences of Ukraine (project 0120U100855). AL acknowledges Swedish Research Council for financial support, and partial support from a grant from Ministry of Research and Innovation of Romania (CNCS - UEFISCDI, project number PN-III-P4-ID-PCCF-2016-0050, within PNCDI III). FM and LV acknowledge financial support from Progetto Fondazione di Sardegna (Grant CUP: F72F20000230007) and Regione Autonoma della Sardegna (RASSR81788-2017).

ACKNOWLEDGMENTS

Grants of computing time from the Swedish national infrastructure for computing (SNIC) are gratefully acknowledged. We are grateful to our Reviewers whose comments and suggestions helped to improve and clarify this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2022.836994/full#supplementary-material

REFERENCES

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* 1-2, 19–25. doi:10.1016/j.softx.2015.06.001

Ames, B. N., and Dubin, D. T. (1960). The Role of Polyamines in the Neutralization of Bacteriophage Deoxyribonucleic Acid. *J. Biol. Chem.* 235, 769–775. doi:10.1016/s0021-9258(19)67936-6

Atzori, A., Liggi, S., Laaksonen, A., Porcu, M., Lyubartsev, A. P., Saba, G., et al. (2016). Base Sequence Specificity of Counterion Binding to DNA: What Can MD Simulations Tell Us? *Can. J. Chem.* 94, 1181–1188. doi:10.1139/cjc-2016-0296

Auffinger, P., and Hashem, Y. (2007). Nucleic Acid Solvation: from outside to Insight. *Curr. Opin. Struct. Biol.* 17, 325–333. doi:10.1016/j.sbi.2007.05.008

Auffinger, P., and Westhof, E. (2000). Water and Ion Binding Around RNA and DNA (C,G) oligomers11Edited by I. Tinoco. *J. Mol. Biol.* 300, 1113–1131. doi:10.1006/jmbi.2000.3894

Avogadro (2018). Avogadro: an Open-Source Molecular Builder and Visualization Tool. Version 1.2. Available at: http://avogadro.cc/.

Bignon, E., Chan, C.-H., Morell, C., Monari, A., Ravanat, J.-L., and Dumont, E. (2017). Molecular Dynamics Insights into Polyamine-DNA Binding Modes: Implications for Cross-Link Selectivity. *Chem. Eur. J.* 23, 12845–12852. doi:10.1002/chem.201702065

Blagoi, Y. P., Galkin, V. L., Gladchenko, V. L., Kornilova, S. V., Sorokin, V. A., and Shkorbatov, A. G. (1991). *The Complexes of Nucleic Acids and Metals in the Solutions*. Kyiv: Naukova Dumka.

Bloomfield, V. A. (1996). DNA Condensation. *Curr. Opin. Struct. Biol.* 6, 334–341. doi:10.1016/S0959-440X(96)80052-2

Bloomfield, V. A. (1997). DNA Condensation by Multivalent Cations. *Biopolymers* 44, 269–282. doi:10.1002/(sici)1097-0282(1997)44:3<269::aid-bip6>3.0.co;2-t

Carrivain, P., Cournac, A., Lavelle, C., Lesne, A., Mozziconacci, J., Paillusson, F., et al. (2012). Electrostatics of DNA Compaction in Viruses, Bacteria and Eukaryotes: Functional Insights and Evolutionary Perspective. *Soft Matter* 8, 9285–9301. doi:10.1039/c2sm25789k

Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E. I., Cruzeiro, V. W. D., et al. (2018). *AMBER 2018*. San Francisco: University of California.

Cesare Marincola, F., Virno, A., Randazzo, A., Mocci, F., Saba, G., and Lai, A. (2009). Competitive Binding Exchange between Alkali Metal Ions (K+, Rb+, and Cs+) and Na+ions Bound to the Dimeric Quadruplex [d(G4T4G4)]2: a23Na and1H NMR Study. *Magn. Reson. Chem.* 47, 1036–1042. doi:10.1002/mrc.2509

Chattoraj, D. K., Gosule, L. C., and Schellman, J. A. (1978). DNA Condensation with Polyamines. *J. Mol. Biol.* 121, 327–337. doi:10.1016/0022-2836(78)90367-4

D'Agostino, L. (2018). Native DNA Electronics: Is it a Matter of Nanoscale Assembly? *Nanoscale* 10, 12268–12275. doi:10.1039/c8nr03153c

Dai, L., Mu, Y., Nordenskiöld, L., and Van Der Maarel, J. R. C. (2008). Molecular Dynamics Simulation of Multivalent-Ion Mediated Attraction between DNA Molecules. *Phys. Rev. Lett.* 100, 1–4. doi:10.1103/PhysRevLett.100.118301

Dans, P. D., Walther, J., Gómez, H., and Orozco, M. (2016). Multiscale Simulation of DNA. *Curr. Opin. Struct. Biol.* 37, 29–45. doi:10.1016/j.sbi.2015.11.011

Darden, T., York, D., and Pedersen, L. (1993). Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 98, 10089. doi:10.1063/1.464397

Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., et al. (1981). Structure of a B-DNA Dodecamer: Conformation and Dynamics. *Proc. Natl. Acad. Sci.* 78, 2179–2183. doi:10.1073/pnas.78.4.2179

El Hassan, M. A., and Calladine, C. R. (1998). Two Distinct Modes of Protein-Induced Bending in DNA 1 1 Edited by J. Karn. *J. Mol. Biol.* 282, 331–343. doi:10.1006/jmbi.1998.1994

Estévez-Torres, A., and Baigl, D. (2011). DNA Compaction: Fundamentals and Applications. *Soft Matter* 7, 6746–6756. doi:10.1039/c1sm05373f

Firpo, M. R., and Mounce, B. C. (2020). Diverse Functions of Polyamines in Virus Infection. *Biomolecules* 10, 628. doi:10.3390/biom10040628

Franklin, R. E., and Gosling, R. G. (1953). The Structure of Sodium Thymonucleate Fibres. II. The Cylindrically Symmetrical Patterson Function. *Acta Crystallogr.* 6, 678–685. doi:10.1107/s0365110x53001940

Gosule, L. C., and Schellman, J. A. (1978). DNA Condensation with Polyamines. *J. Mol. Biol.* 121, 311–326. doi:10.1016/0022-2836(78)90366-2

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012). Avogadro: an Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* 4, 17. doi:10.1186/1758-2946-4-17

Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* 18, 1463–1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h

Hoover, W. G. (1985). Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A.* 31, 1695–1697. doi:10.1103/PhysRevA.31.1695

Ivani, I., Dans, P. D., Noy, A., Pérez, A., Faustino, I., Hospital, A., et al. (2016). Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* 13, 55–58. doi:10.1038/nmeth.3658

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79, 926–935. doi:10.1063/1.445869

Joung, I. S., and Cheatham, T. E. (2008). Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* 112, 9020–9041. doi:10.1021/jp8001614

Kabir, A., and Suresh Kumar, G. (2013). Binding of the Biogenic Polyamines to Deoxyribonucleic Acids of Varying Base Composition: Base Specificity and the Associated Energetics of the Interaction. *PLoS One* 8, e70510. doi:10.1371/journal.pone.0070510

Kanemura, A., Yoshikawa, Y., Fukuda, W., Tsumoto, K., Kenmotsu, T., and Yoshikawa, K. (2018). Opposite Effect of Polyamines on *In Vitro* Gene Expression: Enhancement at Low Concentrations but Inhibition at High Concentrations. *PLoS One* 13, e0193595–11. doi:10.1371/journal.pone.0193595

Kitagawa, T., Nishio, T., Yoshikawa, Y., Umezawa, N., Higuchi, T., Shew, C.-Y., et al. (2021). Effects of Structural Isomers of Spermine on the Higher-Order Structure of Dna and Gene Expression. *Int. J. Mol. Sci.* 22, 2355. doi:10.3390/ijms22052355

Kornyshev, A. A., Lee, D. J., Leikin, S., and Wynveen, A. (2007). Structure and Interactions of Biological Helices. *Rev. Mod. Phys.* 79, 943–996. doi:10.1103/RevModPhys.79.943

Kornyshev, A. A., and Leikin, S. (1999). Electrostatic Zipper Motif for DNA Aggregation. *Phys. Rev. Lett.* 82, 4138–4141. doi:10.1103/PhysRevLett.82.4138

Korolev, N., Allahverdi, A., Lyubartsev, A. P., and Nordenskiöld, L. (2012). The Polyelectrolyte Properties of Chromatin. *Soft Matter* 8, 9322–9333. doi:10.1039/c2sm25662b

Korolev, N., Lyubartsev, A. P., Laaksonen, A., and Nordenskiöld, L. (2004). A Molecular Dynamics Simulation Study of Polyamine? and Sodium?DNA. Interplay between Polyamine Binding and DNA Structure. *Eur. Biophys. J.* 33, 671–682. doi:10.1007/s00249-004-0410-7

Korolev, N., Lyubartsev, A. P., Nordenskiöld, L., and Laaksonen, A. (2001). Spermine: an "invisible" Component in the Crystals of B-DNA. A Grand Canonical Monte Carlo and Molecular Dynamics Simulation Study. *J. Mol. Biol.* 308, 907–917. doi:10.1006/jmbi.2001.4642

Kumar, R., and Grubmüller, H. (2015). Do_x3dna: a Tool to Analyze Structural Fluctuations of dsDNA or dsRNA from Molecular Dynamics Simulations: Fig. 1. *Bioinformatics* 31, 2583–2585. doi:10.1093/bioinformatics/btv190

Lavery, R., Maddocks, J. H., Pasi, M., and Zakrzewska, K. (2014). Analyzing Ion Distributions Around DNA. *Nucleic Acids Res.* 42, 8138–8149. doi:10.1093/nar/gku504

Levine, B. G., Stone, J. E., and Kohlmeyer, A. (2011). Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units-Radial Distribution Function Histogramming. *J. Comput. Phys.* 230, 3556–3569. doi:10.1016/j.jcp.2011.01.048

Lu, X.-J., and Olson, W. K. (2003). 3DNA: A Software Package for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic Acid Structures. *Nucleic Acids Res.* 31, 5108–5121. doi:10.1093/nar/gkg680

Maleev, V. I., Semenov, M. A., Gasan, A. I., and Kashpur, V. A. (1993). Physical Properties of the DNA-Water System. *Biofizika* 38, 768–790.

McConnell, K. J., and Beveridge, D. L. (2000). DNA Structure: What's in Charge? *J. Mol. Biol.* 304, 803. doi:10.1006/jmbi.2000.4167

Mocci, F., Laaksonen, A., Engelbrecht, L., Vasiliu, T., and Perepelytsya, S. (2021). "DNA-polyamine Interactions: Insight from Molecular Dynamics Simulations on the Sequence-specific Binding of Spermidine3+," in *Soft Matter Systems for Biomedical Applications*.

Mocci, F., and Laaksonen, A. (2012). Insight into Nucleic Acid Counterion Interactions from inside Molecular Dynamics Simulations Is "worth its Salt". *Soft Matter* 8, 9268–9284. doi:10.1039/c2sm25690h

Mocci, F., Laaksonen, A., Lyubartsev, A., and Saba, G. (2004). Molecular Dynamics Investigation of 23Na NMR Relaxation in Oligomeric DNA Aqueous Solution. *J. Phys. Chem. B* 108, 16295–16302. doi:10.1021/jp047744+

Mocci, F., and Saba, G. (2003). Molecular Dynamics Simulations of A · T-Rich Oligomers: Sequence-specific Binding of Na in the Minor Groove of B-DNA. *Biopolymers* 68, 471–485. doi:10.1002/bip.10334

Mounce, B. C., Olsen, M. E., Vignuzzi, M., and Connor, J. H. (2017). Polyamines and Their Role in Virus Infection. *Microbiol. Mol. Biol. Rev.* 81, 1–12. doi:10.1128/mmbr.00029-17

Nosé, S. (1984). A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* 52, 255–268. doi:10.1080/00268978400101201

Parrinello, M., and Rahman, A. (1981). Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* 52, 7182–7190. doi:10.1063/1.328693

Pasi, M., Maddocks, J. H., and Lavery, R. (2015). Analyzing Ion Distributions Around DNA: Sequence-Dependence of Potassium Ion Distributions from Microsecond Molecular Dynamics. *Nucleic Acids Res.* 43, 2412–2423. doi:10.1093/nar/gkv080

Patel, M. M., and Anchordoquy, T. J. (2006). Ability of Spermine to Differentiate between DNA Sequences-Preferential Stabilization of A-Tracts. *Biophysical Chem.* 122, 5–15. doi:10.1016/j.bpc.2006.02.001

Perepelytsya, S. (2018). Hydration of Counterions Interacting with DNA Double helix: a Molecular Dynamics Study. *J. Mol. Model.* 24, 171. doi:10.1007/s00894-018-3704-x

Perepelytsya, S., Uličný, J., Laaksonen, A., and Mocci, F. (2019). Pattern Preferences of DNA Nucleotide Motifs by Polyamines Putrescine2+, Spermidine3+ and Spermine4+. *Nucleic Acids Res.* 47, 6084–6097. doi:10.1093/nar/gkz434

Ponomarev, S. Y., Thayer, K. M., and Beveridge, D. L. (2004). Ion Motions in Molecular Dynamics Simulations on DNA. *Proc. Natl. Acad. Sci.* 101, 14771–14775. doi:10.1073/pnas.0406435101

Post, C. B., and Zimm, B. H. (1979). Internal Condensation of a Single DNA Molecule. *Biopolymers* 18, 1487–1501. doi:10.1002/bip.1979.360180612

Roos, W. H., Ivanovska, I. L., Evilevitch, A., and Wuite, G. J. L. (2007). Viral Capsids: Mechanical Characteristics, Genome Packaging and Delivery Mechanisms. *Cell. Mol. Life Sci.* 64, 1484–1497. doi:10.1007/s00018-007-6451-1

Rouzina, I., and Bloomfield, V. A. (1996). Macroion Attraction Due to Electrostatic Correlation between Screening Counterions. 1. Mobile Surface-Adsorbed Ions and Diffuse Ion Cloud. *J. Phys. Chem.* 100, 9977–9989. doi:10.1021/jp960458g

Saenger, W. (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag. Available at: http://books.google.com/books?id=opBqAAAAMAAJ.

Schlick, T. (2002). *Molecular Modeling and Simulation: An Interdisciplinary Guide*. New York: Springer US. Available at: http://link.springer.com/10.1007/978-0-387-22464-0.

Tsunaka, Y., Kajimura, N., Tate, S., and Morikawa, K. (2005). Alteration of the Nucleosomal DNA Path in the crystal Structure of a Human Nucleosome Core Particle. *Nucleic Acids Res.* 33, 3424–3434. doi:10.1093/nar/gki663

van Dam, L., Korolev, N., and Nordenskiöld, L. (2002). Polyamine-nucleic Acid Interactions and the Effects on Structure in Oriented DNA Fibers. *Nucleic Acids Res.* 30, 419–428. doi:10.1093/nar/30.2.419

Várnai, P., and Zakrzewska, K. (2004). DNA and its Counterions: A Molecular Dynamics Study. *Nucleic Acids Res.* 32, 4269–4280. doi:10.1093/nar/gkh765

Vasiliu, T., Craciun, B. F., Neamtu, A., Clima, L., Isac, D. L., Maier, S. S., et al. (2021). In Silico study of PEI-PEG-Squalene-dsDNA Polyplex Formation: The Delicate Role of the PEG Length in the Binding of PEI to DNA. *Biomater. Sci.* 9, 6623–6640. doi:10.1039/d1bm00973g

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/jcc.20035

Watson, J. D., and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 171, 737–738. doi:10.1038/171737a0

William, H., Andrew, D., and Klaus, S. (1996). Visual Molecular Dynamics. *J. Mol. Graphics* 14, 33–38. doi:10.1016/0263-7855(96)00018-5

Yoo, J., and Aksimentiev, A. (2016). The Structure and Intermolecular Forces of DNA Condensates. *Nucleic Acids Res.* 44, 2036–2046. doi:10.1093/nar/gkw081

Yoo, J., Kim, H., Aksimentiev, A., and Ha, T. (2016). Direct Evidence for Sequence-dependent Attraction between Double-Stranded DNA Controlled by Methylation. *Nat. Commun.* 7, 1–7. doi:10.1038/ncomms11045

Young, M. A., Jayaram, B., and Beveridge, D. L. (1997). Intrusion of Counterions into the Spine of Hydration in the Minor Groove of B-DNA: Fractional Occupancy of Electronegative Pockets. *J. Am. Chem. Soc.* 119, 59. doi:10.1021/ja960459m

Zdorevskyi, O. O., and Perepelytsya, S. M. (2020). Dynamics of K+ Counterions Around DNA Double helix in the External Electric Field: A Molecular Dynamics Study. *Eur. Phys. J. E* 43, 77. doi:10.1140/epje/i2020-12000-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 28 February 2022
doi: 10.3389/fchem.2022.823246

Environment Effects on X-Ray Absorption Spectra With Quantum Embedded Real-Time Time-Dependent Density Functional Theory Approaches

*Matteo De Santis, Valérie Vallet and André Severo Pereira Gomes**

CNRS, UMR 8523–PhLAM–Physique des Lasers, Atomes et Molécules, University Lille, Lille, France

OPEN ACCESS

Edited by:
Heribert Reis, National Hellenic Research Foundation, Greece

Reviewed by:
Aurélien De La Lande, UMR8000 Laboratoire de Chimie Physique (LCP), France
Yi Yao, Duke University, United States

***Correspondence:**
André Severo Pereira Gomes
andre.gomes@univ-lille.fr

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *27 November 2021*
Accepted: *04 February 2022*
Published: *28 February 2022*

Citation:
De Santis M, Vallet V and Gomes ASP (2022) Environment Effects on X-Ray Absorption Spectra With Quantum Embedded Real-Time Time-Dependent Density Functional Theory Approaches. Front. Chem. 10:823246. doi: 10.3389/fchem.2022.823246

In this work we implement the real-time time-dependent block-orthogonalized Manby-Miller embedding (rt-BOMME) approach alongside our previously developed real-time frozen density embedding time-dependent density functional theory (rt-TDDFT-in-DFT FDE) code, and investigate these methods' performance in reproducing X-ray absorption spectra (XAS) obtained with standard rt-TDDFT simulations, for model systems comprised of solvated fluoride and chloride ions ($[X@(H_2O)_8]^-$, X = F, Cl). We observe that for ground-state quantities such as core orbital energies, the BOMME approach shows significantly better agreement with supermolecular results than FDE for the strongly interacting fluoride system, while for chloride the two embedding approaches show more similar results. For the excited states, we see that while FDE (constrained not to have the environment densities relaxed in the ground state) is in good agreement with the reference calculations for the region around the K and L_1 edges, and is capable of reproducing the splitting of the $1s^1(n+1)p^1$ final states ($n+1$ being the lowest virtual p orbital of the halides), it by and large fails to properly reproduce the $1s^1(n+2)p^1$ states and misses the electronic states arising from excitation to orbitals with important contributions from the solvent. The BOMME results, on the other hand, provide a faithful qualitative representation of the spectra in all energy regions considered, though its intrinsic approximation of employing a lower-accuracy exchange-correlation functional for the environment induces non-negligible shifts in peak positions for the excitations from the halide to the environment. Our results thus confirm that QM/QM embedding approaches are viable alternatives to standard real-time simulations of X-ray absorption spectra of species in complex or confined environments.

Keywords: real-time propagation, time-dependent density functional theory, frozen density embedding, block-orthogonalized manby-miller embedding, X-ray absorption spectroscopy, halides

1 INTRODUCTION

X-ray absorption spectroscopy (XAS) is a powerful technique to probe the structural and electronic properties of molecules from an atomistic picture, since the absorbing photons in the X-ray energy range promote excitations of the core electrons to unoccupied or continuum states. The resulting absorption peaks are called edges in XAS and are labelled according to the origin of the core state, for instance K edge for 1s, L_1 for 2s, L_2 edge for $2p_{1/2}$ and L_3 edge for the $2p_{3/2}$. The spectra features near these edges are called X-ray absorption near-edge structures (XANES). Both the energy range and the spectral shapes directly provide information on the oxidation state, local symmetry, and coordination environment of a selected analyte in the gas, liquid, or solid phase (de Groot, 2001; Bunker, 2009; Bokhoven and Lamberti, 2016; Zimmermann et al., 2020). For instance, K edges correspond to $1s \rightarrow (n + 1)p$ dipole transitions, $n + 1$, being the first virtual p level, implying that, in a simple picture, the edge position is a direct measure of molecular valence states, thus allowing us to monitor the effect of the local environment on a given atom within an analyte. The interpretation of such environmental interplay calls for electronic structure calculations that allow us to access the atomic and molecular energy levels. More specifically, the theoretical modeling of XAS spectra implies the calculations of core-valence excitations.

Within quantum chemical methods, density functional theory (DFT)-based approaches such as time-dependent density functional theory (TDDFT) in its linear response (LR) (frequency domain) formulation, currently offer the best compromise between cost and accuracy for calculating electronic excitations (Norman and Dreuw, 2018; Besley, 2021). While a brute-force application of TDDFT to XAS would be prohibitively expensive as a large number of states (valence excitations, resonance, etc.) need to be determined before arriving at the energy regions pertaining to the core excitations, the introduction of restricted-excitation window TDDFT (in which one can restrict the calculation to access only configurations in which particular core electrons are excited (Stener et al., 2003; Besley and Asmuruf, 2010; Zhang et al., 2012)) or the complex polarization propagator approach (Ekström et al., 2006; Ekström and Norman, 2006; Jiemchooroj et al., 2007; Villaume et al., 2010; Pedersen et al., 2014; Fahleson et al., 2016; Rinkevicius et al., 2016) (from which one can obtain the spectral profiles for a given range of frequencies of the external perturbations from the imaginary part of the dipole polarizability), has allowed TDDFT to be routinely used in simulating XAS.

An alternative to the frequency domain approaches above that is gaining attention in recent years is that of the real-time formulation of TDDFT (rt-TDDFT) (Goings et al., 2017; Li et al., 2020), in which time-dependent properties (such as electronically excited states) are obtained based on integrating the time-dependent Kohn-Sham (TDKS) equations in time. While the theoretical underpinnings, strengths, and limitations in respect to accuracy are similar to traditional linear response (LR) TDDFT methods for obtaining electronic spectra, rt-TDDFT provides fully time-resolved solutions that can potentially incorporate non-linear effects, and also allows for strong external perturbations. With that, rt-TDDFT is used to compute not only spectroscopic properties including XAS (Lopata et al., 2012; Kadek et al., 2015; Li et al., 2020) but also the time and space-resolved electronic response to arbitrary external stimuli (e.g., electron charge dynamics after laser excitation) (Eberly et al., 1991; Cheng et al., 2006; Keldysh, 2017; Mokkath, 2020).

However, as soon as one wishes to treat molecules surrounded by an environment (e.g., species in solution or in otherwise confined spaces), the structural model for the system of interest might become too large to be treated with plain DFT approaches. In such cases, subsystem or embedding approaches (Gomes and Jacob, 2012) appear as a computationally efficient strategy: they allow for combinations of different levels of theory for the subsystems of interest and their surroundings, thereby reducing the overall computational cost. Furthermore, embedding approaches allow for selectively switching on or off interactions between different subsystems, and thus can offer a powerful way to understand the physics of chemistry of a particular process, when the analysis of a full (supermolecular) calculation may prove much more cumbersome to analyze. This is the case when analyzing electronically excited states of confined systems, which can involve transitions within particular subsystems as well as between the different subsystems–and in core states in particular since core states can be embedded into states representing the continuum.

There have been several propositions to couple rt-TDDFT methods to embedding approaches, perhaps the most widely used ones involving the coupling between a quantum subsystem and a classical environment (Lipparini and Mennucci, 2021), described by continuum models (Pipolo and Corni, 2016; Gil et al., 2019) or classical force fields (QM/MM) (Marques et al., 2003; Morzan et al., 2014; Wu et al., 2017; Parise et al., 2018). Although the obvious advantages are in cost reduction, these approaches may not properly describe specific interactions such as hydrogen bonds (for the continuum models) or rely on the availability of an appropriate classical force field (for QM/MM). Classical approaches in any case will be limited in their ability to properly describe phenomena in which a quantum description of the environment is important (such as charge delocalization, coupled excitations, or excitations across many parts of the systems not confined to a small fragment). The alternative in this case is to use quantum embedding theories (QM/QM) (Gomes and Jacob, 2012; Jacob and Neugebauer, 2014; Wesolowski et al., 2015; Sun and Chan, 2016; Goez et al., 2018), and among the fully quantum mechanical approaches to include environmental effects in the molecular response property, we note the family of subsystem DFT approaches (Jacob and Neugebauer, 2014; Wesolowski et al., 2015), to which the frozen density embedding (FDE) scheme is a member. It corresponds to a partitioning of a given system into a set of subsystems that can be, for instance, all represented within the Kohn-Sham framework, which interact through a local embedding potential. A subsystem DFT formulation of the real-time methodology (rt-TDDFT-in-DFT)

has been presented in a seminal work by Pavanello *et al.* together with its formulation within the FDE framework (Krishtal et al., 2015; Krishtal and Pavanello, 2016; Kumar et al., 2017; Genova et al., 2017).

This initial formulation, based upon plane-wave basis representations for the different subsystems, has been shown to properly capture the coupling in the response of the different subsystems, through the dependency of the time-dependent embedding potential on the time-dependent electron densities of all (or a subset of) subsystems, whenever such a coupling is of importance. It should be noted that such couplings between the response of subsystems to external perturbation can also be taken into account in a frequency-domain formulation, but at the expense of determining second (or higher) order derivatives of the interaction energy (Neugebauer, 2007; Neugebauer, 2009a; Neugebauer, 2009b; Höfener et al., 2012; König and Neugebauer, 2013; Pavanello, 2013). That said, applications of linear-response or real-time TDDFT-in-DFT showed that in many cases, the coupling between the response of different subsystems can be ignored and a so-called "uncoupled" TDDFT-in-DFT approach can yield accurate results (Gomes et al., 2008; Gomes et al., 2013; Olejniczak et al., 2017; De Santis et al., 2020a), provided the coupling between subsystems in the ground state is well described by the embedding potential representing the subsystems' interaction. FDE-based calculation has been shown to perform well for situations in which there are no strong interactions between subsystems, such as covalent bonds. This makes it possible in general to describe interactions such as hydrogen bonds, though in certain cases the approximations intrinsic to FDE, due to the use of approximate kinetic energy density functionals (KEDFs) (Beyhan et al., 2010; Grimmel et al., 2019) in the description of the embedding potential, prevent it from accurately describing stronger non-covalent interactions (Fux et al., 2010; Bouchafra et al., 2018a). While in such cases, a pragmatic solution is to enlarge the active subsystem, that can be potentially problematic in respect to the increase of computational costs, especially if one is interested in replacing DFT by higher-level approaches such as coupled clusters to describe the subsystem of interest.

Another QM/QM family of embedding approaches closely connected to the subsystem DFT approaches mentioned above involves the use of projection operator techniques (Goodpaster et al., 2010; Goodpaster et al., 2011; Manby et al., 2012; Ding et al., 2017; Tölle et al., 2019a; Tölle et al., 2019b; Lee et al., 2019; Graham et al., 2020; Niemeyer et al., 2020; Scholz et al., 2020), and by foregoing the use of the approximate KEDFs, shows better performance in describing strong interactions. These approaches, in some variants also referred to as Manby-Miller embedding (MME), have been shown to be particularly adept at allowing the fragmentation of a particular system through covalent bonds. More recently, block-orthogonalized MME (BOMME) (Ding et al., 2017) has been introduced to alleviate issues that plagued prior MME variants (Manby et al., 2012). BOMME allows one to treat the target system with a high-level Fock matrix and the remaining degrees of freedom with a less expensive Fock matrix by reducing the quality of the basis set and exchange. A combination of BOMME with rt-TDDFT has been recently proposed by Koh *et al.* (Koh et al., 2017). They demonstrated that rt-BOMME can capture both intermolecular and intramolecular couplings and their induced effects, namely solvent shifts, on spectra of chromophores.

However, in that work only processes involving valence electrons have been considered. Given the interest of core spectra as a means to characterize species in complex environments, it is of great interest to explore the behavior of rt-BOMME for describing XAS. We note that the same is also the case for FDE or TDDFT-in-DFT since these have also, to the best of our knowledge, not yet been explored for core excitations.

Thus, the main goal of this work is to describe the first investigation of the performance of rt-TDDFT-in-DFT and rt-BOMME for core excitations. To do so, we have extended our recently developed Psi4Numpy-based rt-TDDFT-in-DFT to implement the BOMME and rt-BOMME methods. As discussed below, in this manuscript we shall focus on the K and L_1 edges of hydrated halides, since halogenated species and their interaction with species in solution or at interfaces are of particular interest in atmospheric sciences (Finlayson-Pitts, 2013; Pillar-Little et al., 2013; Simpson et al., 2015; Kong et al., 2017; Finlayson-Pitts, 2019; Bartels-Rausch et al., 2021; Yu and Li, 2021). Here, however, in order to simplify our discussion we have considered relatively simple model systems representing the first hydration shell of the halides ($[F(H_2O)_8]^-$ and $[Cl(H_2O)_8]^-$) that nevertheless can gauge the ability of the different embedding methods to describe interactions of varying strengths between halides and their environment. Also, due to the scarce experimental data for XAS on such systems, we shall restrict ourselves to comparisons of two limiting cases: the free ions and the rt-TDDFT calculations on the supermolecular system (which then serve as our benchmarks).

2 MATERIALS AND METHODS

2.1 Theoretical Background

The Frozen Density (FDE) (Wesolowski and Warshel, 1993; Gomes and Jacob, 2012; Jacob and Neugebauer, 2014; Wesolowski et al., 2015) and Block-Orthogonalized Manby-Miller embedding (BOMME) approaches (Ding et al., 2017), and their extension to the rt-TDDFT framework have been described in previous works (Koh et al., 2017; De Santis et al., 2020a). In this section, after brief recapitulation of the rt-TDDFT method, we will outline analogies and differences of rt-FDE and rt-BOMME approaches.

In rt-TDDFT, the one-electron density matrix $\boldsymbol{D}(t)$ representing, in the algebraic approximation, the time-dependent electron density evolves in time according to

$$\boldsymbol{D}(t) = \boldsymbol{U}(t, t_0)\, \boldsymbol{D}(t_0)\, \boldsymbol{U}(t, t_0)^{\dagger}, \tag{1}$$

where $\boldsymbol{U}(t, t_0)$ is the matrix representation of the time-evolution operator:

$$\boldsymbol{U}(t, t_0) = \hat{T} \exp\left(-i \int_{t_0}^{t} \boldsymbol{F}(t')dt' \right). \tag{2}$$

The real-time approach is based on the repeated application of **Eq. 1** over a discretized time-domain. Time discretization allows us to devise advantageous representations of $\boldsymbol{U}$ to be employed in real computer codes. In this work, we employ the exponential midpoint ansatz, which has been successfully employed in the study of valence and core excitations (Lopata and Govind, 2011; Lopata et al., 2012) Extensive discussion on the computational strategies employed to carry out the time-evolution propagation can be found in the seminal work by Castro and co-workers (Castro et al., 2004). In rt-TDDFT, the Fock matrix is defined as

$$\boldsymbol{F}(t) = \boldsymbol{h}_0 + \boldsymbol{G}[\boldsymbol{D}(t)] + v_{\text{ext}}(t), \tag{3}$$

where $\boldsymbol{h}_0$ represents the one-electron operator while $\boldsymbol{G}$ is the two-electron term

$$\boldsymbol{G}[\boldsymbol{D}(t)] = \boldsymbol{J}[\boldsymbol{D}(t)] + c_x\boldsymbol{K}[\boldsymbol{D}(t)] + c_x\boldsymbol{V}_{\text{xc}}[\boldsymbol{D}(t)], \tag{4}$$

c_x being the fraction of the Hartree-Fock exchange in the exchange correlation potential V_{xc}. It is worth noting the Fock matrix appearing in **Eq. 3** has an implicit time-dependence due to a time-dependent density matrix, and the explicit time-dependence due to the external potential $v_{\text{ext}}(t)$.

An embedding mean-field approach is based on the mapping of two different domains within the total system, into two different-quality levels of theory to be employed in each domain. This can be realized by assigning a high-level Fock matrix to the subsystem to be treated accurately, while letting the remaining part be described by a low-level Fock matrix in a reduced basis set. The block-orthogonalized (BO) partitioning scheme proposed by Ding and co-workers (Ding et al., 2017) relies on a projected basis in place of the conventional atomic-orbital (AO) partitioning to define the high- and low-level components of the system. Such a scheme proved to be suitable to remove the artifacts related to the embedding scheme while keeping the expression of the low-level Fock as simple as:

$$\tilde{\boldsymbol{h}}_0 = \boldsymbol{O}^T\boldsymbol{h}_0\boldsymbol{O}, \quad \tilde{\boldsymbol{G}}^{\text{Low}}[\tilde{\boldsymbol{D}}] = \boldsymbol{O}^T\boldsymbol{G}^{\text{Low}}\boldsymbol{O}, \quad \text{where } \tilde{\boldsymbol{D}} = \boldsymbol{O}\boldsymbol{D}\boldsymbol{O}^T. \tag{5}$$

In **Eq. 5**, quantities expressed in the block-orthogonalized basis are denoted by tildes, and $\boldsymbol{O}$ is the transformation matrix from the non-orthogonal AO basis set to the BO basis set:

$$\boldsymbol{O} = \begin{pmatrix} \boldsymbol{I}^{\text{AA}} & -\boldsymbol{P}^{\text{AB}} \\ \boldsymbol{0} & \boldsymbol{I}^{\text{BB}} \end{pmatrix}. \tag{6}$$

The sub-blocks appearing in the transformation matrix are the identity matrices $\boldsymbol{I}^{\text{AA}}$ and $\boldsymbol{I}^{\text{BB}}$ having dimensions of n_a and n_b, mapping subsystem A and B basis sets respectively, and the projection matrix $\boldsymbol{P}^{AB} = (\boldsymbol{S}^{\text{AA}})^{-1}\boldsymbol{S}^{\text{AB}}$, in which $\boldsymbol{S}^{AB}$ is the AO overlap between the subsystems. Here and hereafter the AA (BB) block denote the subsystem with high- (low-)level theory. The Fock matrix in the BO basis reads as:

$$\boldsymbol{F} = \tilde{\boldsymbol{h}}_0 + \tilde{\boldsymbol{G}}^{\text{Low}}[\tilde{\boldsymbol{D}}] + (\tilde{\boldsymbol{G}}^{\text{High}}[\tilde{\boldsymbol{D}}^{\text{AA}}] - \tilde{\boldsymbol{G}}^{\text{Low}}[\tilde{\boldsymbol{D}}^{\text{AA}}]). \tag{7}$$

In this context different schemes for the calculation of the exchange term (in $\boldsymbol{G}^{\text{High}}$) are available. Following Koh et al. (2017), we adopted the simplest scheme for $E_{\text{EX}}[\tilde{\boldsymbol{D}}^{\text{AA}}]$ which takes into account only the exact exchange interaction within the AA block:

$$E_{\text{EX0}} = -\frac{1}{4}\sum_{\mu\kappa\nu\lambda}(\mu\kappa|\nu\lambda)\boldsymbol{D}^{\text{AA}}_{\mu\nu}\boldsymbol{D}^{\text{AA}}_{\kappa\lambda}. \tag{8}$$

In the Frozen Density formulation of DFT the entire system is partitioned into N subsystems, and the total density $\rho_{\text{tot}}(\boldsymbol{r})$ is represented as the sum of electron densities of the various subsystems [i.e., $\rho_a(\boldsymbol{r})$ (a = 1, .., N)]. In this work we restrict our consideration to a simplified model in which the total density is partitioned in only two contributions as

$$\rho_{\text{tot}}(\boldsymbol{r}) = \rho_{\text{I}}(\boldsymbol{r}) + \rho_{\text{II}}(\boldsymbol{r}). \tag{9}$$

The total energy of the system can then be written as

$$E_{\text{tot}}[\rho_{\text{I}}, \rho_{\text{II}}] = E_{\text{I}}[\rho_{\text{I}}] + E_{\text{II}}[\rho_{\text{II}}] + E_{\text{int}}[\rho_{\text{I}}, \rho_{\text{II}}], \tag{10}$$

with the energy of each subsystem (E_i [ρ_i], with i = I, II) given according to the usual definition in DFT as

$$E_i[\rho_i] = \int \rho_i(\boldsymbol{r})v^i_{\text{nuc}}(\boldsymbol{r})\mathrm{d}^3r + \frac{1}{2}\iint\frac{\rho_i(\boldsymbol{r})\rho_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3r\mathrm{d}^3r' + \\ +E_{\text{xc}}[\rho_i] + T_s[\rho_i] + E^i_{\text{nuc}}. \tag{11}$$

In the above expression, $v^i_{\text{nuc}}(\boldsymbol{r})$ is the nuclear potential due to the set of atoms which defines the subsystem and E^i_{nuc} is the related nuclear repulsion energy. $T_s[\rho_i]$ is the kinetic energy of the auxiliary non-interacting system, which is, within the Kohn-Sham (KS) approach, commonly evaluated using the KS orbitals. The interaction energy is given by the expression:

$$E_{\text{int}}[\rho_{\text{I}}, \rho_{\text{II}}] = \int \rho_{\text{I}}(\boldsymbol{r})v^{\text{II}}_{\text{nuc}}(\boldsymbol{r})\mathrm{d}^3r + \int \rho_{\text{II}}(\boldsymbol{r})v^{\text{I}}_{\text{nuc}}(\boldsymbol{r})\mathrm{d}^3r + E^{\text{I,II}}_{\text{nuc}} \\ + \iint\frac{\rho_{\text{I}}(\boldsymbol{r})\rho_{\text{II}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3r\mathrm{d}^3r' + E^{\text{nadd}}_{\text{xc}}[\rho_{\text{I}}, \rho_{\text{II}}] + T^{\text{nadd}}_s[\rho_{\text{I}}, \rho_{\text{II}}], \tag{12}$$

with $v^{\text{I}}_{\text{nuc}}$ and $v^{\text{II}}_{\text{nuc}}$ as the nuclear potentials due to the set of atoms associated with subsystems I and II, respectively. The repulsion energy for nuclei belonging to different subsystems is described by the $E^{\text{I,II}}_{\text{nuc}}$ term. The non-additive contributions are defined as:

$$X^{\text{nadd}}[\rho_{\text{I}}, \rho_{\text{II}}] = X[\rho_{\text{I}} + \rho_{\text{II}}] - X[\rho_{\text{I}}] - X[\rho_{\text{II}}], \tag{13}$$

with $X = E_{\text{xc}}$, T_s. These terms arise because both exchange-correlation and kinetic energy, in contrast to the Coulomb interaction, are not linear functionals of the density.

The electron density of a given fragment (ρ_{I} or ρ_{II} in this case) can be determined by minimizing the total energy functional (**Eq. 10**) with respect to the density of the fragment while keeping the density of the other subsystem frozen. This procedure is the essence of the FDE scheme and leads to a set of Kohn-Sham-like equations (one for each subsystem)

$$\left[-\frac{\nabla^2}{2} + v^{\text{KS}}_{\text{eff}}[\rho_{\text{I}}](\boldsymbol{r}) + v^{\text{I}}_{\text{emb}}[\rho_{\text{I}}, \rho_{\text{II}}](\boldsymbol{r})\right]\phi^{\text{I}}_k(\boldsymbol{r}) = \varepsilon^{\text{I}}_k\phi^{\text{I}}_k(\boldsymbol{r}) \tag{14}$$

which are coupled by the embedding potential term $v^{\mathrm{I}}_{\mathrm{emb}}(\boldsymbol{r})$, that carries all dependence on the other fragment's density. In the framework of FDE theory, $v^{\mathrm{I}}_{\mathrm{emb}}(\boldsymbol{r})$ is explicitly given by

$$\begin{aligned} v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}](\boldsymbol{r}) &= \frac{\delta E_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}(\boldsymbol{r})} \\ &= v^{\mathrm{II}}_{\mathrm{nuc}}(\boldsymbol{r}) + \int \frac{\rho_{\mathrm{II}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3r' + \frac{\delta E^{\mathrm{nadd}}_{\mathrm{xc}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}(\boldsymbol{r})} \\ &\quad + \frac{\delta T^{\mathrm{nadd}}_{s}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}(\boldsymbol{r})}, \end{aligned} \tag{15}$$

where the non-additive exchange-correlation and kinetic energy contributions are defined as the difference between the associated exchange-correlation and kinetic potentials defined using $\rho_{\mathrm{tot}}(\boldsymbol{r})$ and $\rho_{\mathrm{I}}(\boldsymbol{r})$. It is worth noting that only the density for the total system is available so that potentials requiring KS orbitals as input are excluded.

For the exchange-correlation potential, one may make use of accurate density functional approximations and its quality is therefore similar to that of ordinary KS. The potential for the non-additive kinetic term ($\delta T^{\mathrm{nadd}}_{s}[\rho]/\delta\rho_{\mathrm{I}}(\boldsymbol{r})$, in **Eq. 15**) is more problematic as it relies on less accurate orbital-free kinetic energy density functionals (KEDFs).

In this context, the Thomas-Fermi (TF) kinetic energy functional (Thomas, 1927) or the GGA functional PW91k (Lembarki and Chermette, 1994), are customarily employed. The reader interested in applicability and shortcomings of the functionals associated to $T^{\mathrm{nadd}}_{s}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$ term can refer to Fux et al. (2010) and references therein.

In general, the FDE scheme provides a set of coupled equations for the subsystems that have to be solved iteratively. Typically, the "freeze-and-thaw" (FnT) procedure is employed, meaning that the electron density of the active subsystem is determined, keeping the electron density of the other subsystems frozen, and is then frozen when the electron density of the other subsystems is worked out. The subsystems' densities are converged by repeatedly applying the procedure.

We conclude this section highlighting the main differences between FDE and BOMME approaches. The FDE approach employing an explicit embedding potential allows us to optimize the subsystem of interest limiting the basis set to the sole "active" basis subset. We have already mentioned that the embedding potential relies on the KEDFs, which are in general less accurate than the exchange-correlation counterpart. On the contrary, in BOMME formulation, the self-consistent calculation is carried out in the supermolecular basis, and the embedding is handled implicitly in the calculation. As far as the coupling between subsystems is concerned, the FDE approach is trivial when attempting to estimate the interaction energy of the subsystem and eventually evaluate the net effect of the environment polarization performing an unrelaxed calculation (keeping the environment frozen). It is important to note that since the total density is obtained as the sum of subsystem densities, the partitioning reflects the mean values of observables. In the BOMME approach, the high-level system (AA block) and its environment (BB block) are optimized on the same footing, thus disentangling them could result in a cumbersome procedure. Nevertheless it could be possible to investigate the contribution of the different domains to the overall value of an observable using localization techniques.

2.2 Computational Details

In the ADF (te Velde et al., 2001) calculations, all of which were performed with version `2019.307` (Autschbach et al., 2019), we have employed the AUG/ATZP basis sets for the halogens, and the single-z without polarization (SZ) basis set for the water molecules (Van Lenthe and Baerends, 2003). In supermolecular calculations, we employed the B3LYP functional. The ADF FDE and FnT calculations were performed via the PyADF scripting framework (Jacob et al., 2011). The halogen subsystem has been calculations with the B3LYP functional and the water molecules with BLYP. Since the use of different density functionals for different subsystems in an FnT calculation is currently not possible from within the ADF implementation (Jacob et al., 2008), a PyADF script to carry out such calculations is provided as part of the dataset accompanying this manuscript (De Santis et al., 2021a), and in this case we employed a convergence criteria on the energy of 1×10^{-6}. In all cases, the Thomas-Fermi and BLYP functionals have been used to calculate the non-additive kinetic energy and exchange-correlation contributions to the embedding potentials, respectively. We employed supermolecular integration grids of normal (6.0) accuracy in all calculations.

In the (rt-)BOMME and (rt-)FDE calculations in the Psi4Numpy (Smith et al., 2018) framework, we employed version `1.3.2` of the Psi4 code (Smith et al., 2020) as a computational backend. We have employed the equivalent functionals for the ADF calculations for the halogen and water subsystems (B3LYP and BLYP, respectively). As for basis sets, we employed aug-cc-pVTZ (Dunning, 1989; Kendall et al., 1992; Woon and Dunning, 1994) and STO-3G (Hehre et al., 1969) basis sets for the halogen and the water cluster, respectively. We note that for the Psi4 calculations the basis sets employed are those provided by the code's own basis set library.

For the real-time simulations, the electronic ground state of the halogen-water complex, calculated in absence of an external electric field, was perturbed by an analytic δ-function pulse with a strength of $\kappa = 5.0 \times 10^{-4}$ a.u. along the three directions, x, y, z. The induced dipole moment has been collected for 56,000 time steps with a length of 0.025 a.u. per time step, corresponding to 33.9 fs of simulation. The choice of such a fine-grained time grid ensures in principle an observable frequency up to 3419.5 eV in the power spectrum distribution. In the case of the fluorine-water complex, the near-edge structure is located in the range of 665–700 eV, thus the time-dependent dipole moment was down-sampled halving the amount of sampling points. The use of Padé approximant-based Fourier transform allowed us to further reduce the length of the signal to be sampled corresponding to an "effective" dipole moment of 24 and 29 fs for the fluorine- and chlorine-water complex respectively. In both cases prior to Fourier transformation an exponential damping $e^{-\lambda \cdot t}$ with $\lambda = 3.0 \times 10^{-4}$ was applied.

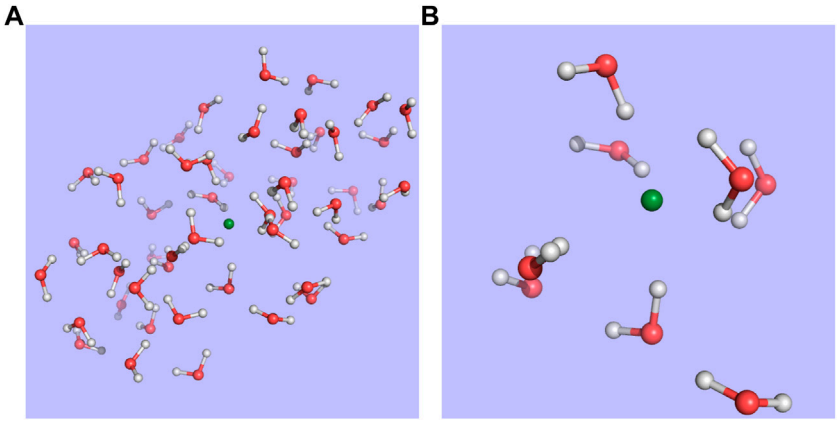


FIGURE 1 | Structure for `snapshot 1` of the fluoride-water droplet system taken from Bouchafra et al. (Bouchafra et al., 2018a; Bouchafra et al., 2018b) **(A)** and the model used in this work **(B)**, in which the halide and the eight nearest water molecules making up its first solvation shell were extracted from the 50-water droplet.

The code implementing the rt-FDE in the Psi4Numpy framework used in this work is part of the PyBertha package (De Santis et al., 2020a; De Santis et al., 2020b; De Santis et al., 2021b) (revision `3c752072`). The code implementing the (rt-) BOMME approach is under version control (Git) but does not yet have a public release version (one is envisaged for 2022). The simulations described in the paper have been carried out with revision `3c4c334b`.

The structures employed in the calculations were taken from the structures generated by Bouchafra et al. (Bouchafra et al., 2018a; Bouchafra et al., 2018b) for halogens in 50-water droplets–in particular, `snapshot 619` for chloride and `snapshot 1` for fluoride–and, for reasons of computational cost, we have only kept the nearest 8 water molecules that correspond to the first solvation shell. This setup is exemplified in **Figure 1**.

We provide as part of the **Supplemental Information** a comparison of the effect of replacing the aug-cc-pVTZ with the aug-cc-pCVTZ basis set (for BOMME and supermolecule calculations), as well as performance metrics for the real-time simulations.

3 RESULTS AND DISCUSSION

We proceed now to the presentation and discussion of our results. Before doing so, we recall that since we are interested in the relative performance of the embedding methods with respect to a calculation on the whole system, and not in a comparison to the experiment, we have opted to disregard both relativistic effects and statistical sampling of different solute configurations (e.g., by considering different snapshots from molecular dynamics simulations) as done by some of us in Bouchafra *et al.* (Bouchafra et al., 2018a), which we aim to consider in subsequent work. Second, we chose to focus only on transitions from the core s orbitals of the halogens, that is, the 1 s (K edges) for F^- and Cl^-, and the 2 s (L_1 edge) of Cl^-, since they provide sufficient information for our method comparisons.

3.1 Ground States

Before investigating the outcome of the real-time propagation of the wavefunctions, it is instructive to analyze the differences between the different models: isolated atoms, embedding approaches (FDE and BOMME), and standard (supermolecular) DFT calculations. To this end, we shall focus on the comparison of core orbital energies, on the one hand, since there is a direct connection between them and how environment effects are incorporated (see discussion on theoretical approaches in **Section 2.1**), and, on the other hand, their values provide an approximation to the ionization potentials–though the very important effects of wavefunction relaxation will still be missing due to the creation of the core hole.

While core orbitals are naturally rather localized, they are nevertheless quite sensitive to changes in the surroundings of the atom due to the presence of the solvent molecules, as we can see from the comparison of values for the isolation anions and the supermolecular systems. With the embedding approaches we expect the orbital energies to be much closer to the supermolecular values, since they introduce the different interactions (electrostatic, kinetic energy, and exchange-correlation) between the halides and the water molecules, albeit in more or less approximated manners. Consequently, the closer an embedding approach yields orbital energies to supermolecular ones, the better suited it can be considered to replace the supermolecular calculation.

Before we can proceed to a comparison between the FDE and BOMME results shown in **Table 1**, we should note that our Psi4Numpy-based code, in which both embedding approaches are implemented, does not yet implement the "freeze-and-thaw" (FnT) procedure for the FDE case. While that posed no problem in its first application to the rt-TDDFT-in-DFT simulation of a neutral system (De Santis et al., 2020a), prior work by some of us (Gomes et al., 2008; Gomes et al., 2013; Bouchafra et al., 2018a) has shown that for charged systems such as those considered here, the manner in which the environment density has been constructed is important, and that a relaxation of both the active subsystem and the environment densities via FnT can

TABLE 1 | Orbital energies (ϵ, in eV) of the core s orbitals the halogen atom in the $[X(H_2O)n]^-$ clusters, (X = F^-, Cl^-), obtained with different models: (0) isolated halogen atoms; (1) DFT-in-DFT without the relaxation of the solvent (FDE); (2) DFT-in-DFT with the relaxation of the solvent environment (FnT); and (3) block-orthogonalized Manby-Miller embedding (BOMME). In addition to the energies obtained with embedding, we provide energy differences with respect to reference supermolecular DFT calculations ($\Delta\epsilon = \epsilon_{sup} - \epsilon_{model}$), represented by $\Delta\epsilon_0$, $\Delta\epsilon_1$, $\Delta\epsilon_2$, and $\Delta\epsilon_3$, respectively. Due to the fact that for technical reasons, in our FDE implementation based on the Psi4Numpy framework, we are currently not able to perform an embedding scheme (2), we provide results for (1) and (2) obtained with the ADF code.

Orbital energies (eV)										
Framework	X	Orbital	Iso. (0)	FDE (1)	FnT (2)	BOMME (3)	$\Delta\epsilon_0$	$\Delta\epsilon_1$	$\Delta\epsilon_2$	$\Delta\epsilon_3$
ADF	F^-	1s	−659.67	−661.16	−661.59		−2.57	−0.67	0.46	
	Cl^-	1s	−2753.62	−2755.05	−2755.48		−1.55	−0.12	0.32	
		2s	−248.46	−249.92	−250.36		−1.59	−0.12	0.31	
Psi4Numpy	F^-	1s	−659.88	−661.96	−662.13		−2.42	−0.34		−0.17
	Cl^-	1s	−2754.83	−2756.28		−2756.25	−1.50	−0.05		−0.08
		2s	−248.78	−250.27		−250.25	−1.55	−0.06		−0.08

improve the results over a pure FDE calculation, in which the environment's density and electrostatic potential have been obtained in the absence of the halides.

To estimate the effect of relaxing the environment on the orbital energies, and indirectly part of its influence on the simulation of core spectra (the other part coming from the effect on the halide virtual orbitals), we also present in **Table 1**, results obtained with the ADF code, in which both FDE and FnT calculations have been carried out. We see that the FDE calculations tend to overestimate the effect of the environment, via overall more attractive embedding potentials, reflected in lower orbital energies than the supermolecular case, whereas FnT reverses this trend but overcorrects somewhat and yields energies which are slightly higher than the supermolecular ones. For fluoride, FDE and FnT differ by roughly 1.2 eV, whereas for chloride there is a much smaller important difference, of around 0.4 eV. In addition to being smaller in magnitude, the shift for chloride is roughly the same for both 1 and 2 s orbitals, an observation that is consistent with prior work (Gomes et al., 2013) in which we observed that the embedding potential shifted orbital energies in a nearly constant manner across different occupied orbitals.

Comparing the differences between FDE and supermolecule results between ADF and Psi4Numpy, we see a similar trend in that FDE overestimates the effect of the environment. From the comparison of $\Delta\epsilon_0$ for the two codes, we see that discrepancies of around 0.15 eV (for fluoride) and 0.05 eV (for chloride) can be attributed to differences inherent to the two sets of calculations (Slater vs. Gaussian basis sets, etc.), with values calculated with ADF showing larger discrepancies between isolated and supermolecular calculations than Psi4Numpy. If we correct for these differences, we see that for chloride the $\Delta\epsilon_1$ values are consistent between codes, though for fluoride even taking into account such corrections, non-negligible differences between codes, of around 0.15 eV, remain. From this comparison, we believe that we can conclude that, if we were able to carry out such calculations, the Psi4Numpy FnT $\Delta\epsilon_2$ would likely be of around 0.2–0.3 eV for chloride, and 0.4–0.5 eV for fluoride.

The BOMME results show a similar trend to the FDE in overestimating the effect of the environment with respect to supermolecular results, and that such an overestimation is larger for fluoride than for chloride. The magnitude of such an effect, however, is about half of that of FDE for fluoride (-0.17 eV vs -0.34 eV), and roughly equivalent to that of FDE (less than -0.1 eV) for chloride. The differences between BOMME and FDE are consistent with what is known in the literature between the more reliable behavior of projection-based embedding (such as BOMME) in describing cases in which there are strong interactions between the different subsystems with respect to FDE, which suffers from the limited accuracy of the non-additive kinetic energy density functionals used to calculate the non-additive kinetic energy contribution to the embedding potential (Gomes and Jacob, 2012).

From the discussion above, and assuming that the dominant effects in the electronic spectra would come mainly from the energy differences between the core and low-lying virtuals either on the halogen (for both BOMME and FDE) or on the environment (for BOMME), we can expect to see that BOMME excitation energies would be consistently closer to the supermolecular results than FDE, but that such a difference would decrease for chloride. In the following we

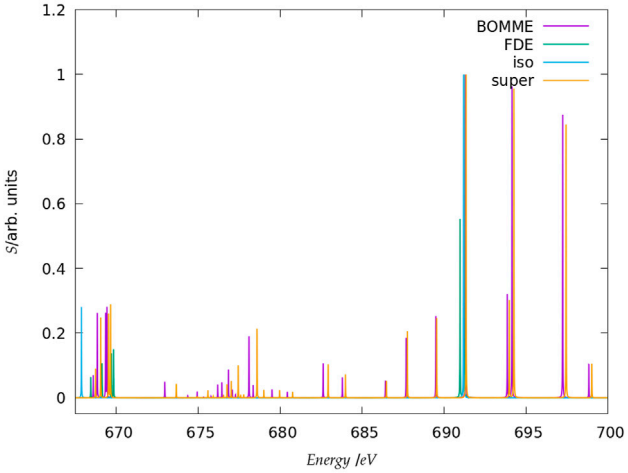


FIGURE 2 | Simulated K-edge spectra for the fluoride model system, over the roughly 30 eV interval starting at the free ion edge peak. It should be noted that here the peak heights (in arbitrary units) for each family of models: free ion (=iso), FDE, BOMME, and supermolecule (super) have been scaled, with a height of 1 assigned to the most intense transition.

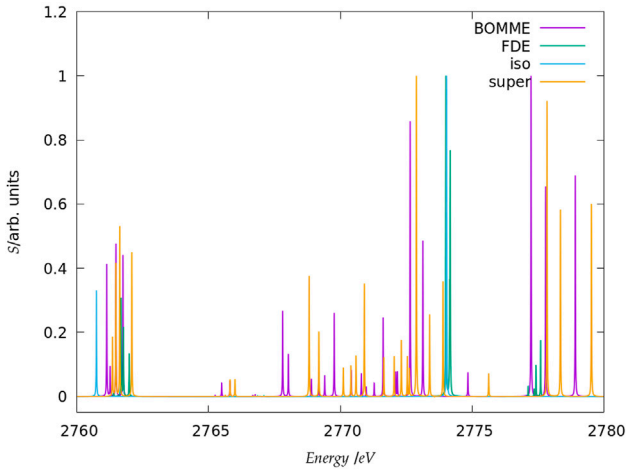


FIGURE 3 | Simulated K-edge spectra for the chloride model system, over the roughly 30 eV interval starting at the free ion edge peak. It should be noted that here the peak heights (in arbitrary units) for each family of models: free ion (=iso), FDE, BOMME, and supermolecule (=super) have been scaled, with a height of 1 assigned to the most intense transition.

shall see to what extent this picture holds true. In any case, for the core orbital energies, the FDE results seem to provide a fortuitous error cancellation that places the relatively simple FDE model on par with the much more sophisticated BOMME.

3.2 Core Excited States

Before discussing the behavior of the different approaches for the core states, we note that in the following we shall focus on the edge region for the K edges of both systems, considering energies spanning a somewhat broad window (20–30 eV higher than the first peaks, in order to have a wider region in which to compare the different models) and the L_1 edge of chloride. Furthermore, in the discussion below we shall focus on combined contributions from the *x*, *y*, *z* components of the perturbing field. We present a breakdown of these by component of the perturbing field, along with the spectra for the whole regions under consideration (**Figures 2**, **3**, **5**) in the **Supplemental Information**.

3.2.1 K Edge

The spectra for the K edges of fluoride and chloride are shown in **Figure 4**. Starting with the simplest systems, for the isolated anions, we note that, as expected, the K-edge spectra corresponds to transitions from the 1 s orbitals to the first virtual halide p orbitals ((n + 1)p). The second peak in the energy range considered, on the other hand, corresponds to transitions from the 1 s to higher-lying halide p-type orbitals ((n + 2)p).

Second, at the other extreme we have the supermolecular calculations on the microsolvated anions (which here serve as a benchmark to which the embedding approaches will be compared). The first remarkable difference from the free ions is that there is an environment-induced shift in the first region (**Figures 4A,B**), which at around 1 eV is fairly similar between systems, but about half of what would be expected from the difference in orbital energies between the free ions and the supermolecular systems ($\Delta\varepsilon_0$). This is the first indication that core orbital energies are useful for understanding the K-edge absorption spectra from a semi-quantitative viewpoint at best.

The asymmetrical first hydration shell environment also breaks the atomic symmetry, which has as consequence the lifting of the selection rules for the atom, and introduces differential interactions with the different p orbitals that become occupied in the excited states. As a result of that, for both fluoride and chloride, in the supermolecular calculations we observe four transitions within roughly a 1 eV interval, with spacings of around 0.5 eV between the first three peaks (with the fourth being much closer to the third).

For the region corresponding to the second free ion transition (**Figures 4E,F**), we observe a similar situation to that of the first, with the environment inducing a symmetry breaking of the higher-lying p orbitals. In the case of fluoride, for the supermolecular calculations we observe five peaks, one rather close to that of the free ion (around 691.5 eV), followed by two other peaks around 2 eV higher (at 694 eV), and two more peaks between 697 and 699 eV. In the case of chloride, we also observe five peaks, one nearest to that of the free ion, another peak around 2775.5 eV, and then three additional peaks between 2778 and 2780 eV. Finally, in between the two free ion peaks (**Figures 4C,D**), in the supermolecular calculation we have a region that contains several peaks.

Considering now the FDE calculations–and recalling that these correspond to a situation in which the density of the environment has not been relaxed in the presence of the anions–we see, for fluoride, a semi-quantitative agreement with the supermolecular calculation; for the lower energy region (**Figures 4A,B**), the first peak appears in slightly (around 0.2 eV) lower energies than the supermolecular ones, while the second, third, and fourth peaks appear at slightly higher energies. For chloride, the peak positions are overall closer (0.1 eV or less) to the supermolecular one than for fluoride, but now the energies of the first three peaks are slightly overestimated with respect to the supermolecule, while in the fourth we see a slight underestimation. We note that this is in line with the better agreement between FDE and supermolecular 1 s orbital energies than for fluoride. Such a tendency was also observed by Bouchafra et al. (Bouchafra et al., 2018a), though for valence ionizations. It was shown that an FDE model in which only the halide belonged to the active subsystem was not a good representation for the solvated ion, due to the strong water-fluoride interactions in the valence regions, whereas for chloride (and other heavier halides) this simple FDE model containing no explicit halide-water interactions was quite good. In the higher energy region (**Figures 4E,F**), we see that the peaks from FDE calculations are also close to that of the free ion and of the supermolecule (around 2774 eV), though for FDE we observe another two peaks, just over 2774 eV (that show a very small splitting), and four others between 2777 and 2778 eV. The behavior of FDE for this higher energy range is, therefore, in stark contrast to the lowest energy range considered, since there is not even qualitative agreement with the supermolecular results.

Having in mind that in the supermolecular case the complete system is allowed to respond to the external perturbation, but that

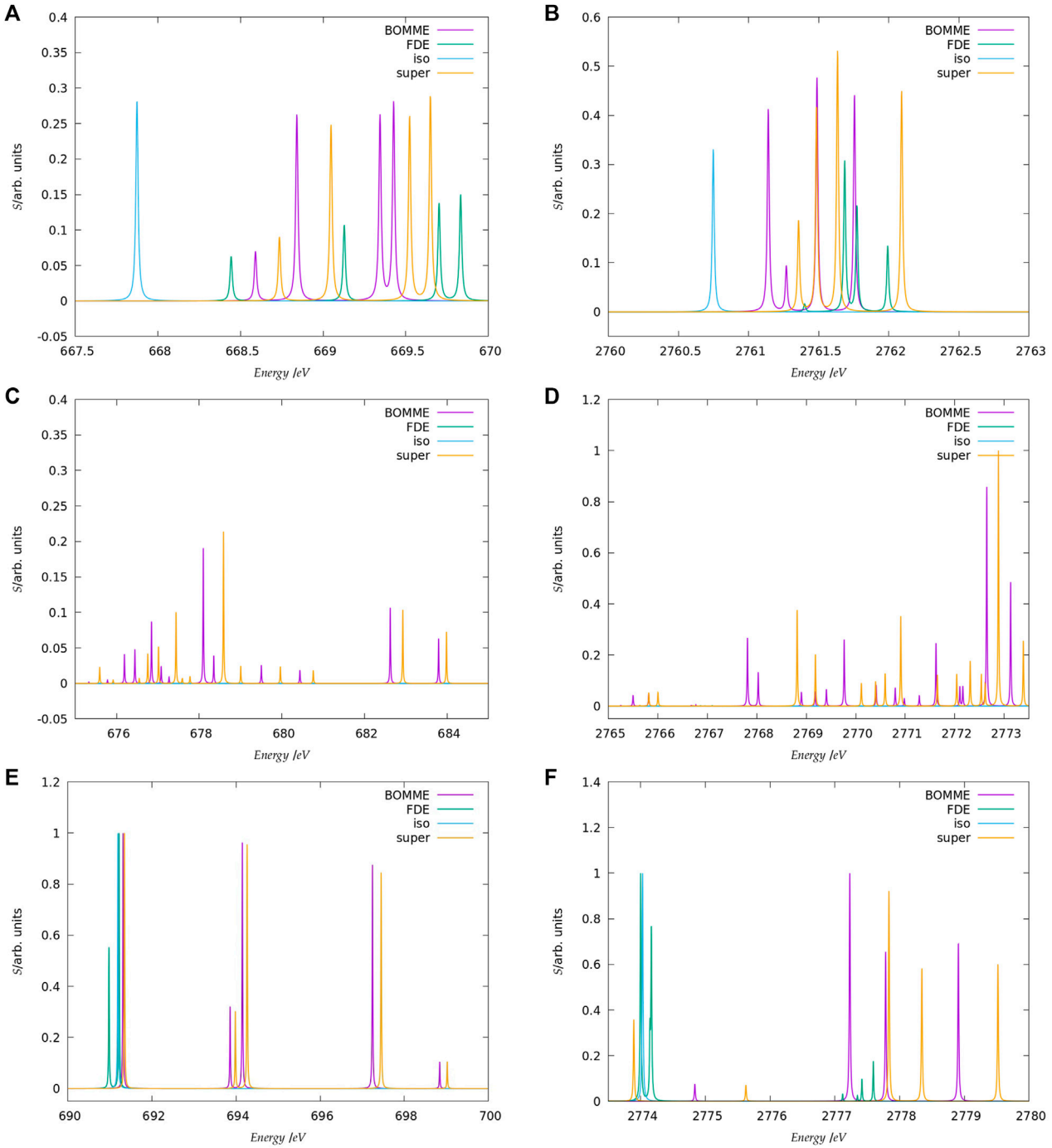


FIGURE 4 | Details on the three main energy ranges for the K-edge spectra of F^- **(A,C,E)** and Cl^- **(B,D,F)** model systems selected from the spectra shown in **Figures 2**, **3**.

by construction the response of the environment is absent in the FDE case, this discrepancy provides the first indication of the importance of the response of the environment for higher energies. This is further underscored by the fact that the FDE calculations show no peaks in the intermediate energy range considered. Consequently, we can safely say that the intermediate energy range is, in effect, dominated by excitations from the halide to virtuals with strong (if not dominant) contributions from the environment. Furthermore, while the behavior of FDE is in line with the difference between the isolated and FDE orbital energies, in particular for the lower energy part of the spectra, the situation is less clear-cut with respect to a comparison to the supermolecule. We consider the discrepancies in this case to partly arise from the lack of relaxation for fluoride virtual orbitals

and partly from the lack of coupling between the response of the subsystems, as discussed below.

Now comparing supermolecular and BOMME calculations, we see that for both systems the latter provides an overall improvement over FDE–already in qualitative terms, with BOMME we are able to capture the contributions to the solvent to the different electronic states, and, furthermore, in all energy ranges BOMME systematically approaches (underestimates) the excitation energies. In quantitative terms, for fluoride, BOMME clearly performs better than FDE; this can already be seen from the orbital energy differences, and in the low-energy range, the differences in absolute are not very large but BOMME does show smaller differences. For the larger energy range, where FDE is not even qualitatively correct, BOMME shows discrepancies of around 0.1 eV or less.

For the intermediate energy range, on the other hand, we see more significant differences between the BOMME and supermolecular energies. We attribute this to the use of a GGA functional for the environment, since GGAs tend to underestimate excitation energies with respect to hybrid functionals (Besley, 2021), and it is precisely in this region that contributions from the environment have a prominent role. While the goal of BOMME is to replace a high-level description of the environment for a lower-level one, our results are the first indication that for core excited states in which the environment plays an important role, the quality of the low-level of theory may matter much more than for valence states.

The trends outlined above for fluoride are also seen for chloride; if for the low-energy region BOMME does not bring about as significant an improvement over FDE as for fluoride, for the other two energy ranges BOMME provides a semi-quantitative agreement due to its systematic behavior. We see, however, that for the higher energy range, we have often energy differences between BOMME and the supermolecule in the order of 1.0 eV, and nearly so for many of the states in the intermediate region. This further underscores the importance of the quality of the density functionals employed for core energies, and in particular deeper cores such as the chloride K edge.

We can now compare these results to what one could expect, the simple argument put forward above that the halide orbital energies would provide a dominant contribution to the excited states. From the orbital energies alone, we would expect that both BOMME and FDE would yield excitation energies larger than supermolecular ones; considering only the low-energy range, for both halides this is at odds with the BOMME results, which always show lower energies than the supermolecule, but the simple orbital picture is more consistent with the FDE results, since for three peaks out of four they appear at slightly higher excitation energies (for fluoride, the first FDE peak appears at a lower energy than the supermolecule, whereas for chloride that happens for the fourth peak). We consider this is yet another piece of evidence of the importance of the virtual orbitals from the environment to characterize the excitation energies obtained with BOMME, since such contributions are absent by construction in the case of FDE.

To deepen the discussion on the origin of differences between embeddding and supermolecular results, it is useful to analyse the molecular orbitals involved in the K edges. In the real-time framework, similarly to LR, the absorption cross-section can be interpreted in terms of occupied-virtual molecular orbital (MO) pairs. This approach has been proposed originally by Repisky and coworkers (Repisky et al., 2015) and implemented in the relativistic code ReSpect (Repisky et al., 2020). The scheme has been slightly reworked by Bruner et al. (Bruner et al., 2016) accelerating the methodology applying the Padé approximants to the Fourier transform of the deconvolution of the induced dipole into molecular orbital pairs. The method gives the MO contribution to the dipole strength function at all frequencies. The relative areas under each peak correspond to MO contribution to an excitation at that frequency, which gives a representation of transitions consistent with the linear response.

In order to have a more visual interpretation of transitions, one can return to transition density plots. We recall that the simplest direct approach (Hofmann and Kümmel, 2012) is to evaluate the transition density at a preset frequency according to:

$$\rho(\boldsymbol{r},\omega) \propto -\mathrm{Im}\{\delta\tilde{\rho}(\boldsymbol{r},\omega)\}, \tag{16}$$

where $\delta\tilde{\rho}(\boldsymbol{r},\omega)$ is the Fourier transform of the time-dependent (TD)-induced density $\delta\rho(\boldsymbol{r}, t) = \rho(\boldsymbol{r}, t) - \rho(\boldsymbol{r}, 0)$. Many research groups have recently contributed to this topic (Rossi et al., 2017; Schelter and Kümmel, 2018; Sinha-Roy et al., 2018; Jornet-Somoza and Lebedeva, 2019), focusing on low-frequency excitations. To the best of our knowledge, these methods have not been applied yet in the high-frequency range.

In the **Supplemental Information**, we can observe from the isosurface plot, $\delta\tilde{\rho}(\boldsymbol{r},\omega_i)$ for ω_i = 668.586 eV, that the surrounding waters contribute to the K edge. At the same time, we should point out that the Fourier transform of the TD-induced dipole (using GNU Octave (Eaton et al., 2020) routines) is of exceptionally poor quality in the frequency range of the fluorine K edge. It can be argued reasonably that also $\delta\tilde{\rho}(\boldsymbol{r},\omega_i)$ is not of the best quality possible. With that, we have been unable to apply the methodology developed by Schelter and co-workers, in which the accurate value of the oscillator strength is extracted from the refined dipole strength function (DSF) of the TD-induced dipole moment. The lack of high quality DSF makes it difficult to provide an accurate estimation of both the transition dipole moment and transition density. This further prevents us from applying more elaborate methods, namely natural transition orbitals. These difficulties should be addressed in dedicated future works.

In the meantime, in the present work we provide in the **Supplemental Information** an analysis of an LR-TDDFT calculation on the supermolecular system, performed with the NWChem code (Valiev et al., 2010) (using the same basis sets and density functionals as the Psi4 calculations), from which we can determine that lower-energy K-edge transitions of $[F(H_2O)_8]^-$ indeed involve virtual orbitals with contributions from orbitals centered on the oxygen atoms, in line with the qualitative picture we have managed to extract from the isosurface plots of $\delta\tilde{\rho}(\boldsymbol{r},\omega_i)$.

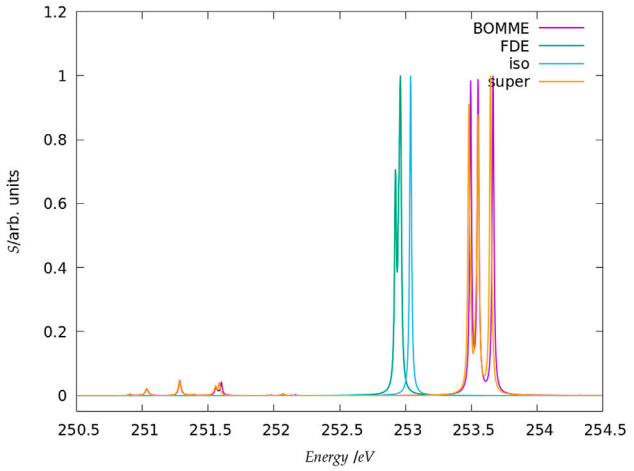


FIGURE 5 | Simulated L_1-edge spectra for the fluoride model system in the edge region (contrary to the K edge, no peaks of appreciable intensity have been observed at higher energies). It should be noted that here the peak heights (in arbitrary units) for each family of models: free ion (=iso), FDE, BOMME, and supermolecule (super) have been scaled, with a height of 1 assigned to the most intense transition.

3.2.2 L_1 Edge

The spectra for the chloride L_1 edge region are shown in **Figure 5**. Unlike the case for the K edge, here we do not show a larger energy range since, for energies between 245 and 260 eV, there are no other peaks with appreciable intensity other than those in the picture. A comparison between the free ion at around 253 eV and the supermolecular peaks here show that the energy shift due to the environment is not as marked (around 0.5 eV higher) but nevertheless sufficient to clearly characterize the interaction with the waters through the splitting of the peaks. In this case, we have a near perfect agreement between the supermolecular and BOMME results in terms of peak positions, at around 253.5 eV but also for low intensity transitions around 251.5 eV. We also observe that, unlike for the K edge, here the BOMME results slightly overestimate the supermolecular ones, in agreement with the orbital energy differences ($\Delta\epsilon_3$) in **Table 1**. We consider that this point, and the absence of other peaks as in the K edge that would indicate more or fewer important contributions from the environment, the L1 edge is almost exclusively dominated by halide to halide transitions. The FDE results, on the other hand, underestimate the effect of the environment and show almost no difference to the free ion results, apart from the fact that a splitting of the peak, much less significant than that seen for BOMME or the supermolecule, is also seen. This is a further indication that the FDE approach has not properly captured the perturbations to the virtual orbitals of chloride induced by the solvent.

4 CONCLUSION

In this manuscript we have carried out an investigation, to the best of our knowledge for the first time, of the accuracy of fully quantum mechanical (QM/QM), DFT-based embedding approaches–namely, Frozen Density embedding (FDE) and block-orthogonalized Manby-Miller embedding (BOMME) approaches–in the description of core excitation spectra (XAS), obtained by the real-time propagation of the electron density, for model systems representing the hydration of halide ions, comprising the halide ions (fluoride and chloride) as active subsystems, and the eight water molecules in the first solvation shells as the environment.

We note that the BOMME approach and its real-time variant has been implemented within the Psi4Numpy framework in which some of us had previously implemented the real-time TDDFT-in-DFT FDE method, thereby facilitating a rigorous one-to-one comparison between approaches.

From our comparison of the two embedding methods to reference DFT calculations on the whole model systems, we observe first that the BOMME approach can better describe the fluoride core orbital energies in the fluoride-water system than FDE, due to its better handling of the stronger interactions between subsystems, while for chloride both BOMME and FDE perform rather similarly.

In the case of real-time simulations, we have found that the rt-BOMME approach follows the behavior of the reference rt-TDDFT calculations in a very systematic manner across all energy ranges investigated, and as such can potentially become very useful in the investigation of core spectra of species in confined or complex environments.

We observe that rt-BOMME tends to slightly underestimate the supermolecular results around the edge region for both K and L_1 edges, where excitations mostly take place between orbitals belonging to the halides. On the other hand, we observe much more important discrepancies in higher energy regions for which, it turns out, the environment plays a more important role.

We attribute these discrepancies to the fact that in rt-BOMME the environment is described with a lower-accuracy GGA functional, a functional class which tends to underestimate core excitation energies due to larger self-interaction errors in comparison to the hybrid functionals which were employed for the active subsystem, and for the whole system in the supermolecular calculations. Our results call for particular attention, in the case of core spectra, in choosing the density functional for the environment, in order to minimize artifacts in the simulations.

The rt-TDDFT-in-DFT simulations carried out under the constraint that the density of the environment has not been relaxed, have nevertheless shown performances similar to the rt-BOMME and reference rt-TDDFT simulations for the (pre-)edge regions. However, since in our implementation the response of the environment is also lacking, large parts of the spectra are either inaccessible or are not correctly described. We intend to address this issue, and introduce the coupling between the response of the different subsystems, in subsequent work.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://zenodo.org/record/5729961.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

We acknowledge funding from projects Labex CaPPA (ANR-11-LABX-0005-01) and CompRIXS (ANR-19-CE29-0019, DFG JA 2329/6-1), the I-SITE ULNE project OVERSEE and MESONM International Associated Laboratory (LAI) (ANR-16-IDEX-0004), and support from the French national supercomputing facilities (grant DARI A0090801859).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2022.823246/full#supplementary-material

REFERENCES

Autschbach, J., Baseggio, O., Bashford, D., Bérces, A., Bickelhaupt, F. M., Bo, C., et al. (2019). *ADF 2019.3, SCM, Theoretical Chemistry*. Amsterdam, The Netherlands: Vrije Universiteit. Avalableat: http://www.scm.com. Optionally (you may add the following list of authors and contributors: E.J. Baerends, T. Ziegler, A.J. Atkins(.)

Bartels-Rausch, T., Kong, X., Orlando, F., Artiglia, L., Waldner, A., Huthwelker, T., et al. (2021). Interfacial Supercooling and the Precipitation of Hydrohalite in Frozen NaCl Solutions as Seen by X-ray Absorption Spectroscopy. *The Cryosphere* 15, 2001–2020. doi:10.5194/tc-15-2001-2021

Besley, N. A., and Asmuruf, F. A. (2010). Time-dependent Density Functional Theory Calculations of the Spectroscopy of Core Electrons. *Phys. Chem. Chem. Phys.* 12, 12024–12039. doi:10.1039/C002207A

Besley, N. A. (2021). Modeling of the Spectroscopy of Core Electrons with Density Functional Theory. *Wires Comput. Mol. Sci.* 11. doi:10.1002/wcms.1527

Beyhan, S. M., Götz, A. W., Jacob, C. R., and Visscher, L. (2010). The Weak Covalent Bond in NgAuF (Ng=ar, Kr, Xe): A challenge for Subsystem Density Functional Theory. *J. Chem. Phys.* 132, 044114. doi:10.1063/1.3297886

Bouchafra, Y., Shee, A., Réal, F., Vallet, V., and Gomes, A. S. P. (2018b). *Predictive Simulations of Ionization Energies of Solvated Halide Ions with Relativistic Embedded Equation of Motion Coupled-Cluster Theory*. Genève: Dataset. doi:10.5281/zenodo.1477004

Bouchafra, Y., Shee, A., Réal, F., Vallet, V., and Gomes, A. S. P. (2018a). Predictive Simulations of Ionization Energies of Solvated Halide Ions with Relativistic Embedded Equation of Motion Coupled Cluster Theory. *Phys. Rev. Lett.* 121, 266001. doi:10.1103/PhysRevLett.121.266001

Bruner, A., LaMaster, D., and Lopata, K. (2016). Accelerated Broadband Spectra Using Transition Dipole Decomposition and Padé Approximants. *J. Chem. Theor. Comput.* 12, 3741–3750. doi:10.1021/acs.jctc.6b00511

Bunker, G. (2009). *Introduction to XAFS*. Cambridge University Press. doi:10.1017/cbo9780511809194

Castro, A., Marques, M. A. L., and Rubio, A. (2004). Propagators for the Time-dependent Kohn-Sham Equations. *J. Chem. Phys.* 121, 3425–3433. doi:10.1063/1.1774980

Cheng, C.-L., Evans, J. S., and Van Voorhis, T. (2006). Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* 74, 155112. doi:10.1103/PhysRevB.74.155112

de Groot, F. (2001). High-resolution X-ray Emission and X-ray Absorption Spectroscopy. *Chem. Rev.* 101, 1779–1808. doi:10.1021/cr9900681

De Santis, M., Belpassi, L., and Storchi, L. (2021). Pyberthaembedrt, a Code to Perform Real-Time Frozen Density Embedding Calculations. Avalableat: https://github.com/BERTHA-4c-DKS/pybertha/tree/master/pyberthaembedrt.

De Santis, M., Vallet, V., and Gomes, A. S. P. (2021). *Dataset: Environment Effects on X-ray Absorption Spectra with Quantum Embedded Real-Time Time-dependent Density Functional Theory Approaches*. doi:10.5281/zenodo.5729961

De Santis, M., Belpassi, L., Jacob, C. R., Gomes, A. S. P., Tarantelli, F., Visscher, L., et al. (2020). Environmental Effects with Frozen-Density Embedding in Real-Time Time-dependent Density Functional Theory Using Localized Basis Functions. *J. Chem. Theor. Comput.* 16, 5695–5711. doi:10.1021/acs.jctc.0c00603

De Santis, M., Storchi, L., Belpassi, L., Quiney, H. M., and Tarantelli, F. (2020). Pyberthart: A Relativistic Real-Time Four-Component Tddft Implementation Using Prototyping Techniques Based on python. *J. Chem. Theor. Comput.* 16, 2410–2429. doi:10.1021/acs.jctc.0c00053

Ding, F., Manby, F. R., and Miller, T. F. (2017). Embedded Mean-Field Theory with Block-Orthogonalized Partitioning. *J. Chem. Theor. Comput.* 13, 1605–1615. doi:10.1021/acs.jctc.6b01065

Dunning, T. H., Jr. (1989). Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms boron through Neon and Hydrogen. *J. Chem. Phys.* 90, 1007–1023. doi:10.1063/1.456153

Eaton, J. W., Bateman, D., Hauberg, S., and Wehbring, R. (2020). *GNU Octave Version 5.2.0 Manual: A High-Level Interactive Language for Numerical Computations*.

Eberly, J. H., Javanainen, J., and Rzążewski, K. (1991). Above-threshold Ionization. *Phys. Rep.* 204, 331–383. doi:10.1016/0370-1573(91)90131-5

Ekström, U., Norman, P., Carravetta, V., and Ågren, H. (2006). Polarization Propagator for X-ray Spectra. *Phys. Rev. Lett.* 97, 143001. doi:10.1103/PhysRevLett.97.143001

Ekström, U., and Norman, P. (2006). X-ray Absorption Spectra from the Resonant-Convergent First-Order Polarization Propagator Approach. *Phys. Rev. A.* 74, 042722. doi:10.1103/PhysRevA.74.042722

Fahleson, T., Ågren, H., and Norman, P. (2016). A Polarization Propagator for Nonlinear X-ray Spectroscopies. *J. Phys. Chem. Lett.* 7, 1991–1995. doi:10.1021/acs.jpclett.6b00750

Finlayson-Pitts, B. J. (2019). Multiphase Chemistry in the Troposphere: It All Starts . . . and Ends . . . with Gases and Ends with Gases. *Int. J. Chem. Kinet.* 51, 736–752. doi:10.1002/kin.21305

Finlayson-Pitts, B. J. (2013). Chlorine Chronicles. *Nat. Chem* 5, 724. doi:10.1038/nchem.1717

Fux, S., Jacob, C. R., Neugebauer, J., Visscher, L., and Reiher, M. (2010). Accurate Frozen-Density Embedding Potentials as a First Step towards a Subsystem Description of Covalent Bonds. *J. Chem. Phys.* 132, 164101. doi:10.1063/1.3376251

Genova, A., Ceresoli, D., Krishtal, A., Andreussi, O., DiStasio, R. A., and Pavanello, M. (2017). eQE: An Open-source Density Functional Embedding Theory Code for the Condensed Phase. *Int. J. Quan. Chem.* 117. doi:10.1002/qua.25401

Gil, G., Pipolo, S., Delgado, A., Rozzi, C. A., and Corni, S. (2019). Nonequilibrium Solvent Polarization Effects in Real-Time Electronic Dynamics of Solute Molecules Subject to Time-dependent Electric fields: A New Feature of the Polarizable Continuum Model. *J. Chem. Theor. Comput.* 15, 2306–2319. doi:10.1021/acs.jctc.9b00010

Goez, A., and Neugebauer, J. (2018). "Embedding Methods in Quantum Chemistry,". *Frontiers of Quantum Chemistry*. Editors M. J. Wójcik, H. Nakatsuji, B. Kirtman, and Y. Ozaki (Singapore: Springer Singapore), 139–179. doi:10.1007/978-981-10-5651-2_7

Goings, J. J., Lestrange, P. J., and Li, X. (2017). Real-time Time-dependent Electronic Structure Theory. *Wires Comput. Mol. Sci.* 8, e1341. doi:10.1002/wcms.1341

Gomes, A. S. P., Jacob, C. R., Réal, F., Visscher, L., and Vallet, V. (2013). Towards Systematically Improvable Models for Actinides in Condensed Phase: the Electronic Spectrum of Uranyl in Cs2UO2Cl4 as a Test Case. *Phys. Chem. Chem. Phys.* 15, 15153–15162. doi:10.1039/C3CP52090K

Gomes, A. S. P., Jacob, C. R., and Visscher, L. (2008). Calculation of Local Excitations in Large Systems by Embedding Wave-Function Theory in Density-Functional Theory. *Phys. Chem. Chem. Phys.* 10, 5353–5362. doi:10.1039/b805739g

Gomes, A. S. P., and Jacob, C. R. (2012). Quantum-chemical Embedding Methods for Treating Local Electronic Excitations in Complex Chemical Systems. *Annu. Rep. Prog. Chem. Sect. C: Phys. Chem.* 108, 222–277. doi:10.1039/C2PC90007F

Goodpaster, J. D., Ananth, N., Manby, F. R., and Miller, T. F. (2010). Exact Nonadditive Kinetic Potentials for Embedded Density Functional Theory. *J. Chem. Phys.* 133, 084103. doi:10.1063/1.3474575

Goodpaster, J. D., Barnes, T. A., and Miller, T. F. (2011). Embedded Density Functional Theory for Covalently Bonded and Strongly Interacting Subsystems. *J. Chem. Phys.* 134, 164108. doi:10.1063/1.3582913

Graham, D. S., Wen, X., Chulhai, D. V., and Goodpaster, J. D. (2020). Robust, Accurate, and Efficient: Quantum Embedding Using the Huzinaga Level-Shift Projection Operator for Complex Systems. *J. Chem. Theor. Comput.* 16, 2284–2295. doi:10.1021/acs.jctc.9b01185

Grimmel, S. A., Teodoro, T. Q., and Visscher, L. (2019). Is it Worthwhile to Go beyond the Local-density Approximation in Subsystem Density Functional Theory? *Int. J. Quan. Chem.* 120. doi:10.1002/qua.26111

Hehre, W. J., Stewart, R. F., and Pople, J. A. (1969). Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.* 51, 2657–2664. doi:10.1063/1.1672392

Höfener, S., Gomes, A. S. P., and Visscher, L. (2012). Molecular Properties via a Subsystem Density Functional Theory Formulation: A Common Framework for Electronic Embedding. *J. Chem. Phys.* 136, 044104. doi:10.1063/1.3675845

Hofmann, D., and Kümmel, S. (2012). Self-interaction Correction in a Real-Time Kohn-Sham Scheme: Access to Difficult Excitations in Time-dependent Density Functional Theory. *J. Chem. Phys.* 137, 064117. doi:10.1063/1.4742763

Jacob, C. R., Beyhan, S. M., Bulo, R. E., Gomes, A. S. P., Götz, A. W., Kiewisch, K., et al. (2011). PyADF - a Scripting Framework for Multiscale Quantum Chemistry. *J. Comput. Chem.* 32, 2328–2338. doi:10.1002/jcc.21810

Jacob, C. R., and Neugebauer, J. (2014). Subsystem Density-Functional Theory. *Wires Comput. Mol. Sci.* 4, 325–362. doi:10.1002/wcms.1175

Jacob, C. R., Neugebauer, J., and Visscher, L. (2008). A Flexible Implementation of Frozen-Density Embedding for Use in Multilevel Simulations. *J. Comput. Chem.* 29, 1011–1018. doi:10.1002/jcc.20861

J. A. V. Bokhoven and C. Lamberti (Editors) (2016). *X-Ray Absorption and X-Ray Emission Spectroscopy* (John Wiley & Sons). doi:10.1002/9781118844243

Jiemchooroj, A., Ekström, U., and Norman, P. (2007). Near-edge X-ray Absorption and Natural Circular Dichroism Spectra of L-Alanine: A Theoretical Study Based on the Complex Polarization Propagator Approach. *J. Chem. Phys.* 127, 165104. doi:10.1063/1.2800024

Jornet-Somoza, J., and Lebedeva, I. (2019). Real-time Propagation Tddft and Density Analysis for Exciton Coupling Calculations in Large Systems. *J. Chem. Theor. Comput.* 15, 3743–3754. doi:10.1021/acs.jctc.9b00209

Kadek, M., Konecny, L., Gao, B., Repisky, M., and Ruud, K. (2015). X-ray Absorption Resonances Near L2,3-Edges from Real-Time Propagation of the Dirac-Kohn-Sham Density Matrix. *Phys. Chem. Chem. Phys.* 17, 22566–22570. doi:10.1039/C5CP03712C

Keldysh, L. V. (2017). Multiphoton Ionization by a Very Short Pulse. *Phys.-Usp.* 60, 1187–1193. doi:10.3367/ufne.2017.10.038229

Kendall, R. A., Dunning, T. H., Jr, and Harrison, R. J. (1992). Electron Affinities of the First-row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* 96, 6796–6806. doi:10.1063/1.462569

Koh, K. J., Nguyen-Beck, T. S., and Parkhill, J. (2017). Accelerating Realtime TDDFT with Block-Orthogonalized Manby-Miller Embedding Theory. *J. Chem. Theor. Comput.* 13, 4173–4178. doi:10.1021/acs.jctc.7b00494

Kong, X., Waldner, A., Orlando, F., Artiglia, L., Huthwelker, T., Ammann, M., et al. (2017). Coexistence of Physisorbed and Solvated HCl at Warm Ice Surfaces. *J. Phys. Chem. Lett.* 8, 4757–4762. doi:10.1021/acs.jpclett.7b01573

König, C., and Neugebauer, J. (2013). Exciton Coupling Mechanisms Analyzed with Subsystem Tddft: Direct vs Pseudo Exchange Effects. *J. Phys. Chem. B* 117, 3480–3487. doi:10.1021/jp3105419

Krishtal, A., Ceresoli, D., and Pavanello, M. (2015). Subsystem Real-Time Time Dependent Density Functional Theory. *J. Chem. Phys.* 142, 154116. doi:10.1063/1.4918276

Krishtal, A., and Pavanello, M. (2016). Revealing Electronic Open Quantum Systems with Subsystem TDDFT. *J. Chem. Phys.* 144, 124118. doi:10.1063/1.4944526

Kumar, S. P., Genova, A., and Pavanello, M. (2017). Cooperation and Environment Characterize the Low-Lying Optical Spectrum of Liquid Water. *J. Phys. Chem. Lett.* 8, 5077–5083. doi:10.1021/acs.jpclett.7b02212

Lee, S. J. R., Welborn, M., Manby, F. R., and Miller, T. F. (2019). Projection-based Wavefunction-In-DFT Embedding. *Acc. Chem. Res.* 52, 1359–1368. doi:10.1021/acs.accounts.8b00672

Lembarki, A., and Chermette, H. (1994). Obtaining a Gradient-Corrected Kinetic-Energy Functional from the Perdew-Wang Exchange Functional. *Phys. Rev. A.* 50, 5328–5331. doi:10.1103/PhysRevA.50.5328

Li, X., Govind, N., Isborn, C., DePrince, A. E., and Lopata, K. (2020). Real-time Time-dependent Electronic Structure Theory. *Chem. Rev.* 120, 9951–9993. doi:10.1021/acs.chemrev.0c00223

Lipparini, F., and Mennucci, B. (2021). Hybrid Qm/classical Models: Methodological Advances and New Applications. *Chem. Phys. Rev.* 2, 041303. doi:10.1063/5.0064075

Lopata, K., and Govind, N. (2011). Modeling Fast Electron Dynamics with Real-Time Time-dependent Density Functional Theory: Application to Small Molecules and Chromophores. *J. Chem. Theor. Comput.* 7, 1344–1355. doi:10.1021/ct200137z

Lopata, K., Van Kuiken, B. E., Khalil, M., and Govind, N. (2012). Linear-response and Real-Time Time-dependent Density Functional Theory Studies of Core-Level Near-Edge X-ray Absorption. *J. Chem. Theor. Comput.* 8, 3284–3292. doi:10.1021/ct3005613

Manby, F. R., Stella, M., Goodpaster, J. D., and Miller, T. F. (2012). A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theor. Comput.* 8, 2564–2568. doi:10.1021/ct300544e

Marques, M. A. L., López, X., Varsano, D., Castro, A., and Rubio, A. (2003). Time-dependent Density-Functional Approach for Biological Chromophores: The Case of the green Fluorescent Protein. *Phys. Rev. Lett.* 90, 258101. doi:10.1103/physrevlett.90.258101

Mokkath, J. H. (2020). Photo-induced Electron Transfer between a Metal Nanoparticle and a Collection of Molecular Emitters. *Chem. Phys. Lett.* 758, 137905. doi:10.1016/j.cplett.2020.137905

Morzan, U. N., Ramírez, F. F., Oviedo, M. B., Sánchez, C. G., Scherlis, D. A., and Lebrero, M. C. G. (2014). Electron Dynamics in Complex Environments with Real-Time Time Dependent Density Functional Theory in a QM-MM Framework. *J. Chem. Phys.* 140, 164105. doi:10.1063/1.4871688

Neugebauer, J. (2007). Couplings between Electronic Transitions in a Subsystem Formulation of Time-dependent Density Functional Theory. *J. Chem. Phys.* 126, 134116. doi:10.1063/1.2713754

Neugebauer, J. (2009). On the Calculation of General Response Properties in Subsystem Density Functional Theory. *J. Chem. Phys.* 131, 084104. doi:10.1063/1.3212883

Neugebauer, J. (2009). Subsystem-based Theoretical Spectroscopy of Biomolecules and Biomolecular Assemblies. *ChemPhysChem* 10, 3148–3173. doi:10.1002/cphc.200900538

Niemeyer, N., Tölle, J., and Neugebauer, J. (2020). Approximate versus Exact Embedding for Chiroptical Properties: Reconsidering Failures in Potential and Response. *J. Chem. Theor. Comput.* 16, 3104–3120. doi:10.1021/acs.jctc.0c00125

Norman, P., and Dreuw, A. (2018). Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* 118, 7208–7248. doi:10.1021/acs.chemrev.8b00156

Olejniczak, M., Bast, R., and Gomes, A. S. P. (2017). On the Calculation of Second-Order Magnetic Properties Using Subsystem Approaches in a Relativistic Framework. *Phys. Chem. Chem. Phys.* 19, 8400–8415. doi:10.1039/C6CP08561J

Parise, A., Alvarez-Ibarra, A., Wu, X., Zhao, X., Pilmé, J., and de la Lande, A. (2018). Quantum Chemical Topology of the Electron Localization Function in the Field of Attosecond Electron Dynamics. *J. Phys. Chem. Lett.* 9, 844–850. doi:10.1021/acs.jpclett.7b03379

Pavanello, M. (2013). On the Subsystem Formulation of Linear-Response Time-dependent Dft. *J. Chem. Phys.* 138, 204118. doi:10.1063/1.4807059

Pedersen, M. N., Hedegård, E. D., Olsen, J. M. H., Kauczor, J., Norman, P., and Kongsted, J. (2014). Damped Response Theory in Combination with

Polarizable Environments: The Polarizable Embedding Complex Polarization Propagator Method. *J. Chem. Theor. Comput.* 10, 1164–1171. doi:10.1021/ct400946k

Pillar-Little, E. A., Guzman, M. I., and Rodriguez, J. M. (2013). Conversion of Iodide to Hypoiodous Acid and Iodine in Aqueous Microdroplets Exposed to Ozone. *Environ. Sci. Technol.* 47, 10971–10979. doi:10.1021/es401700h

Pipolo, S., and Corni, S. (2016). Real-time Description of the Electronic Dynamics for a Molecule Close to a Plasmonic Nanoparticle. *J. Phys. Chem. C* 120, 28774–28781. doi:10.1021/acs.jpcc.6b11084

Repisky, M., Komorovsky, S., Kadek, M., Konecny, L., Ekström, U., Malkin, E., et al. (2020). ReSpect: Relativistic Spectroscopy DFT Program Package. *J. Chem. Phys.* 152, 184101. doi:10.1063/5.0005094

Repisky, M., Konecny, L., Kadek, M., Komorovsky, S., Malkin, O. L., Malkin, V. G., et al. (2015). Excitation Energies from Real-Time Propagation of the Four-Component Dirac-Kohn-Sham Equation. *J. Chem. Theor. Comput.* 11, 980–991. doi:10.1021/ct501078d

Rinkevicius, Z., Sandberg, J. A. R., Li, X., Linares, M., Norman, P., and Ågren, H. (2016). Hybrid Complex Polarization Propagator/molecular Mechanics Method for Heterogeneous Environments. *J. Chem. Theor. Comput.* 12, 2661–2667. doi:10.1021/acs.jctc.6b00255

Rossi, T. P., Kuisma, M., Puska, M. J., Nieminen, R. M., and Erhart, P. (2017). Kohn-sham Decomposition in Real-Time Time-dependent Density-Functional Theory: An Efficient Tool for Analyzing Plasmonic Excitations. *J. Chem. Theor. Comput.* 13, 4779–4790. doi:10.1021/acs.jctc.7b00589

Schelter, I., and Kümmel, S. (2018). Accurate Evaluation of Real-Time Density Functional Theory Providing Access to Challenging Electron Dynamics. *J. Chem. Theor. Comput.* 14, 1910–1927. doi:10.1021/acs.jctc.7b01013

Scholz, L., Tölle, J., and Neugebauer, J. (2020). Analysis of Environment Response Effects on Excitation Energies within Subsystem-based Time-dependent Density-functional Theory. *Int. J. Quan. Chem.* 120, e26213. doi:10.1002/qua.26213

Simpson, W. R., Brown, S. S., Saiz-Lopez, A., Thornton, J. A., and von Glasow, R. (2015). Tropospheric Halogen Chemistry: Sources, Cycling, and Impacts. *Chem. Rev.* 115, 4035–4062. doi:10.1021/cr5006638

Sinha-Roy, R., García-González, P., López Lozano, X., Whetten, R. L., and Weissker, H.-C. (2018). Identifying Electronic Modes by Fourier Transform from δ-Kick Time-Evolution TDDFT Calculations. *J. Chem. Theor. Comput.* 14, 6417–6426. doi:10.1021/acs.jctc.8b00750

Smith, D. G. A., Burns, L. A., Simmonett, A. C., Parrish, R. M., Schieber, M. C., Galvelis, R., et al. (2020). Psi4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. *J. Chem. Phys.* 152, 184108. doi:10.1063/5.0006002

Smith, D. G. A., Burns, L. A., Sirianni, D. A., Nascimento, D. R., Kumar, A., James, A. M., et al. (2018). Psi4numpy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theor. Comput.* 14, 3504–3511. doi:10.1021/acs.jctc.8b00286

Stener, M., Fronzoni, G., and de Simone, M. (2003). Time Dependent Density Functional Theory of Core Electrons Excitations. *Chem. Phys. Lett.* 373, 115–123. doi:10.1016/s0009-2614(03)00543-8

Sun, Q., and Chan, G. K.-L. (2016). Quantum Embedding Theories. *Acc. Chem. Res.* 49, 2705–2712. doi:10.1021/acs.accounts.6b00356

te Velde, G., Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G., et al. (2001). Chemistry with Adf. *J. Comput. Chem.* 22, 931–967. doi:10.1002/jcc.1056

Thomas, L. H. (1927). The Calculation of Atomic fields. *Math. Proc. Camb. Phil. Soc.* 23, 542–548. doi:10.1017/S0305004100011683

Tölle, J., Böckers, M., and Neugebauer, J. (2019). Exact Subsystem Time-dependent Density-Functional Theory. *J. Chem. Phys.* 150, 181101. doi:10.1063/1.5097124

Tölle, J., Böckers, M., Niemeyer, N., and Neugebauer, J. (2019). Inter-subsystem Charge-Transfer Excitations in Exact Subsystem Time-dependent Density-Functional Theory. *J. Chem. Phys.* 151, 174109. doi:10.1063/1.5121908

Valiev, M., Bylaska, E. J., Govind, N., Kowalski, K., Straatsma, T. P., Van Dam, H. J. J., et al. (2010). NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Computer Phys. Commun.* 181, 1477–1489. doi:10.1016/j.cpc.2010.04.018

Van Lenthe, E., and Baerends, E. J. (2003). Optimized Slater-type Basis Sets for the Elements 1-118. *J. Comput. Chem.* 24, 1142–1156. doi:10.1002/jcc.10255

Villaume, S., Saue, T., and Norman, P. (2010). Linear Complex Polarization Propagator in a Four-Component Kohn-Sham Framework. *J. Chem. Phys.* 133, 064105. doi:10.1063/1.3461163

Wesolowski, T. A., Shedge, S., and Zhou, X. (2015). Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* 115, 5891–5928. doi:10.1021/cr500502v

Wesolowski, T. A., and Warshel, A. (1993). Frozen Density Functional Approach for Ab Initio Calculations of Solvated Molecules. *J. Phys. Chem.* 97, 8050–8053. doi:10.1021/j100132a040

Woon, D. E., and Dunning, T. H., Jr (1994). Gaussian Basis Sets for Use in Correlated Molecular Calculations. Iv. Calculation of Static Electrical Response Properties. *J. Chem. Phys.* 100, 2975–2988. doi:10.1063/1.466439

Wu, X., Teuler, J.-M., Cailliez, F., Clavaguéra, C., Salahub, D. R., and de la Lande, A. (2017). Simulating Electron Dynamics in Polarizable Environments. *J. Chem. Theor. Comput.* 13, 3985–4002. doi:10.1021/acs.jctc.7b00251

Yu, Z., and Li, Y. (2021). Marine Volatile Organic Compounds and Their Impacts on marine Aerosol-A Review. *Sci. Total Environ.* 768, 145054. doi:10.1016/j.scitotenv.2021.145054

Zhang, Y., Biggs, J. D., Healion, D., Govind, N., and Mukamel, S. (2012). Core and Valence Excitations in Resonant X-ray Spectroscopy Using Restricted Excitation Window Time-dependent Density Functional Theory. *J. Chem. Phys.* 137, 194306. doi:10.1063/1.4766356

Zimmermann, P., Peredkov, S., Abdala, P. M., DeBeer, S., Tromp, M., Müller, C., et al. (2020). Modern X-ray Spectroscopy: XAS and XES in the Laboratory. *Coord. Chem. Rev.* 423, 213466. doi:10.1016/j.ccr.2020.213466

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Chemistry

ORIGINAL RESEARCH
published: 07 April 2022
doi: 10.3389/fchem.2022.858946

Reassessing the Role of σ Holes in Noncovalent Interactions: It is Pauli Repulsion that Counts

Małgorzata M. Szczęśniak[1]* *and Grzegorz Chałasinski*[2]*

[1]Department of Chemistry, Oakland University, Rochester, MI, United States, [2]Faculty of Chemistry, Warsaw University, Warszawa, Poland

OPEN ACCESS

Edited by:
Piotr Żuchowski,
Nicolaus Copernicus University in Toruń, Poland

Reviewed by:
Fangfang Pan,
Central China Normal University, China
Cina Foroutan-Nejad,
Institute of Organic Chemistry (PAN), Poland

***Correspondence:**
Małgorzata M. Szczęśniak
bryant@oakland.edu
Grzegorz Chałasinski
chalbie@tiger.chem.uw.edu.pl

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: 20 January 2022
Accepted: 16 February 2022
Published: 07 April 2022

Citation:
Szczęśniak MM and Chałasinski G (2022) Reassessing the Role of σ Holes in Noncovalent Interactions: It is Pauli Repulsion that Counts. Front. Chem. 10:858946. doi: 10.3389/fchem.2022.858946

A number of prototypical weak electron donor–electron acceptor complexes are investigated by the Symmetry Adapted Perturbation Theory, some of which belong to novel classes of weak bonds such as halogen and chalcogen bonds. Also included are complexes involving strong Lewis acids such as BeO and AuF. The common view in the literature is to associate these novel bonds with a variety of "holes", σ, π, δ, or positive areas in their electrostatic potential maps. The presumption is that these positive areas of the electrostatic potential are indicative of the electrostatic nature of these noncovalent bonds. The electrostatic view extends to the explanations of the directionality of approaches between the subsystems forming these bonds. This work demonstrates that one common feature of these electrostatic potential "holes" is the local depletion of electron density of which the best detector is the first-order Pauli repulsion. The minimization of this repulsion determines the bond directionality and its relative angular rigidity. In relatively strong complexes of BeO with rare gases, where BeO shows a clear cavity in electron density—an ultimate "σ hole"—the electrostatic effect does not control the bending potential—the exchange repulsion does. In halogen bonds, the halogen atom is nonspherical, displaying an axial "σ hole" in its electrostatic potential. However, in no examined case, from rare gas acting as an electron donor to a polar donor to an anionic donor, is the electrostatic energy responsible for the directionality of the halogen bond. In fact, it is not even maximized in the direction of the σ hole in N_2-ClF and NH_3-ClF. Yet, in all the cases, the exchange repulsion is minimized in the direction of the σ hole. The minimized exchange repulsion associated with the subtle and less subtle depletions of the electron density occur on the nodal planes or on the intersections thereof in the highest occupied molecular orbitals of Lewis acids, provided that the systems are closed-shell. The role of nodal planes in covalent and coordinate covalent bonds is well recognized. This work points to their similarly equal importance in certain types of donor–acceptor noncovalent interactions.

Keywords: sigma holes, Pauli repulsion, noncovalent interactions, donor–acceptor interactions, SAPT

INTRODUCTION

The weak donor–acceptor interactions encompass a wide swath of noncovalent interactions from noble-gas molecule to halogen-bonded to metalophilic. New types are being discovered by examinations of crystal structures (Ramasubbu et al., 1986; Bauzá et al., 2013; Saha and Desiraju, 2017; Saha et al., 2018) while new names are added to the lexicons [for the latest list see (Zierkiewicz et al., 2021)]. One common feature of these interactions is the presence of broadly defined Lewis acid as an electron acceptor from an electron-rich Lewis base. The sources of electron deficiency may vary from the incomplete octet (as in BF_3), low-lying LUMO [as in typical CT complex of Benesi and Hildebrand—benzene-I_2 (Benesi and Hildebrand, 1948)], lower-lying T_{2g} ligand positions in O_h transition metal complexes, electron deficiency due to the positive quadrupole moment (as in dihalogens), to electropositive regions discussed by Legon (Legon, 1999), which morphed into ubiquitous "σ holes" (as in many types of novel bonds from halogen to tetrel) or actual cavities in electron distribution (as in BeO). In this work, we argue that these sources of electron deficiency have a common origin. For historical perspective, the terms σ hole initially referred to empty or half-empty σ-symmetry orbitals (Koch et al., 1987).

In today's popular meaning of the role of "σ holes," one refers to the map of electrostatic potential on a threshold density surface and to the locus of positive potential that is used to explain both the bonding and directionality of the interactions with a Lewis base. This explanation echoes the way that the positive partial charge on protons was used to explain the directionality of hydrogen bonding. Consequently, one may conclude that all the noncovalent interactions are driven by electrostatics. In fact, some authors do claim that everything *is* electrostatics (Politzer et al., 2015) and the rest is "*noumenon*"—the Kantian term introduced by Kozuch and Martin to describe unobservable explanations (Kozuch and Martin, 2013).

This simplified understanding does not explain why these noncovalent bonds are rigid. Stone in a recent paper entitled "Are halogen bonds electrostatically driven?" shows that for one selected Lewis acid, ClF, with a variety of bases, the directionality is determined by the exchange repulsion (Stone, 2013). Arguably, the first notice that the weak interactions with halogens were unusually rigid, i.e., have unusually high bending constants, was that by Klemperer and coworkers (Harris et al., 1974) in Ar-ClF. The first explanation of this phenomenon by Sadlej et al. (1993) centered on the pronounced dip in the exchange repulsion in the linear geometry that determines the shape of the potential curve. The dip in the repulsive term, as the work argues, allows the subsystems to approach more closely and bind stronger by induction and dispersion attractive components (no need for electrostatic component!). The dip in the exchange repulsion implies a depletion of the electron density in this position. The electron density depletion may also manifest itself as positive value of the electrostatic potential at the density map in this region. At this point, one may ask oneself what is the more sensitive and foolproof measure of the shape of electron density and the Lewis acid capability: the electrostatic potential, i.e., the interaction of a proton (positive point charge) with the charge distribution of a molecule, or the first-order exchange repulsion between the same charge distribution and a probe—such as, for example, a two-electron atom—the simplest Lewis base?

There exists an entire cottage industry around the analysis of the density shapes in terms of its gradient and Laplacian (Bader, 1994). This analysis is best suited for the detection of bond covalency and is not going to be employed here. The present work will focus on the role of Pauli exchange repulsion in probing the shapes of density in the sites of Lewis acidity and its consequences to the directionality of weak donor–acceptor interactions.

METHODS

The interaction energies between probing systems and Lewis acids are described in terms of the Symmetry Adapted Perturbation Theory (SAPT) (Jeziorski et al., 1994). The interaction energy is defined as a sum of electrostatics (es), exchange (exch), induction (ind), exchange-induction (ex-ind), dispersion (disp), and its exchange counterpart (ex-disp),

$$E_{SAPT0} = E_{es}^{(10)} + E_{exch}^{(10)} + E_{ind,r}^{(20)} + E_{ex-ind,r}^{(20)} + E_{disp}^{(20)} + E_{ex-disp}^{(20)} + \delta_{HF} \tag{1a}$$

$$E_{SAPT2+} = E_{SAPT0} + E_{es,r}^{(12)} + E_{exch}^{(11)} + E_{exch}^{(12)} + E_{ind}^{(22)} + E_{ex-ind}^{(22)} + E_{disp}^{(21)} + E_{disp}^{(22)} \tag{1b}$$

where "r" in the subscripts refers to the response-inclusive treatment of induction and electrostatics. The numbers in the superscripts in the $E^{(VW)}$ notation correspond to the orders in the intermolecular perturbation operator V and the intramonomer correlation operator W. The last term δ_{HF} stands for higher-order induction and/or residual effects. Two levels of SAPT are employed: SAPT0 means that the interaction energy terms are derived from the product of Hartree-Fock (HF) subsystem wave functions. We use terminology employed in the Psi4 program (Parrish et al., 2017; Hohenstein et al., 2011). In SAPT0, the terms on the right in **Eq. 1a** sum up to the HF + Dispersion level of theory. A higher level of theory is obtained if the monomer wave functions are correlated to the 2nd order in the intramonomer correlation operator W. This results in SAPT2. The level designated SAPT2+ (Parker et al., 2014) means that the treatment of the dispersion energy is enhanced up to the inclusion of triples (although the exchange counterpart of dispersion is not correlated). In the graphs below when SAPT2+ is employed, we fold the intramonomer correlation contributions into the respective terms and combine the induction with exchange (denoting it $E_{ind\text{-}PT}$), as well as the dispersion with its exchange counterpart. The residual term δ_{HF} is shown separately for the reasons described below. The monomer geometries were taken from experiment (Huber et al., 2022) and the intersystem distances were either at the sum of van der Waals radii or at their equilibrium, if known

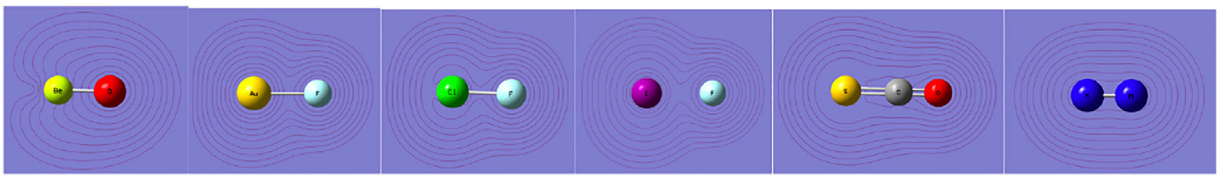


FIGURE 1 | Density contours (0.001 au) of BeO, AuF, Cl-F, I-F, SCO, and N_2 molecules from TPSS/AVQZ-derived densities.

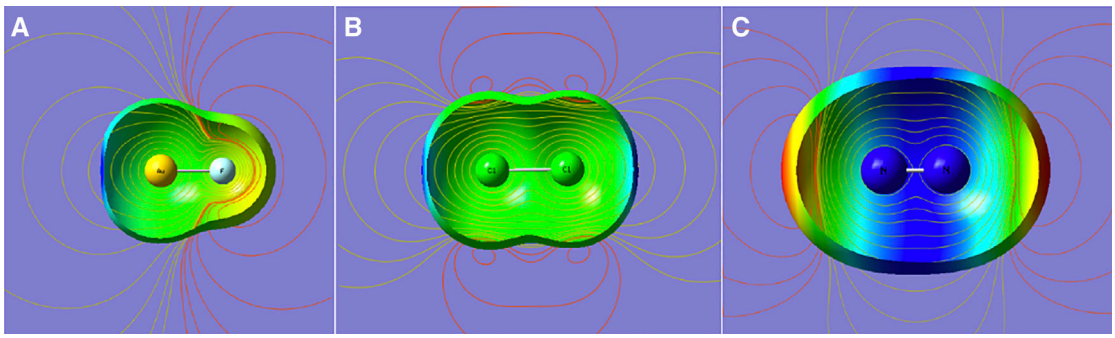


FIGURE 2 | Density (0.001 au), electrostatic potential mapped on the density surface (blue for positive and red for negative), and electrostatic potential contours (yellow for positive; red for negative). From **(A–C)**: polar molecule (AuF), nonpolar molecule with positive quadrupole (Cl_2), and with negative quadrupole (N_2).

(Evans et al., 2010; Hapka et al., 2013). The SAPT computations were performed using the Psi4 (Parrish et al., 2017) program suite, which employs the density fitting and Laplace transform for the evaluation of energy components (Hohenstein et al., 2011). The basis set was primarily aug-cc-pVQZ (Pritchard et al., 2019) except for the heavier atoms, I and Au, for which the def2-QZVPPD basis set was used (Pritchard et al., 2019) (this choice was dictated by the existence of density fitting basis sets in SAPT).

RESULTS AND DISCUSSION

Convex, Concave, or Polar-Flattened?

We begin with a look at the shapes of electron density—a simple but crude instrument. We focus on density around a Lewis acidic site. **Figure 1** shows the isodensity plots for a number of molecules discussed in this work. Going from the left to right, the density contours show the deep depletion in BeO, still visible depletion in AuF, a slight depletion in ClF, especially in a short region of Cl, an axial flattening of density in I-F around the iodine, convex contours in SCO around sulfur with some axial flattening, and finally the convex contours in N_2 with no flattening. The axial flattening, also called polar flattening in literature (Sedlak et al., 2015), does not occur in N_2; rather, its density protrudes axially in the lone pair region. In fact, the quadrupole moment of N_2 is negative, thus indicating that the negative charge extends further along the molecular axis.

Now, how is this picture reflected in the electrostatic potential mapped on the threshold density?

Three limited cases are presented in **Figure 2**: a dipolar and two quadrupolar cases. It shows the contours of the electrostatic potential that emanate from a dipole and two quadrupoles of opposite signs. One may conclude that the positive "σ hole" in the left two and the negative "σ bump" offer no unique insights except for pointing to the sign of the leading multipole moment of the molecule. Arguably, when the density surface shows an axial indentation, as in AuF in **Figure 2**, the magnitude of electrostatic potential at this point is enhanced. Nevertheless, the insights from this picture are qualitative at best.

Recent years have seen an explosive growth (Web of Science Core Collection, 2022) of literature on the features of electrostatic potential painted on the density (Cavallo et al., 2016; Kolář and Hobza, 2016). The types of holes that emerged from these studies range from σ to π to δ (Angarov and Kozuch, 2018). Simultaneously, the names of bonds associated with these "holes" exploded up to 8 by the last count (Zierkiewicz et al., 2021). The "hole" properties such as the spatial extent of the positive region (Kolář et al., 2015), the value of the maximal electrostatic potential, V_{smax}, at a certain threshold density (Politzer et al., 2013), and the extent of the hole's "polar flattening" (Sedlak et al., 2015) have been correlated with the bonding strength of a complex.Simultaneously, there were claims that only the density and its associated electrostatics represent "phenomena" and all the other computable and physically interpretable effects represent "unicorns" (Politzer et al., 2015). This literature avalanche gives an impression that all of the donor–acceptor interactions are electrostatically controlled, which is incomplete if not misleading. There are some

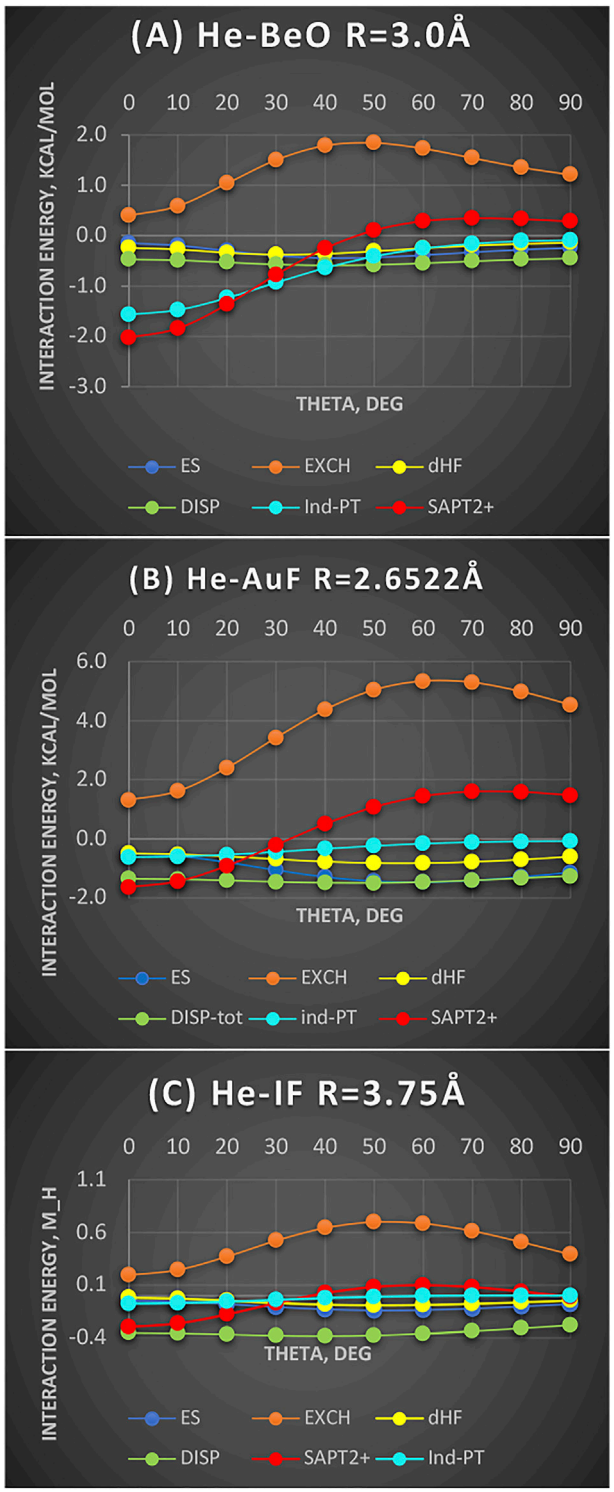


FIGURE 3 | Interaction energy terms for angular displacement in Jacobi coordinates: **(A)** He-BeO R = 3.0 Å (units kcal/mol), **(B)** He-AuF R = 2.6522Å (units kcal/mol), and **(C)** He-IF R = 3.75 Å (units mH).

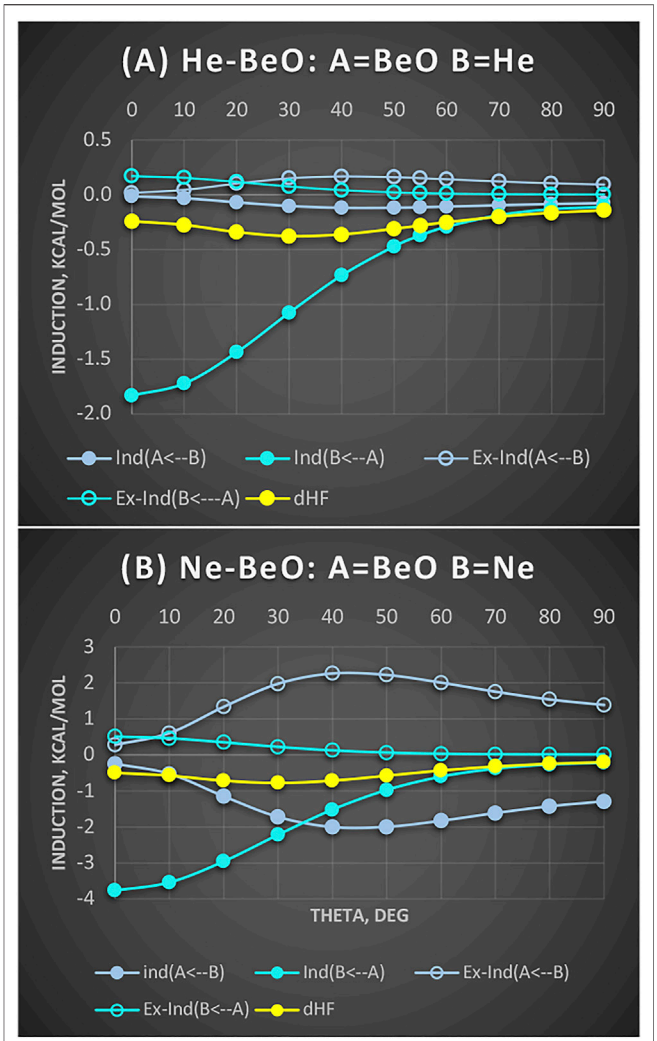


FIGURE 4 | He-BeO **(A)** and Ne-BeO **(B)** at R = 3.0 Å: Partitioning of induction terms into mutual contributions of A = BeO and B=He/Ne. Solid circles are induction terms and open circles are exchange-induction terms. Values in kcal/mol.

dissenting voices from this chorus stressing the importance of the exchange effect (Adhikari and Scheiner, 2012; Stone, 2013), the ground-state charge-transfer interactions (Wang et al., 2014; Rezac and de la Lande, 2017; Thirman et al., 2018), orbital interactions (Pascoe et al., 2017), or covalency contributions (Bora et al., 2017).

The electrostatic effect and "σ holes", which are expected to explain the directionality of these interactions, are purely classical concepts. To fully deconstruct the properties of the "hole" and to detect the existence of density depletions in the direction of the "hole", we need a more sophisticated tool—a quantum effect of exchange. This requires the participation of at least a two-electron interacting partner.

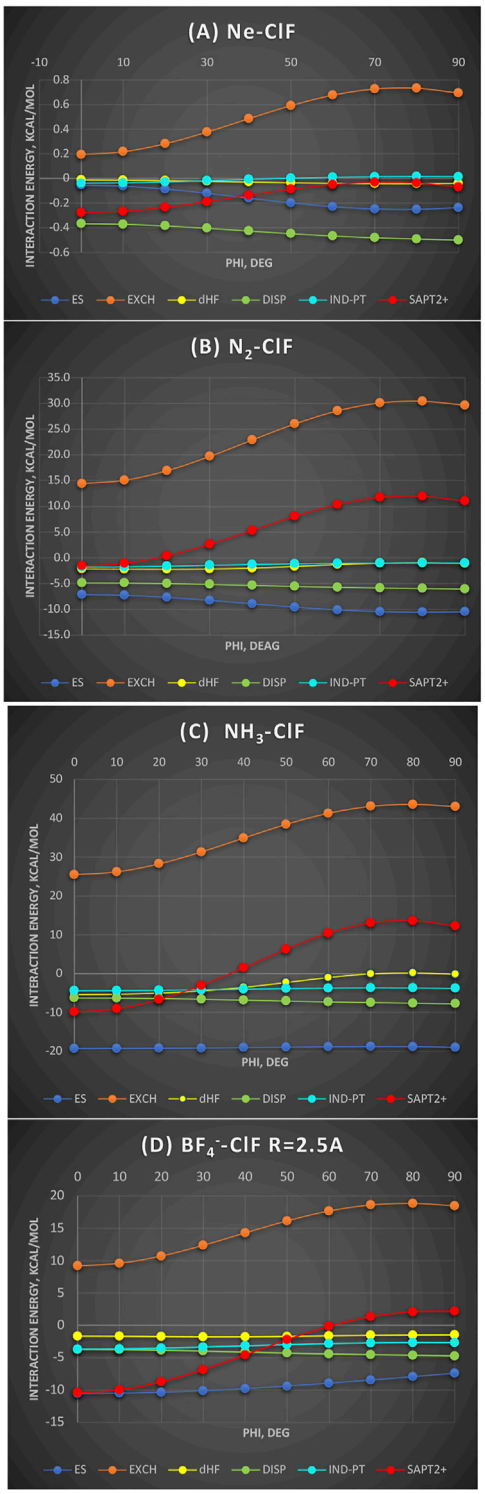


FIGURE 5 | SAPT energy components as a function of ClF bending around the Cl atom (denoted phi); values in kcal/mol. **(A)** Ne-ClF at the sum of van der Waals radii distance 3.3 Å; **(B)** N_2-ClF at the N-Cl distance 2.5 Å; **(C)** NH_3-ClF at the N-Cl distance 2.5 Å; **(D)** BF_4^--ClF at the intersystem F^--Cl distance 2.5 Å.

He-Molecule Interactions

Figure 3 shows three complexes, He-BeO, He-AuF, and He-IF. In the first complex, the density hole is evident and so deep that it can accommodate a He atom but not Ne. In the second one, the density still has some depletion but to a much lesser degree. In the third one, only a flattening is seen (**Figure 1**). We will show that regardless of the depth of the density shape, the same pattern persists if the energy components are considered. The systems are also chosen by steeply decreasing dipole moments from 6.6D (BeO) to 4.5D (AuF) to 2.7D (IF) (as per CCSD computation). Despite these widely varying electrical properties, the nature of bending potentials remains the same. It is the exchange repulsion term that displays a deep "hole" for the linear configuration and thus dictates the shape of the overall interaction energy (shown as SAPT2+ in **Figure 3**). The electrostatic term favors nonlinear geometries and disfavors the linear ones. It is of no importance since it is of the overlap character. The induction term's importance reflects the size of the dipole moment: the effect is maximized in He-BeO and minimized in He-IF. The dispersion energy is an important attractive contributor but its angular variations are minimal.

The same behavior could be expected for the radial coordinate: Namely, the depletion of the exchange repulsion allows the subsystems to approach more closely and benefit from enhanced dispersion and induction effects.

The *cavity in the interaction potential* of He-BeO was first demonstrated in the work of Hapka et al. (2013) in the form of zero-interaction energy boundary showing a deep indentation for the linear geometry on the Be side. They estimated the binding energy of nearly 1,900 cm^{-1} or 5.4 kcal/mol. Ne is similarly bound to BeO but penetrates less into the cavity by 0.267 Å according to Nunzi et al. (2017). The Ar-BeO study estimated the well depth of 3,990 cm^{-1} or 11.4 kcal/mol (Tebai et al., 2014), quite unusual for the rare gas binding to a molecule. The SAPT conclusion is that the minimum in He-BeO is largely due to the exchange depletion reinforced by the induction effect. The work of (Nunzi et al., 2017) goes even further by postulating that the unique strength of He-BeO as compared to Ne-BeO, for example, is due to the ability of BeO to back-donate to the vacant 2p orbitals of He—and the inability to do so to Ne because its 2p orbitals are already occupied. Our observation is that (1) Ne is simply too large to penetrate as deep into the density depletion in BeO as He could and (2) the induction effect is actually stronger in Ne than in He.

We examine this intriguing hypothesis of back-transfer below (**Figure 4**). First, one expects that if there is an overlap between the occupied π MOs of BeO with vacant 2p orbitals of He, it should favor the linear configurations. Second, it should manifest itself in the Molecule → Atom induction contribution. These contributions (along with a residual δ_{HF}) are shown in **Figure 4** for both He and Ne complexes with BeO taken at the same distance R = 3.0 Å, which is the equilibrium for Ne-BeO.

It appears that BeO → He/Ne contributions are strongly attractive, favoring the linear structure. Their respective exchange counterparts cancel very small fractions of these attractions.

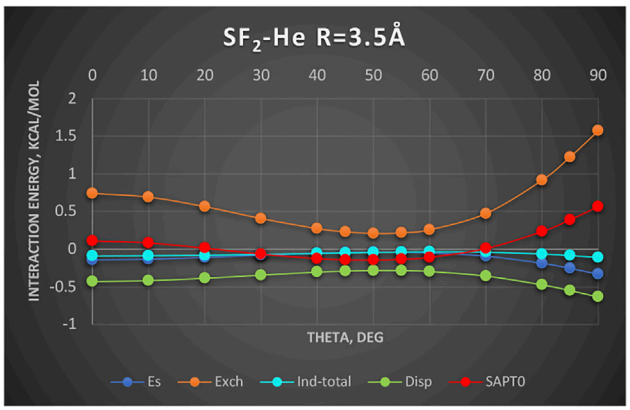


FIGURE 6 | SAPT0 energy components of SF_2-He as a function of theta. Complex is described in Jacobi coordinates: R = 3.5 Å and theta = 0° corresponds to He located at the C_2 axis on the S side. Values in kcal/mol.

The magnitudes of the induction terms of the BeO → He/Ne type are consistent with ca. 2:1 ratio of static dipole polarizabilities of Ne to He. In both cases, the BeO → He/Ne induction attraction strongly favors the linear structure. What if the donation to virtual 2p orbitals shows up in the residual term δ_{HF}? Here, in neither case is δ_{HF} the most attractive for the linear geometries. Thus, in the SAPT language, the only indication to support the BeO-to-He donation is in very strong BeO → He induction. One observation, however, is worth emphasizing: Namely, in the Ne-BeO complex, we can see very strong, common sense eluding Ne → BeO contribution of ca. 2 kcal/mol at skewed geometries. Thankfully, the exchange counterpart of this term varies as almost its mirror image and nearly completely wipes out the common-sense-defying effect. It is useful to interdict to explain why we use the term "common-sense-defying." This is because the induction effect allows electrons of one monomer to occupy the already occupied orbitals of the partner, thus violating the Pauli principle, if unconstrained by the exchange. The lesson from this is that employing the induction effect alone is error prone. The induction energy must be accompanied by its exchange counterpart, if unphysical effects are to be avoided.

When is the Electrostatic Effect Important?

To determine the effect of the nonsphericity of halogen on the shape of the interaction potential, we examine a series of electron donors (Lewis bases) with one Lewis acid ClF, the latter showing the prominent "σ hole," to establish in what circumstances is electrostatic effect the controlling one. This question was previously explored in the literature (Stone, 2013), but we intend to place it in a wider context.

In **Figure 5**, four Lewis bases are shown interacting with ClF, from very weak to ionic, as functions of the bending angle of ClF around the Cl atom (denoted φ) from linear to perpendicular. It is noteworthy that in neither case is the induction effect controlling the shape of the bending potential and neither is the dispersion energy nor the electrostatic. The Ne-ClF complex's (**Figure 5A**) total SAPT2+ interaction energy is the most binding for the linear geometry owing to the exchange repulsion's minimum of ca. 0.5 kcal/mol compared to the perpendicular arrangement. Indeed, this is the true manifestation of the hole. The electrostatic effect is overlap-dependent and varies in the opposite way. Two σ donors described in what follows, N_2 and NH_3, are considered at the same intersystem distance N---Cl of 2.5 Å, which equals the equilibrium separation in the latter. In the BF_4^--ClF complex, the same intersystem distance 2.5 Å is employed because it is close to the complex's equilibrium separation. In a non-polar N_2 interacting with ClF, the exchange hole amounts to 15 kcal/mol, dominating the total SAPT2+ energy. In fact, one curve appears to be the vertical translation of the other. The only other term that favors the linear geometry is δ_{HF} (see below for more discussion). In the polar NH_3 interacting with ClF, considered at the same intersystem distance, the exchange hole amounts now to 18 kcal/mol. The exchange effect again controls the shape of the bending potential as well as the well depth of this complex. Surprisingly, given the participation of the two polar subsystems, the electrostatic energy is almost constant with angular variation between −19.2 kcal/mol at 0° and −19.0 kcal/mol at 90°. Similarly, the induction effect (more specifically its perturbation theory part, $E_{Ind\text{-}PT}$) is also nearly constant (to 0.6 kcal/mol). To emphasize: the electrostatic and induction effects are *isotropic*, or nearly so, around the Cl end of the molecule. This appears to contradict the electrostatic predominance due to "σ hole" and, in particular, attributing any predictive capabilities to the value of V_{smax}. The second most anisotropic term is the residual δ_{HF} contributions to which we will return below.

Finally, in the ion-dipole complex BF_4^--ClF, one would expect that the electrostatic energy should drive the interaction. However, it is still the exchange hole of over 9 kcal/mol that shapes the potential curve. The electrostatic energy varies only by ca. 3 kcal/mol between 0° and 90°. All the other terms remain constant or nearly so (*cf.* dispersion). Naturally, at longer separations, one would expect the electrostatics to dominate; this is because the electrostatic effect varies more slowly with the intersystem distance than the exchange effect.

To summarize, with the variety of electron donors, ranging from a rare gas to quadrupolar to dipolar to ionic interacting with ClF, the bending potential around the equilibrium separation is controlled by the exchange hole. The effects of the electrostatic "σ hole" of ClF manifesting themselves in the electrostatic and induction interactions are only secondary.

Where Is the Charge-Transfer Contribution?

The complexes considered in the section *When is the Electrostatic Effect Important?* above were halogen-bonded, which are believed to be stabilized also by the charge-transfer contribution, as predicted by the Mulliken's interpretation of the complexes of this type (Mulliken and Person, 1969) (before they became known as halogen-bonded); although the subject remains controversial. There have been several schools of thought on inclusion of charge-transfer energies in the interaction energies beginning with Morokuma partitioning of the HF interaction energy in the early 70s (Morokuma, 1971; Kitaura and Morokuma, 1976), the Stone–Misquitta monomer-to-dimer

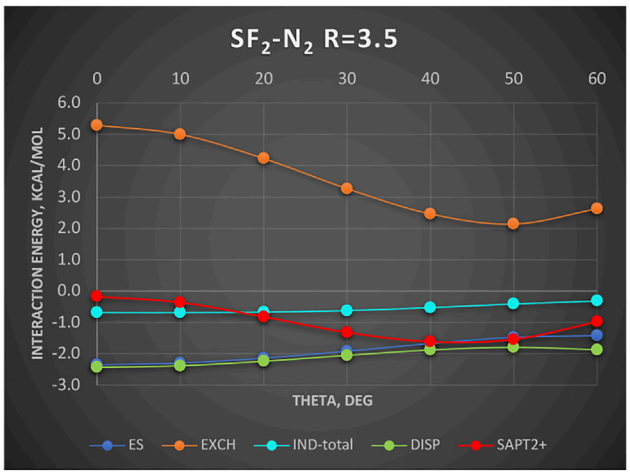


FIGURE 7 | SAPT2+ energy components of SF_2-N_2 as a function of theta. Complex is described in Jacobi coordinates: R = 3.5 Å and theta = 0° corresponds to N_2 oriented at the C_2 axis on the S side. Values in kcal/mol.

basis set extension of the induction effect in SAPT (Stone and Misquitta, 2009) [see also recommendation in Psi4 (Parrish et al., 2017)], Misquitta's attempt to "regularize" one-electron V operator terms (Misquitta, 2013), constraining intermonomer DFT and then relaxing it (Wu and Van Voorhis, 2005; Lao and Herbert, 2016), and the energy decomposition scheme ALMO-EDA in its latest variant (Thirman and Head-Gordon, 2015), just to name a few. Finally, there are also views that the charge transfer is fully included in the induction effect (Politzer et al., 2013; Sedlak et al., 2015).

In the halogen bond context, the charge-transfer contribution is understood as energy lowering resulting from the orbital interaction involving the lone pair of Lewis base (n) with the LUMO of halogen molecule σ^* of the type (Thirman et al., 2018)

$$E_{CT} = \frac{\langle n|\hat{H}|\sigma^*\rangle}{E_n - E_{\sigma^*}} \tag{2}$$

This interaction favors the direction where n and σ^* overlap best, i.e., the linear geometry. It also benefits from the denominator's energy rising of n and lowering σ^* resulting from the interaction. Since perturbation theory is based on the product of unperturbed monomers' wave functions, the latter effect will be absent. If one performs the full dimer HF energy calculations, the orbitals that are still located largely on the respective monomers will be relaxed. Therefore, the term δ_{HF}, which is obtained as

$$\delta_{HF} = \Delta E_{HF} - E_{es}^{(10)} - E_{exch}^{(10)} - E_{ind,r}^{(20)} - E_{ex-ind,r}^{(20)} \tag{3}$$

where ΔE_{HF} stands for the supermolecular interaction energy at the HF level, should capture such effects. A conventional view of δ_{HF} is that it sums up induction and exchange-induction terms of orders higher than the second. However, it also includes other residual effects including a possible divergence of SAPT due to the weak antisymmetry forcing.

In **Figure 5**, we purposely showed perturbation-theory-derived induction effects, $E_{ind\text{-}PT}$, as separate from δ_{HF} to test the hypothesis of δ_{HF} including the charge-transfer effects.

In Ne-ClF, δ_{HF} (**Figure 5A**) is near zero for the linear geometry and more attractive for the bent geometries. Upon closer inspection, we may notice that in the bent geometries, $E_{ind\text{-}PT}$ becomes positive—an unphysical behavior. Thus, δ_{HF} in this case corrects the SAPT's convergence. In BF_4^--ClF (**Figure 5D**), δ_{HF} is constant to a fraction of kcal/mol while $E_{ind\text{-}PT}$ varies by 1 kcal/mol in favor of the linear geometry. Therefore, δ_{HF} is expected to be of induction origin. We should contrast these two cases with what happens in NH_3-ClF. In this instance, when we sum up all the perturbational SAPT terms, the well depth is too shallow by half and δ_{HF} provides the missing half. δ_{HF} is strongly attractive for the linear structure and vanishes for the perpendicular one. This is the typical dependence that one would expect from the charge-transfer effect from n to σ*. In N_2-ClF, considered at the same distance, δ_{HF} binding exceeds that from $E_{ind\text{-}PT}$ and again is the most attractive at the linear geometry. However, it does not vanish at the perpendicular geometry indicating the presence of some residual effects in strongly repulsive part of the interaction SAPT2+ potential.

To conclude, δ_{HF} may contain, depending on the type of a complex, a variety of effects: residual induction and exchange in a weak one with a rare gas, Ne-ClF, or strong but with similarly nearly spherical donor BF_4^--ClF. In the paradigm "charge-transfer", halogen-bonded complex, NH_3-ClF, one-half of the binding energy in the minimum resides in δ_{HF}, i.e., the term whose angular dependence is consistent with the orbital interaction between n and σ*. Thus, we posit that when charge-transfer energy contributions are warranted (e.g., by the right symmetry and energetical balance of donor–acceptor properties), they would manifest themselves in SAPT in its residual term δ_{HF}.

Finally, we should mention that the use δ_{HF} is only compatible with SAPT based on HF subsystem zero-order wave functions because only in this case, as proved by Moszynski et al. (1996), do the induction terms summed up to the infinite order in V converge to the HF interaction energy (sans the first-order energy). As observed before (Modrzejewski et al., 2012), if SAPT is based on the DFT monomer description, the inclusion of δ_{HF}, which is common practice in the literature to improve results, is problematic.

A New Bond Type or the Exchange Hole?

A new and unique bond has been attributed to the 16-group containing atoms O, S, Se, etc., the chalcogen bond (Pascoe et al., 2017; Wang et al., 2009). In this section, we examine an example chalcogen-bond donor SF_2. The electrostatic potential of SF_2 as shown in Angarov and Kozuch (2018) features two positive regions on both sides of the C_2 axis roughly perpendicular to the S–F bonds. These positive regions are supposed to attract basic regions of a Lewis base. To determine what is really behind this attraction, we show in **Figure 6** the SAPT results for the complex SF_2-He. The Jacobi coordinate R = 3.5Å corresponds to the sum of van der Waals radii of He and S and also to the radial

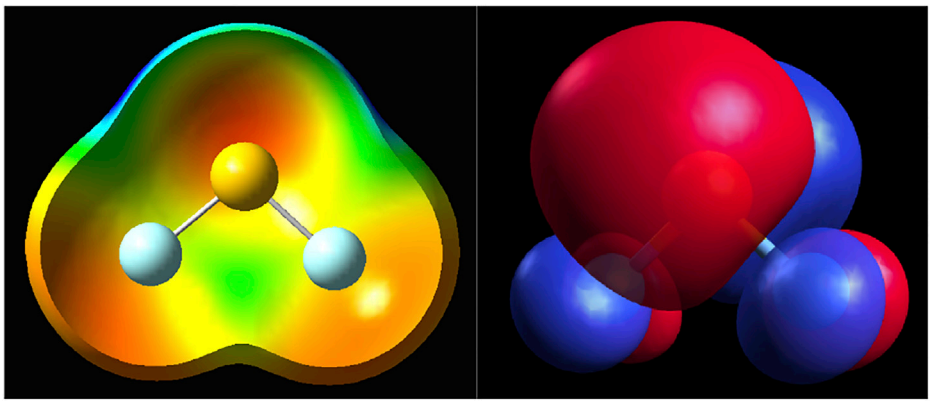


FIGURE 8 | SF_2: The electrostatic potential mapped on the density surface of 0.001 au threshold (left); HOMO orbital (right) from DFT TPSS calculations.

minimum in the SAPT2+ potential surface. The angular potential (here derived at the SAPT0 level of theory) is shallow with a weak minimum at theta = 50°. The minimum closely follows the angular variation of the exchange repulsion. Both the electrostatic and induction effects contribute minimally, especially in the minimum region. The only attractive term of some importance is the dispersion energy albeit the least attractive around the theta = 50° angle. Our two-electron probe thus indicates that the controlling effect is the exchange depletion—an exchange hole. The electrostatics and induction are unimportant. Would the same principles apply in the case of nonspherical Lewis base? The work of Angarov and Kozuch (2018) shows that when the Lewis base is a σ-donor HCN molecule, the angle of approach is also tilted ca. 53° away from the C_2 axis just as in the SF_2-He discussed above. Their results show only the total interaction. In a quick look at the actual balance of the energy components, we show in **Figure 7** the interaction with another σ-donor N_2: N_2-SF_2. The SAPT2+ curve again follows closely the exchange depletion at the tilted geometry with theta = 50°, just as in SF_2-He. The electrostatic energy favors the axial geometry (theta = 0°) and only coincidentally matches the value of total interaction energy around the minimum. The dispersion energy shows a similar pattern to SF_2-He: the most attractive at theta = 0 and the least attractive in the tilted geometry.

What these two cases have in common is that the Lewis base approaches the C_2 axis of SF_2 at a tilted angle and this happens regardless of the polarity state of the base. This may create an impression that the base is electrostatically attracted to the "σ hole" on the side of the sulfur atom (or, more accurately, above the S–F bond). In reality, this position is preferable due to the depletion in the exchange repulsion. Could this be viewed as a unique chalcogen bond? To answer this question let us recall a 30-year-old perturbation theory result of another chalcogen: H_2O-Ar (Chałasiński et al., 1991). In this complex, Ar is also tilted with respect to the C_2 axis of H_2O, the fact that was confirmed experimentally by the high-resolution vibration–rotation–tunneling spectroscopy (Cohen and Saykally, 1993). The perturbation-theory paper (Chalasinski et al., 1991) states that the angle is determined by the depletion in the exchange repulsion. Incidentally, the electrostatic potential of H_2O does not feature a sigma hole in this region. The σ holes, if we can call them such, reside on H atoms!

The Roots of the Exchange Holes

The final question concerns the cause of the exchange depletion in the systems studied here and in many other cases discussed in the vast literature of novel non-covalent bonds. The electron density depletion shown in **Figure 1** may be obvious in some molecules, such as BeO, but even in the halogen bond donors, the density depletions are not obvious from the density contours. The first-order exchange repulsion detects regions that are, to a varying degree, Pauli-forbidden and thus serve as a very sensitive probe of the doubly occupied space. The regions that are not occupied invariably involve nodes and nodal planes. It is, for example, axiomatic and taught in undergraduate courses that the nodal planes help distinguish bonding from antibonding orbitals or that the T_{2g} symmetry positions in the octahedral symmetry are more ligand-favorable than E_g because they lie on the intersections on the d_π orbitals of the transition metal. The same principle applies to the exchange holes of the systems studied here (and elsewhere). A good example is the case of SF_2 shown in **Figure 8**. On the left, there is the electrostatic potential mapped on the density, and on the right, the highest occupied molecular orbital (HOMO) of the molecule. The depletions seen above in the exchange repulsion (**Figure 7**) occur on the intersection of the nodal plane of the π orbital in plane of the molecule and another node that runs perpendicular to the S–F bond.

All Lewis acids examined in this work have the same features: the regions of depleted density occur on the intersections of nodal planes (the illustrations are included in the **Supplementary Material**). One easily explainable exception to this rule among the present systems involves AuF (**Supplementary Figure S2**). In AuF, as seen in **Figure 3**, the exchange depletion occurs in the axial position on the Au side whereas the HOMO orbital is d_σ along the molecular axis. However, due to the relativistic

contraction of the 6s orbital in Au, the axially extending part of s + d_{z2} hybrid is minimized. Close-lying d_π and d_δ occupied orbitals all feature one and two nodal planes, respectively, on the Au side of the molecule.

Molecular orbitals are, after all, one-electron functions and, to an extent, *noumena*. However, their symmetries have a deeper meaning as the irreducible representations of the nuclear point group. Our point here is that the depletion of the density on nodal planes has consequences to bonding whether covalent or noncovalent, indeed.

The circumstances of an actual electron vacancy as in a Cl radical, often mentioned as a protagonist of the halogen bond (Politzer et al., 2010; Kolář and Hobza, 2016), are entirely different. A ^{2}P state of Cl is triply degenerate and it takes an interaction with another system to remove this degeneracy. The Pauli repulsion along the singly occupied orbital with a closed-shell system will be smaller than along the doubly occupied one giving rise to multiple, coupled potential energy surfaces (Klos et al., 2004). If that system is a positive point charge, as in the case of the electrostatic potential, the interaction in the direction of singly occupied orbital will also be more repulsive as shown in (Politzer et al., 2010) for electrostatic reasons. The open-shell case brings us back to the original meaning of the electronic hole (Koch et al., 1987), which, by and large, is not governed by the nodal surfaces but by an actual electron vacancy. A transplantation of this concept to the closed-shell case of halogen bonding has been a gross simplification.

CONCLUSION

We investigated a number of prototypical weak electron donor–electron acceptor complexes by the Symmetry Adapted Perturbation Theory, some of which belong to novel classes of weak bonds such as halogen and chalcogen bonds. Also included were weak complexes involving "unnamed as yet bonds" of electron acceptors such as BeO and AuF. The recent literature trend is to associate these novel bonds with a variety of "holes", σ, π, δ, or positive areas in their electrostatic potential maps. The presumption is that these positive areas of the electrostatic potential are indicative of the electrostatic nature of these noncovalent bonds. This naïve electrostatic view extends to the explanations of the directionality of approaches between the subsystems forming these bonds. We show that one common feature of these electrostatic potential "holes" is the local depletion of electron density of which the best detector is the first-order exchange repulsion. The minimization of this repulsion determines the bond directionality and its relative angular rigidity. In relatively strong complexes of BeO with rare gases, where BeO shows a clear cavity in electron density—an ultimate "σ hole"—the electrostatic effect *is not the controlling* component—the exchange repulsion is. In halogen bonds, the halogen atom is nonspherical with an axial "σ hole" in its electrostatic potential, but in no case examined from rare gas to polar to anionic is the electrostatic energy responsible for the directionality of the halogen bond. In fact, it is not even maximized in the direction of the σ hole in N_2-ClF and NH_3-ClF! Yet, in all the cases, the exchange repulsion is minimized in the direction of the σ hole. The IUPAC definition (Desiraju et al., 2013) states that "forces involving the formation of the halogen bond are mainly electrostatic". There are some extra qualifiers to this statement adding polarization, charge transfer, and dispersion, but not a word about the exchange repulsion. It is hoped that the present work will shift this paradigm.

The charge-transfer energy components, which are somewhat controversial, even though the halogen bonds used to be called charge-transfer complexes, were examined here in the framework of SAPT. Although, by design, SAPT is not able to yield charge-transfer interaction as a well-defined energy term, despite some admirable attempts in literature (Misquitta, 2013; Stone and Misquitta, 2009), we postulate that if these effects are important, as in NH_3-ClF, they would reside in δ_{HF}—the residual term encompassing SAPT unaccounted for terms that occur in the well-defined supermolecule's energy. It is also important to stress that δ_{HF} is only strictly compatible with the SAPT based on the HF zero-order wave functions.

In the paradigm chalcogen-bond donor SF_2, the origins of the "hole" in the electrostatic potential [which, by some definitions (Angarov and Kozuch, 2018; Scheiner, 2021), would be called a π-hole] are exactly the same as in the σ hole case—the depletion of electron density on the intersection of nodal planes and consequent minimization of the exchange repulsion.

The minimized exchange repulsion associated with the subtle and less subtle depletions of the electron density occurs on the nodal planes or on the intersections of more than one nodal plane in the highest occupied molecular orbitals of a Lewis acid, provided that the systems are closed-shell. The role of nodal planes in covalent and coordinate covalent bonds is well recognized. This work points to their similarly equal importance in certain types of donor–acceptor noncovalent interactions, such as those studied in the present work.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2022.858946/full#supplementary-material

REFERENCES

Adhikari, U., and Scheiner, S. (2012). Sensitivity of Pnicogen, Chalcogen, Halogen and H-Bonds to Angular Distortions. *Chem. Phys. Lett.* 532, 31–35. doi:10.1016/j.cplett.2012.02.064

Angarov, V., and Kozuch, S. (2018). On the σ, π and δ Hole Interactions: a Molecular Orbital Overview. *New J. Chem.* 42, 1413–1422. doi:10.1039/c7nj03632a

Bader, R. (1994). *Atoms in Molecules: A Quantum Theory*. Oxford: Oxford University Press.

Bauzá, A., Quiñonero, D., Deyà, P. M., and Frontera, A. (2013). Halogen Bonding Versuschalcogen and Pnicogen Bonding: a Combined Cambridge Structural Database and Theoretical Study. *CrystEngComm* 15, 3137–3144. doi:10.1039/C2CE26741A

Benesi, H. A., and Hildebrand, J. H. (1948). Ultraviolet Absorption Bands of Iodine in Aromatic Hydrocarbons. *J. Am. Chem. Soc.* 70, 2832–2833. doi:10.1021/ja01188a528

Bora, P. L., Novák, M., Novotný, J., Foroutan-Nejad, C., and Marek, R. (2017). Supramolecular Covalence in Bifurcated Chalcogen Bonding. *Chem. Eur. J.* 23, 7315–7323. doi:10.1002/chem.201700179

Cavallo, G., Metrangolo, P., Milani, R., Pilati, T., Priimagi, A., Resnati, G., et al. (2016). The Halogen Bond. *Chem. Rev.* 116, 2478–2601. doi:10.1021/acs.chemrev.5b00484

Chałasiński, G., Szczęśniak, M. M., and Scheiner, S. (1991). Ab Initio study of the Intermolecular Potential of Ar–H_2O. *J. Chem. Phys.* 94, 2807–2816. doi:10.1063/1.459857

Cohen, R. C., and Saykally, R. J. (1993). Determination of an Improved Intermolecular Global Potential Energy Surface for Ar-H_2O from Vibration-Rotation-Tunneling Spectroscopy. *J. Chem. Phys.* 98, 6007–6030. doi:10.1063/1.464841

Desiraju, G. R., Ho, P. S., Kloo, L., Legon, A. C., Marquardt, R., Metrangolo, P., et al. (2013). Definition of the Halogen Bond (IUPAC Recommendations 2013). *Pure Appl. Chem.* 85, 1711–1713. doi:10.1351/PAC-REC-12-05-10

Evans, C. J., Wright, T. G., and Gardner, A. M. (2010). Geometries and Bond Energies of the He-MX, Ne-MX, and Ar-MX (M = Cu, Ag, Au; X = F, Cl) Complexes. *J. Phys. Chem. A.* 114 (114), 4446–4454. doi:10.1021/jp912027y

Hapka, M., Kłos, J., Korona, T., and Chałasiński, G. (2013). Theoretical Studies of Potential Energy Surface and Bound States of the Strongly Bound He(1S)-BeO (1Σ+) Complex. *J. Phys. Chem. A.* 117 (30), 6657–6663. doi:10.1021/jp404467b

Harris, S. J., Novick, S. E., Klemperer, W., and Falconer, W. E. (1974). Intermolecular Potential between an Atom and a Diatomic Molecule: The Structure of ArCIF. *J. Chem. Phys.* 61, 193–197. doi:10.1063/1.1681622

Hohenstein, E. G., Parrish, R. M., Sherrill, C. D., Turney, J. M., and Schaefer, H. F. (2011). Large-scale Symmetry-Adapted Perturbation Theory Computations via Density Fitting and Laplace Transformation Techniques: Investigating the Fundamental Forces of DNA-Intercalator Interactions. *J. Chem. Phys.* 135, 174107. doi:10.1063/1.3656681

Huber, K. P., and Herzberg, G. H. (2022). "Constants of Diatomic Molecules," in *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. Editors P.J. Linstrom and W.G. Mallard (Gaithersburg, MD: National Institute of Standards and Technology), 20899. doi:10.18434/T4D303

Jeziorski, B., Moszynski, R., and Szalewicz, K. (1994). Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* 94, 1887–1930. doi:10.1021/cr00031a008

Kitaura, K., and Morokuma, K. (1976). A New Energy Decomposition Scheme for Molecular Interactions within the Hartree-Fock Approximation. *Int. J. Quan. Chem.* 10, 325–340. doi:10.1002/qua.560100211

Klos, J., Szczesniak, M. M., and Chalasinski, G. (2004). Paradigm pre-reactive van der Waals complexes: X-HX and X-H_2(X = F, Cl, Br). *Int. Rev. Phys. Chem.* 23, 541–571. doi:10.1080/01442350500063634

Koch, W., Frenking, G., Gauss, J., Cremer, D., and Collins, J. R. (1987). Helium Chemistry: Theoretical Predictions and Experimental Challenge. *J. Am. Chem. Soc.* 109, 5917–5934. doi:10.1021/ja00254a005

Kolář, M. H., and Hobza, P. (2016). Computer Modeling of Halogen Bonds and Other σ-Hole Interactions. *Chem. Rev.* 116, 5155–5187. doi:10.1021/acs.chemrev.5b00560

Kolář, M., Hostaš, J., and Hobza, P. (2015). Correction: The Strength and Directionality of a Halogen Bond Are Co-determined by the Magnitude and Size of the σ-hole. *Phys. Chem. Chem. Phys.* 17, 23279. doi:10.1039/c5cp90146d

Kozuch, S., and Martin, J. M. (2013). Halogen Bonds: Benchmarks and Theoretical Analysis. *J. Chem. Theor. Comput.* 9 (9), 1918–1931. doi:10.1021/ct301064t

Lao, K. U., and Herbert, J. M. (2016). Energy Decomposition Analysis with a Stable Charge-Transfer Term for Interpreting Intermolecular Interactions. *J. Chem. Theor. Comput.* 12, 2569–2582. doi:10.1021/acs.jctc.6b00155

Legon, A. C. (1999). Prereactive Complexes of Dihalogens XY with Lewis Bases B in the Gas Phase: A Systematic Case for the Halogen Analogue B···XY of the Hydrogen Bond B···HX. *Angew. Chem. Int. Ed.* 38, 2686–2714. doi:10.1002/(sici)1521-3773(19990917)38:18<2686::aid-anie2686>3.0.co;2-6

Misquitta, A. J. (2013). Charge Transfer from Regularized Symmetry-Adapted Perturbation Theory. *J. Chem. Theor. Comput.* 9, 5313–5326. doi:10.1021/ct400704a

Modrzejewski, M., Rajchel, Ł., Szczęśniak, M. M., and Chałasiński, G. (2012). Dispersion-free Component of Non-covalent Interaction via Mutual Polarization of Fragment Densities. *J. Chem. Phys.* 136, 204109. doi:10.1063/1.4705282

Morokuma, K. (1971). Molecular Orbital Studies of Hydrogen Bonds. III. C=O···H-O Hydrogen Bond in H_2CO···H_2O and H_2CO···$2H_2O$. *J. Chem. Phys.* 55, 1236–1244. doi:10.1063/1.1676210

Moszynski, R., Heijmen, T., and Jeziorski, B. (1996). Symmetry-adapted Perturbation Theory for the Calculation of Hartree-Fock Interaction Energies. *Mol. Phys.* 88, 741–758. doi:10.1080/00268979609482451

Mulliken, R. S., and Person, W. B. (1969). *Molecular Complexes*. New York and London: Wiley-Interscience.

Nunzi, F., Cesario, D., Pirani, F., Belpassi, L., Frenking, G., Grandinetti, F., et al. (2017). Helium Accepts Back-Donation in Highly Polar Complexes: New Insights into the Weak Chemical Bond. *J. Phys. Chem. Lett.* 8, 3334–3340. doi:10.1021/acs.jpclett.7b01320

Parker, T. M., Burns, L. A., Parrish, R. M., Ryno, A. G., and Sherrill, C. D. (2014). Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and Performance for Interaction Energies. *J. Chem. Phys.* 140, 094106. doi:10.1063/1.4867135

Parrish, R. M., Burns, L. A., Smith, D. G. A., Simmonett, A. C., DePrince, A. E., III, Hohenstein, E. G., et al. (2017). Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theor. Comput.* 13 (7), 3185–3197. doi:10.1021/acs.jctc.7b00174

Pascoe, D. J., Ling, K. B., and Cockroft, S. L. (2017). The Origin of Chalcogen-Bonding Interactions. *J. Am. Chem. Soc.* 139, 15160–15167. doi:10.1021/jacs.7b08511

Politzer, P., Murray, J. S., and Clark, T. (2015). σ-Hole Bonding: A Physical Interpretation. *Top. Curr. Chem.* 358, 19–42. doi:10.1007/128_2014_568

Politzer, P., Murray, J. S., and Clark, T. (2013). Halogen Bonding and Other σ-hole Interactions: a Perspective. *Phys. Chem. Chem. Phys.* 15, 11178–11189. doi:10.1039/c3cp00054k

Politzer, P., Murray, J. S., and Clark, T. (2010). Halogen Bonding: an Electrostatically-Driven Highly Directional Noncovalent Interaction. *Phys. Chem. Chem. Phys.* 12, 7748–7757. doi:10.1039/c004189k

Pritchard, B. P., Altarawy, D., Didier, B., Gibson, T. D., and Windus, T. L. (2019). New Basis Set Exchange: An Open, Up-To-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* 59 (11), 4814–4820. doi:10.1021/acs.jcim.9b00725

Ramasubbu, N., Parthasarathy, R., and Murray-Rust, P. (1986). Angular Preferences of Intermolecular Forces Around Halogen Centers: Preferred Directions of Approach of Electrophiles and Nucleophiles Around Carbon-Halogen Bond. *J. Am. Chem. Soc.* 108, 4308–4314. doi:10.1021/ja00275a012

Rezac, J., and de la Lande, A. (2017). On the Role of Charge Transfer in Halogen Bonding. *Phys.Chem.Chem.Phys.* 19, 791–803. doi:10.1039/c6cp07475h

Sadlej, J., Chałasiński, G., and Szczęśniak, M. M. (1993). On the Nature of the Interaction Energy in the Ar-ClF Complex. *J. Chem. Phys.* 99, 3700–3706. doi:10.1063/1.466145

Saha, S., and Desiraju, G. R. (2017). σ-Hole and π-Hole Synthon Mimicry in Third-Generation Crystal Engineering: Design of Elastic Crystals. *Chem. Eur. J.* 23, 4936–4943. doi:10.1002/chem.201700813

Saha, S., Mishra, M. K., Reddy, C. M., and Desiraju, G. R. (2018). From Molecules to Interactions to Crystal Engineering: Mechanical Properties of Organic Solids. *Acc. Chem. Res.* 51 (11), 2957–2967. doi:10.1021/acs.accounts.8b00425

Scheiner, S. (2021). Anatomy of π-hole Bonds: Linear Systems. *J. Chem. Phys.* 155, 174302. doi:10.1063/5.0067570

Sedlak, R., Kolář, M. H., and Hobza, P. (2015). Polar Flattening and the Strength of Halogen Bonding. *J. Chem. Theor. Comput.* 11 (11), 4727–4732. doi:10.1021/acs.jctc.5b00687

Stone, A. J. (2013). Are Halogen Bonded Structures Electrostatically Driven? *J. Am. Chem. Soc.* 135, 7005–7009. doi:10.1021/ja401420w

Stone, A. J., and Misquitta, A. J. (2009). Charge-transfer in Symmetry-Adapted Perturbation Theory. *Chem. Phys. Lett.* 473, 201–205. doi:10.1016/j.cplett.2009.03.073

Tebai, Y., Jaidane, N.-E., Ben Abdallah, D., Halvick, P., Stoecklin, T., and Hochlaf, M. (2014). Theoretical Spectroscopic Characterization of the ArBeO Complex. *J. Chem. Phys.* 141, 174305. doi:10.1063/1.4900770

Thirman, J., Engelage, E., Huber, S. M., and Head-Gordon, M. (2018). Characterizing the Interplay of Pauli Repulsion, Electrostatics, Dispersion and Charge Transfer in Halogen Bonding with Energy Decomposition Analysis. *Phys. Chem. Chem. Phys.* 20, 905–915. doi:10.1039/c7cp06959f

Thirman, J., and Head-Gordon, M. (2015). An Energy Decomposition Analysis for Second-Order Møller-Plesset Perturbation Theory Based on Absolutely Localized Molecular Orbitals. *J. Chem. Phys.* 143, 084124. doi:10.1063/1.4929479

Wang, C., Danovich, D., Mo, Y., and Shaik, S. (2014). On the Nature of the Halogen Bond. *J. Chem. Theor. Comput.* 10, 3726–3737. doi:10.1021/ct500422t

Wang, W., Ji, B., and Zhang, Y. (2009). Chalcogen Bond: A Sister Noncovalent Bond to Halogen Bond. *J. Phys. Chem. A.* 113, 8132–8135. doi:10.1021/jp904128b

Web of Science Core Collection (2022). *A Search of "Web of Science Core Collection" for Sigma & Hole Retrieves 3920 Publications as of This Writing (Accessed January 10, 2022).*

Wu, Q., and Van Voorhis, T. (2005). Direct Optimization Method to Study Constrained Systems within Density-Functional Theory. *Phys. Rev. A.* 72, 024502. doi:10.1103/PhysRevA.72.024502

Zierkiewicz, W., Michalczyk, M., and Scheiner, S. (2021). Noncovalent Bonds through Sigma and Pi-Hole Located on the Same Molecule. Guiding Principles and Comparisons. *Molecules* 26, 1740. doi:10.3390/molecules26061740

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Chemistry

ORIGINAL RESEARCH
published: 14 April 2022
doi: 10.3389/fchem.2022.871659

Recombinant Antibody-Based and Computer-Aided Comprehensive Analysis of Antibody's Equivalent Recognition Mechanism of Alternariol and Alternariol Monomethyl Ether

*Zile Wang[1], Ling Chen[2], Pimiao Zheng[1,3], Jianyi Wang[1], Zhenhui Ren[1], Huixia Zhang[1], Liang Zhang[1] and Haiyang Jiang[1]**

[1]Department of Veterinary Pharmacology and Toxicology, Beijing Key Laboratory of Detection Technology for Animal-Derived Food Safety, Beijing Laboratory for Food Quality and Safety, College of Veterinary Medicine, China Agricultural University, Beijing, China, [2]China Institution of Veterinary Drug Control, Beijing, China, [3]College of Veterinary Medicine, Shandong Agricultural University, Taian, China

OPEN ACCESS

Edited by:
Piotr Żuchowski,
Nicolaus Copernicus University, Poland

Reviewed by:
Maojun Jin,
Chinese Academy of Agricultural Sciences (CAAS), China
Lin Luo,
South China Agricultural University, China

***Correspondence:**
Haiyang Jiang
haiyang@cau.edu.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: 08 February 2022
Accepted: 21 March 2022
Published: 14 April 2022

Citation:
Wang Z, Chen L, Zheng P, Wang J, Ren Z, Zhang H, Zhang L and Jiang H (2022) Recombinant Antibody-Based and Computer-Aided Comprehensive Analysis of Antibody's Equivalent Recognition Mechanism of Alternariol and Alternariol Monomethyl Ether. Front. Chem. 10:871659. doi: 10.3389/fchem.2022.871659

Alternariol (AOH) and alternariol monomethyl ether (AME) are two main *Alternaria* mycotoxins that endanger human health. In this study, a single-chain antibody fragment (scFv) capable of equivalently and specifically recognizing AOH and AME was first expressed, and its equivalent recognition mechanism was discussed. According to molecular docking and dynamic simulation, the C9 site, which was always exposed outside the binding cavity, made the structural differences between AOH and AME negligible. Due to the high similarity of structures, AOH and AME interacted with almost the same amino acids on the scFv; thus, the same interaction mode and interaction force were produced. This was considered to be the most critical reason for the equivalent recognition. Thus, the exposure of common structures was considered a potential strategy to obtain the equivalent recognition antibodies, and C9 was considered the key site in the process of hapten modification. These results laid a theoretical foundation for further research on antibodies against *Alternaria* mycotoxins. It could promote the rapid detection of AOH and AME in food and provide a new idea for targeted preparation of antibodies that could recognize multiple hazards with similar structures.

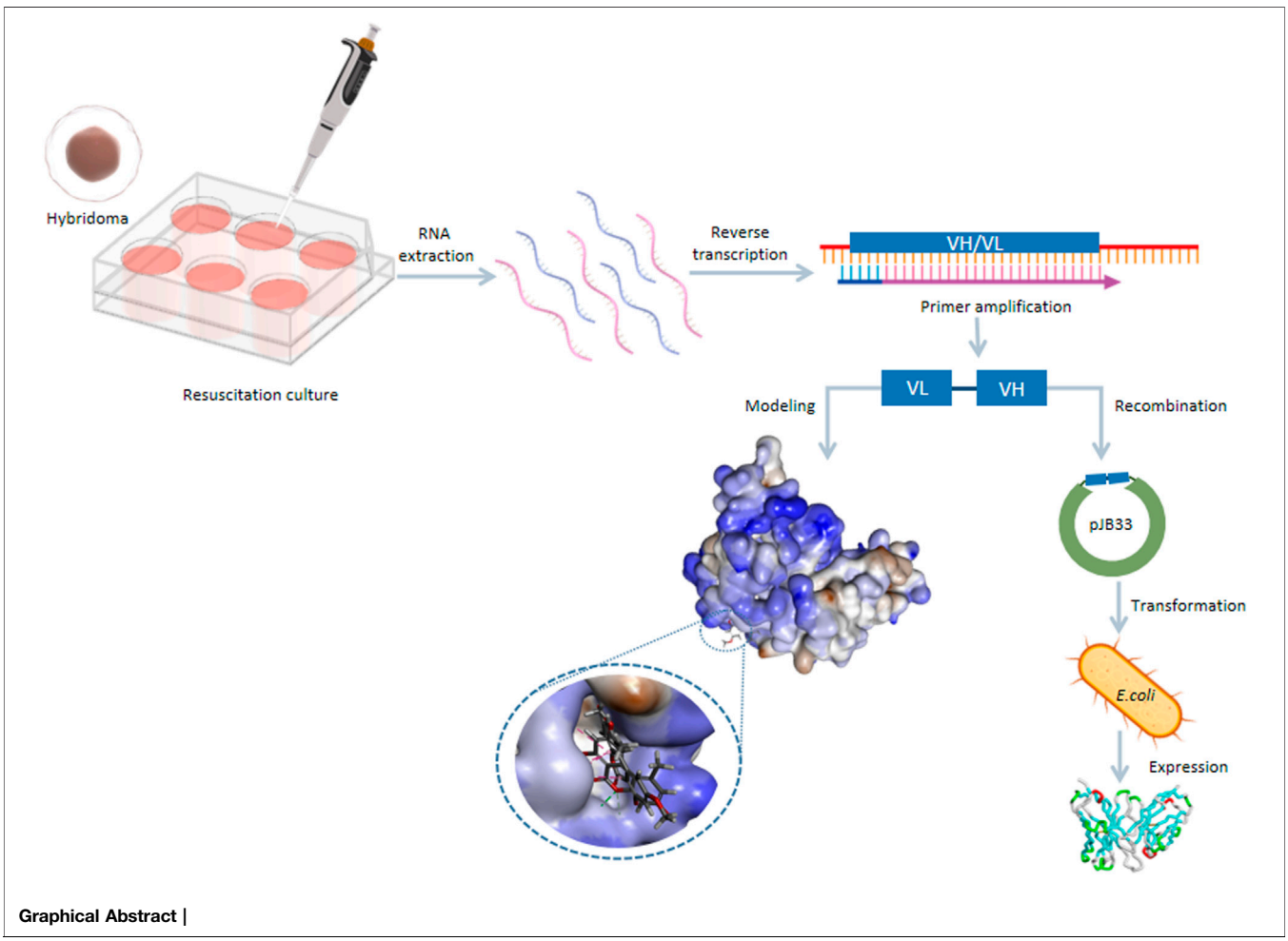
Keywords: molecular interaction, recognition mechanism, alternariol, alternariol monomethyl ether, single-chain antibody fragment

INTRODUCTION

Mycotoxins have teratogenic, mutagenic, and carcinogenic effects even at low concentrations, and long-term consumption of food contaminated with mycotoxins may cause serious consequences, such as cancer (Bhat et al., 2010). In recent years, "emerging mycotoxins" including *Alternaria* mycotoxins are attracting increasing public attention due to their potential toxicity (Gruber-Dorninger et al., 2017). *Alternaria* mycotoxins are secondary metabolites produced by *Alternaria* species. According to the European Food Safety Authority (EFSA), alternariol (AOH) and alternariol monomethyl ether (AME) are two main *Alternaria* mycotoxins that endanger human health (ESFA,

2011). They belong to dibenzo-α-pyrone derivatives and structurally resemble altenuene (ALT) (**Figure 1**). *Alternaria* mycotoxins are widely distributed in nature, and they could be detected in milk, cereals, fruits, vegetables, spices, and herbs (Gambacorta et al., 2019; Li et al., 2020; Orina et al., 2021; Akinyemi et al., 2022). It is reported that *Alternaria* mycotoxins are thermally stable (Estiarte et al., 2018). AOH and AME showed remarkable cytotoxicity to mammalian cells *in vitro* (Solhaug et al., 2014), and they had a synergistic effect when used together in high concentrations. Tenuazonic acid (TeA) exhibits toxicity against several animal species such as mice, chickens, and dogs (EFSA, 2011). In addition, the incidence of esophageal cancer is higher in areas with high AOH and AME food contamination. Currently, there is no regulation for *Alternaria* mycotoxins worldwide except for the Bavarian Health and Food Safety Authority which established a limit of 500 μg/kg for TeA in sorghum/millet-based infant food (Gambacorta et al., 2019). However, the defined threshold of toxicological concern (TTC) of AOH and AME was 2.5 ng/kg body weight, while TeA was 1,500 ng/kg body weight (ESFA, 2011). Therefore, it is of great significance to strengthen the monitoring of the incidence of *Alternaria* mycotoxins (especially AOH and AME) for making policies and ensuring human health.

In the last decade, liquid chromatography–tandem mass spectrometry (LC-MS/MS) (Krausová et al., 2021), immunoassay (Ackermann et al., 2011; Burkin and Kononenko, 2011; Kong et al., 2017; Man et al., 2017; Singh et al., 2018; Wang et al., 2018; Cai et al., 2021; Liang et al., 2021) and other methods with molecularly imprinted polymer or aptamers as core reagents were used to detect *Alternaria* mycotoxins in food and biology samples (Quílez-Alburquerque et al., 2021; Wang et al., 2022). Given the advantages of simple operation, less time consumption, and low cost, immunological methods were considered to be the most suitable method for the on-site detection of *Alternaria* mycotoxins in food samples (Ackermann et al., 2011), and an antibody was supposed to be the most important reagent. However, the preparation of an antibody could be influenced by many factors: it has certain randomness because of the uncertainty of the hapten design, the lack of knowledge about the recognition mechanism, and the ambiguous structure–activity relationship, which make the targeted preparation of the antibody still a considerable challenge (Liu et al., 2007). The reported immunological methods based on monoclonal antibodies (mAbs) or polyclonal antibodies (pAbs) mostly only show good recognition for a certain *Alternaria* mycotoxin. Ackermann



Graphical Abstract |

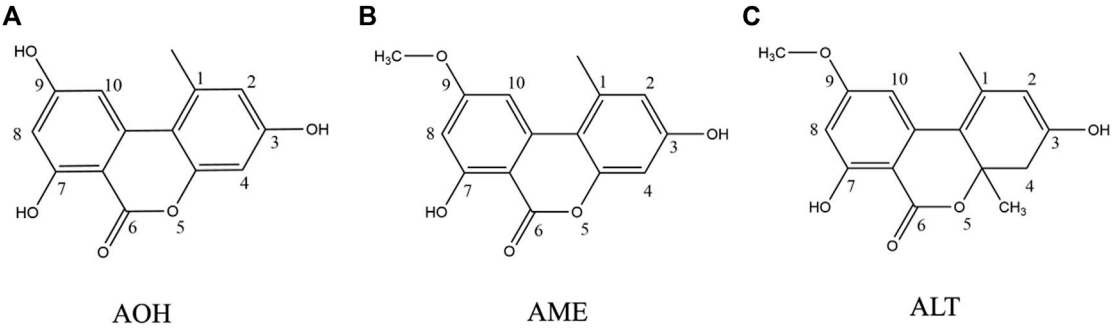


FIGURE 1 | Structure of three main *Alternaria alternata* mycotoxins. **(A)** Alternariol. **(B)** Alternariol monomethyl ether. **(C)** Altenuene.

et al. (2011) prepared mAbs and pAbs via modified AOH. The antibodies could sensitively recognize AOH but almost not recognize AME and ALT. The same result also appeared in the research of Burkin and Kononenko (2011) and Singh et al. (2018). Man et al. (2017) prepared mAbs via modified AME, which could sensitively recognize AME but almost not recognize AOH. Developing a method for simultaneous detection of *Alternaria* mycotoxins is of great significance, but antibodies for multiple recognition have not yet been developed. In our previous work, mAb, which equivalently recognizes AOH and AME, was obtained (Wang et al., 2018), and this may be a breakthrough to develop a broad-spectrum antibody for *Alternaria* mycotoxins.

With the development of bioinformatics, computer-aided and structure-based homologous modeling, molecular docking, etc. could be used to predict protein function; analyze the intermolecular interaction; and could help improve the affinity via virtual mutation calculation (Liu et al., 2021; Li et al., 2022). In this study, a hybridoma cell that secretes mAbs, which equivalently recognizes AOH and AME, was used to construct a single-chain antibody fragment (scFv) gene, and the scFv was successfully expressed. The three-dimensional structure of the scFv was constructed, the mechanism of the scFv's equivalent recognition of AOH and AME was clarified, and the key sites of the hapten design were also discussed. It laid a theoretical foundation for further research on antibodies against *Alternaria* mycotoxins, provided a new idea for targeted preparation of antibodies that could recognize multiple hazards with similar structures and could also promote the simultaneous and rapid detection of *Alternaria* mycotoxins in food.

MATERIALS AND METHODS

Chemicals and Reagents

AOH, AME, ALT, TeA, and tentoxin (TEN) were purchased from J&K Scientific Ltd. (Beijing, China). Hybridoma 13D8 (secretes mAb that specifically and equivalently recognizes AOH and AME) and AOH-BSA coating antigen were previously prepared in our laboratory (Wang et al., 2018). Horseradish peroxidase (HRP)-conjugated His-tag monoclonal antibody was purchased from Proteintech Group, Inc. (Rosemont, United States). The total RNA extraction kit was purchased from Qiagen (Dusseldorf, Germany) and the RT-PCR kit was purchased from ThermoFisher Scientific (Waltham, United States). The plasmid extraction kit was purchased from Tiangen Biotech Co.,, Ltd. (Beijing, China). *Escherichia coli* RV308 was purchased from ATCC. The vector pJB33 used for scFv expression was obtained as a gift from the laboratory of Andreas Plückthun (Biochemisches Institute, Universität Zürich, Switzerland). The restriction enzyme SfiI was purchased from Takara (Dalian, China). DNA polymerase and T4 DNA ligase were purchased from TransGen Biotech (Beijing, China). All the other chemical reagents were purchased from Sinopharm Chemical Reagent Co.,, Ltd. (Shanghai, China).

Software

Discovery Studio 2019 (NeoTrident, China), GaussView 6.0 and Gaussian 09W (Gaussian, United States), ChemDraw (PerkinElmer, United States), Multiwfn 3.7 (dev) code (Lu and Chen, 2012), and VMD visualization program (Humphrey et al., 1996) were used in the study.

Construction of Single-Chain Antibody Fragment Gene

The process of constructing the scFv was similar to the one suggested by Xu et al. (2012). Briefly, the total RNA of the hybridoma cell 13D8 was extracted, and the concentration was determined by using a Nanodrop 2000C spectrophotometer (ThermoFisher, MA, United States). Then, RNA was used as a template to synthesize cDNA through RT-PCR. The template of the cDNA and nested PCR were employed in the amplification of VH and VL. Finally, the splicing of the VL-linker–VH fragment was accomplished by SOE-PCR. The PCR product was identified by agarose gel electrophoresis, and then the gel was collected for obtaining the scFv gene fragments. The gene fragment of the scFv and pJB33 vectors was digested by SfiI, and the two digested products were linked for constructing the recombinant express vector pJB33-scFv. The recombinant vector was transformed into *E. coli* RV308 and then inoculated into the YT culture medium (containing chloramphenicol). The positive colonies were sent to Taihe Biotechnology (Beijing) Co., Ltd., for sequence

identification. Details of the reaction systems are listed in the supplemental materials.

Expression of Single-Chain Antibody Fragment

E. coli RV308, which contained the recombinant plasmid pJB33-scFv, was shaken at 37°C overnight. When the OD_{600} value was 0.8, 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added for induction expression (24°C, 150 rpm, 12 h). Then, ultrasound (200 W) was used to break the bacteria in an ice bath until the solution was clear and transparent, and then the solution was centrifuged at 10,000 rpm for 20 min to collect the supernatant. A Ni agarose resin column was used to purify the supernatant, then SDS-PAGE and Western blot were used for identification, and ic-ELISA was used to determine the half-maximal inhibitory concentration (IC_{50}) of the scFv. The detailed steps of ic-ELISA are listed in the supplementary materials.

Structural Analysis of Alternariol and Alternariol Monomethyl Ether

Three-dimensional structures of AOH and AME were obtained from the PubChem database and then initially prepared in GaussView 6.0. A geometry optimization procedure based on the density functional theory (DFT) was executed in the Gaussian 09W using the M06-2X density functional and TZVP basis set (Garcia et al., 2018). The hydrophobic constant (log P) was shown in ChemDraw. Solvent accessible surface areas (SASA) were extracted by VMD visualization program. The molecular electrostatic potential (ESP)-mapped van der Waals surface was rendered by means of VMD based on the files exported by Multiwfn (Wen et al., 2020). Molecular structural similarity was calculated by "Molecular Overlay" in Discovery Studio 2019 (DS).

Homology Modeling of Single-Chain Antibody Fragment

The spatial model of the scFv was constructed on the basis of the PDB database. First, the complementarity-determining regions (CDRs) of VH and VL were annotated by the "Model Antibody" module in DS based on IMGT. The sequence with the highest similarity and identity was selected as the template to generate scFv spatial models. According to the value of the probability density function (PDF) (automatically generated by DS), the best simulated conformation was selected, then the CDRs of the model were optimized by BLAST to form the final model. The Ramachandran plot diagram and profile-3D score were used to evaluate the final spatial mode.

Molecular Docking and Dynamic Simulation

AOH and AME were docked into the active pocket of the scFv by the "CDOCKER Module" of DS after being drawn by the "Sketch Module" and given the CHARMm force field for structure optimization (Zhang et al., 2019). The active pocket was defined as all atoms within a 10 Å radius of the protein cavity. All water molecules were removed, and hydrogen atoms were added in. The remaining parameters were set to the default values of the program. The AOH–scFv and AME–scFv complexes were given the CHARMm36 force field and explicit periodic boundary solvation. The molecular dynamic simulation was carried out via the "Standard Dynamics Cascade" module of DS.

Virtual Mutation

The complex was given the CHARMm force field, and all amino acids within a 3 Å radius of the binding site were subjected to single-point saturation mutation in the "Design protein" module of DS. When setting parameters, pH and ionic strength were set to 7.45 and 0.15 M, respectively, and the amino acid causing scFv affinity to increase or decrease was predicted according to the mutation results. A mutation energy less than −0.5 could lead to an increase in the affinity, a mutation energy between −0.5 and 0.5 had no significant effect on the affinity, while a mutation energy greater than 0.5 could lead to a decrease in the affinity.

RESULTS AND DISCUSSION

Preparation of Single-Chain Antibody Fragment

The smallest immunoglobulin antigen-binding fragment that keeps a complete antigen-binding site was the Fv fragment, which only consists of a variable (V) region. Then, a soluble and flexible peptide linker was used to stabilize the connection of the V region to form an scFv (Bird et al., 1988). Even though increasingly recombinant production systems have been developed (such as *E. coli*, yeast, insect cell lines, mammalian cells), *E. coli* is still the most commonly used system for the scFv. In this study, *E. coli* RV308 could express soluble scFv, and the scFv could enter the periplasmic cavity under the action of a guiding peptide and fold correctly under the action of an oxidizing environment and various molecular partners. A peptide linker $(G_4S)_3$ could provide enough flexibility for VH and VL domains, so it could form monovalent antigen-binding sites equivalent to Fab fragments of the mAb (Huston et al., 1991).

The concentration of the total RNA was 536.2 ng/μL, and the A260/A280 value was 2.12, which indicated that there was no contamination. As shown in **Supplementary Figure S1**, agarose gel electrophoresis showed that the VH was about 360 bp, VL was about 320 bp, and the scFv was about 750 bp. The full-length fragment of the scFv gene was successfully constructed. SDS-PAGE and Western blot results are shown in **Supplementary Figure S2**. All the samples showed a single band at about 30 kDa, indicating that there were no other miscellaneous proteins, and the scFv was successfully expressed. The concentration of the scFv was 1.4 mg/ml. The IC_{50} values of the scFv for AOH and AME were 509.0 ng/ml and 565.7 ng/ml, respectively (**Supplementary Figure S3**), and there is no cross-recognition for ALT, TeA, and TEN. It proved that the expressed scFv had biological activity. Due to the lack of structure, the affinity of the scFv was lower than that of the parent mAb. The same results also appeared in other

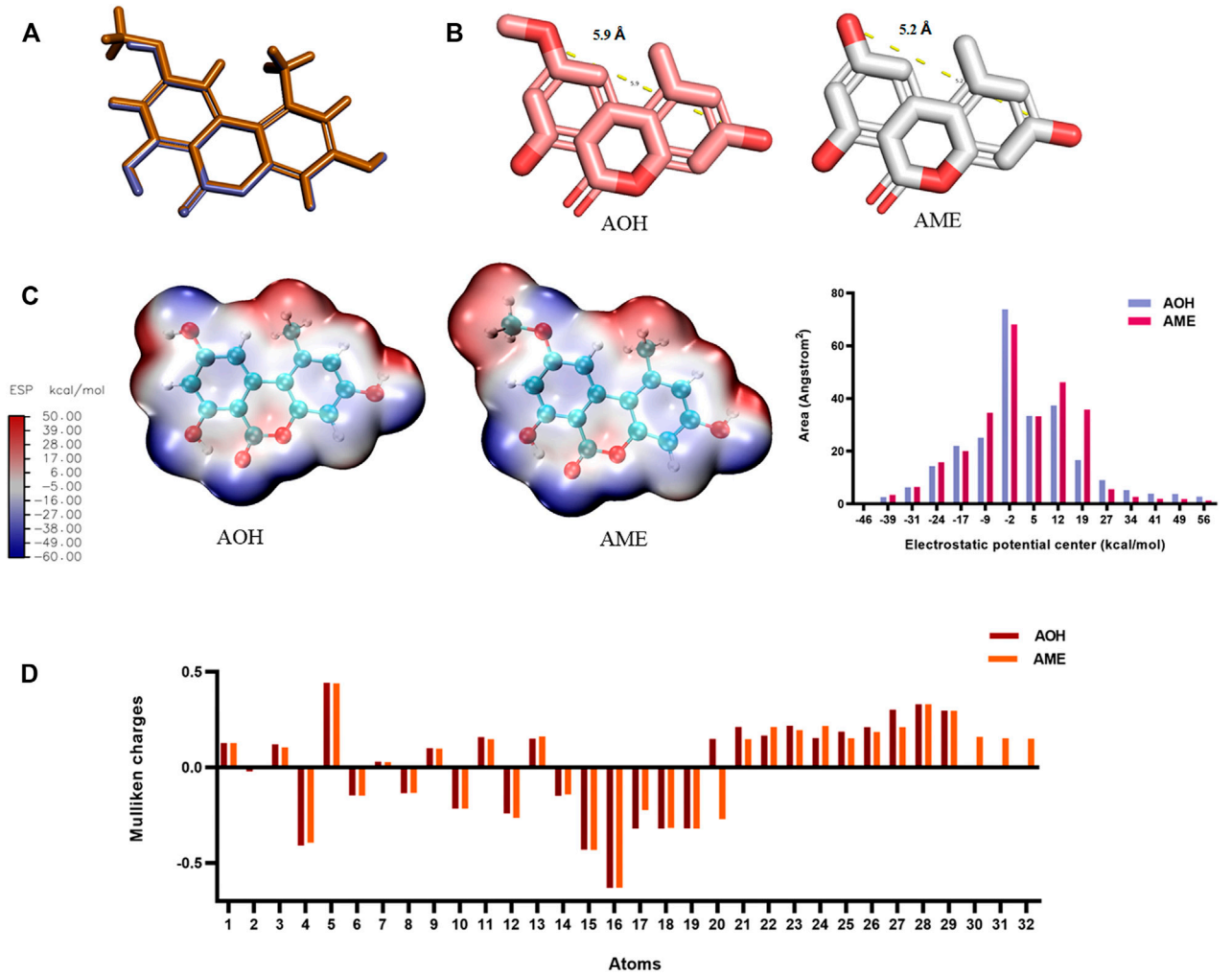


FIGURE 2 | Structural analysis of AOH and AME. **(A)** Molecular overlay (purple represents AOH and yellow represents AME). **(B)** Molecular scale of AOH and AME. **(C)** ESP of AOH and AME. **(D)** Mulliken charge of AOH and AME.

research studies (KorpimäKi et al., 2002; Xie et al., 2020). Although the affinity decreased, the recognition of AOH, AME, ALT, TeA, and TEN was consistent with the parent antibody, which proved the reliability of the results.

Structural Analysis of Alternariol and Alternariol Monomethyl Ether

As shown in **Figure 2A**, the molecular overlay was based on the influence of 50% of the space field and 50% of the electrostatic field. The coincidence degree of AOH and AME reached 98.76%, and there was difference in the C9 group, while the coincidence degree of AOH with ALT, TeA, and TEN was 79.53%, 61.89%, and 68.43%, respectively. As shown in **Figure 2B**, the molecular scales of AOH and AME were calculated, the distance between the C3 hydroxyl group and C9 hydroxyl group of AOH was 5.2 Å, and the distance between the C3 hydroxyl group and C9 methoxy group of AME was 5.9 Å. The difference was negligible. **Figure 2C** shows that the methoxy group was positively charged, while the hydroxyl group was negatively charged. If C9 was the antigen-recognition site, it might greatly affect the recognition result of the antibody. **Figure 2D** shows the Mulliken charge of different atoms of AOH and AME (**Supplementary Figure S4**), with the largest charge difference of numbers 20, 30, 31, and 32, pointing to C9 methoxy of AME. In addition, log P and SASA of AOH and AME were 2.48 and 443.54, and 2.74 and 463.65, respectively, which indicated that AOH and AME had very similar hydrophobic properties. The aforementioned results provided necessary conditions for the appearance of antibody equivalent recognition.

Construction of Single-Chain Antibody Fragment 3D Structure

X-ray, nuclear magnetic resonance (NMR), and cryo-electron microscopy are commonly used to directly obtain the 3D

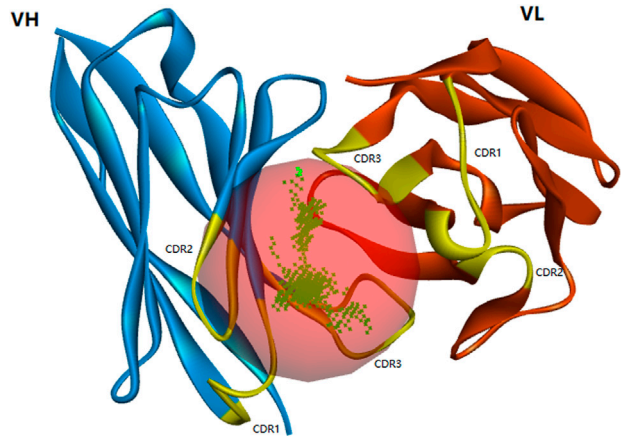


FIGURE 3 | Optimized spatial structure model of scFv (yellow highlights represent the distribution of CDR regions and red balls represent active cavity located in CDR regions).

structure of proteins, but these methods are time-consuming and expensive. Homologous modeling has been widely used (Wang et al., 2020). As shown in **Supplementary Figure S5**, three complementarity-determining regions (CDRs) were annotated in VH and VL, respectively. Blue represented the VL region, and the CDRs were located at 27–33 (CDR1), 51–53 (CDR2), and 89–97 (CDR3), respectively. Green represented the VH region, and the CDRs were located at 153–160 (CDR1), 178–184 (CDR2), and 223–236 (CDR3), respectively. The template 6BDZ with 95% similarity and 87.8% identity was selected from the PDB database, and then the best model with the smallest PDF value was constructed (**Figure 3**). It showed that CDR1, CDR2, and CDR3 of the heavy chain and CDR3 of the light chain formed the binding active cavity.

A reliable model should have more than 90% residues in the allowable area of the Ramachandran plot (Zhang et al., 2019). As shown in **Supplementary Figure S6A**, the blue and purple regions represent the allowed regions, and red represents the disallowed ones. In this model, 99.5% of amino acid residues were located in the allowed regions, while 0.5% of amino acids were distributed in the disallowed regions. The profile-3D evaluation of amino acid sequence-matching degree showed that no amino acids were in the low sub-region structure and mismatched the sequence (**Supplementary Figure S6B**), indicating that the scFv 3D model was reliable.

Molecular Docking and Dynamic Simulation

AOH, AME, and the scFv were subjected to semiflexible docking by CDOCKER (Zhang et al., 2019). As shown in **Figure 4**, AOH and AME entered the binding site of the scFv vertically, the whole parent nucleus was surrounded by the cavity, and C3 hydroxyl was deeply inside to the cavity, while C9 hydroxyl was exposed outside adequately. The donor and acceptor of the hydrogen bond, hydrophobicity, charge, and acid–base around the binding cavity are shown in **Figure 5**. **Figures 5A–D** show that there are abundant hydrogen-bond donors and acceptors around the binding cavity, while AOH and AME have multiple hydroxyl and oxygen atoms, which provides favorable conditions for the formation of a hydrogen bond between AOH/AME and the scFv. The hydrophobicity of the binding cavity is almost weak, while

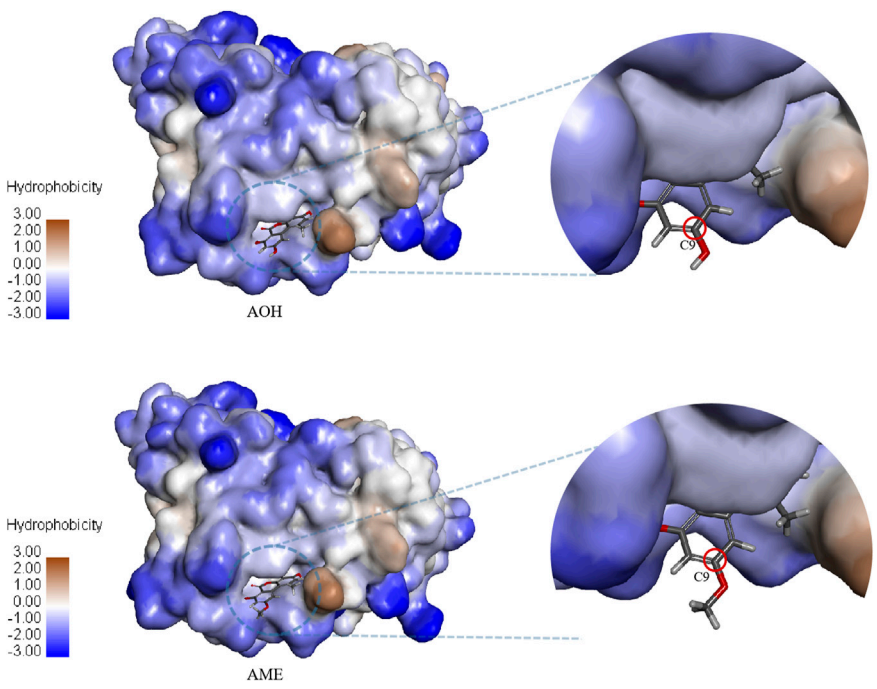


FIGURE 4 | Molecular docking model of AOH/AME and scFv.

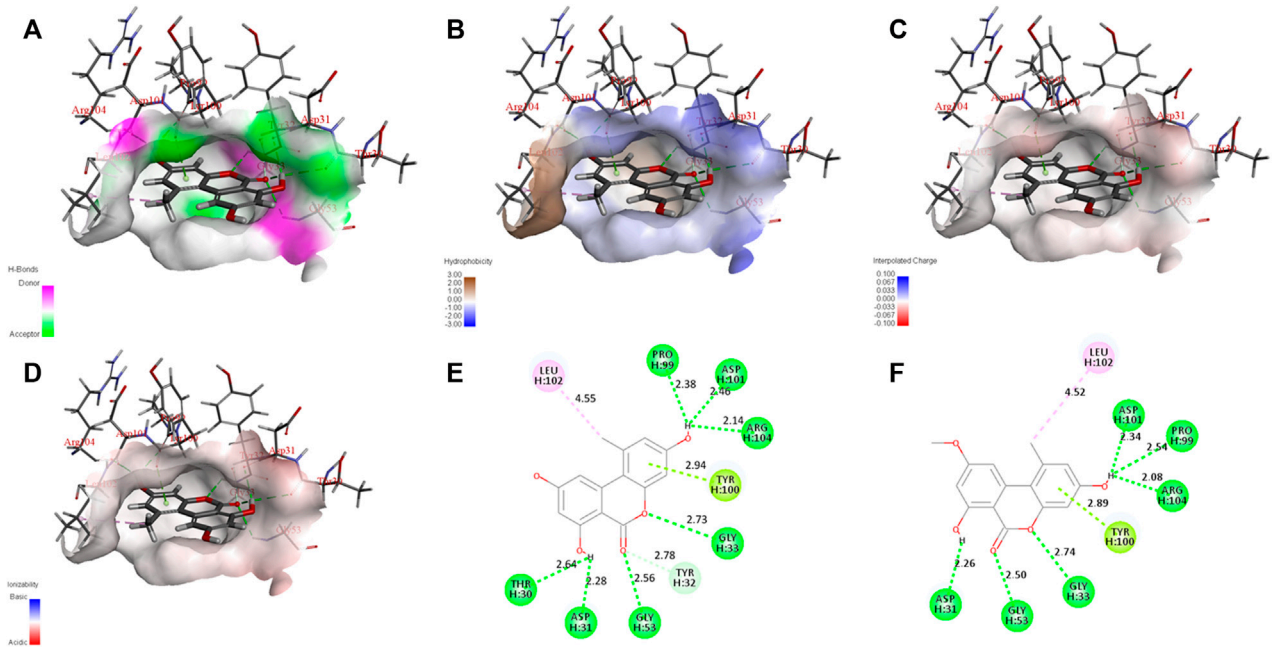


FIGURE 5 | Interaction between AOH/AME and scFv. **(A)** Hydrogen bond donors and acceptors around the binding cavity (green represents the receptor and purple represents the donor.) **(B)** Hydrophobicity around the binding cavity (blue represents hydrophilicity and brown represents hydrophobicity). **(C)** Charge around the binding cavity (blue represents positive electricity and red represents negative electricity). **(D)** The acid–base around the binding cavity (blue represents basic and red represents acidic). **(E)** Two-dimensional diagram of interactions between scFv and AOH. **(F)** Two-dimensional diagram of interactions between scFv and AME (green represents hydrogen bond interaction, cyan represents π-lone pair interaction, and pink represents alkyl interaction).

TABLE 1 | Existing methods for preparing antibodies for AOH and AME.

Antibody type	Compound	Modification method	Modification site	Cross-reaction rate (%)	
mAb	AOH	Mannich	C3 or C7 hydroxyl	AOH (100) AME (0.9)	Ackermann, et al. (2011)
pAb	AOH			AOH (100) AME (<0.5) ALT (<0.5)	
mAb	AME	Carboxyl derivative modification	C3 hydroxyl	AME (100) AOH (2.9)	Man et al. (2017)
mAb	AOH	CDI	C3 or C7 or C9 hydroxyl	AOH (100) AME (24.6)	Kong et al. (2017)
mAb	AOH	Alkylation reaction	C3 or C7 or C9 hydroxyl	AOH (100) AME (97)	Wang et al. (2018)

AOH and AME have multiple benzene ring structures, which may be one of the reasons for the undesirable affinity of antibodies. Moreover, the cavity was a neutral environment; thus, it is difficult to form an electrostatic interaction between AOH/AME and the scFv.

In **Figure 5E**, the molecular docking results show that 10 amino acids interacted with AOH; Thr 30 (2.64 Å) and Asp 31 (2.28 Å) formed hydrogen bonds with C7 hydroxyl groups; Gly 53 (2.56 Å) and Tyr 32 (2.78 Å) formed hydrogen bonds with C6 oxygen; Gly 33 (2.73 Å) formed hydrogen bonds with O5; Tyr 100 (2.94 Å) formed π-lone pair interaction with a benzene ring; Pro 99 (2.38 Å), Asp 101 (2.46 Å), and Arg 104 (2.14 Å) formed hydrogen bonds with C3 hydroxyl groups; and Leu 102 (4.55 Å) formed an alkyl interaction with C1 methyl groups. As shown in **Figure 5F**, eight amino acids interacted with AME; Asp 31 (2.26 Å) formed hydrogen bonds with C7 hydroxyl groups; Gly 53 (2.50 Å) formed hydrogen bonds with C6 oxygen, Gly 33 (2.74 Å) formed hydrogen bonds with O5; Tyr 100 (2.89 Å) formed π-lone pair interaction with a benzene ring; Pro 99 (2.54 Å), Asp 101 (2.34 Å), and Arg 104 (2.08 Å) formed hydrogen bonds with C3 hydroxyl groups; and Leu 102 (4.52 Å) formed alkyl interaction with C1 methyl. These results showed that the interaction between AOH, AME, and the scFv was mainly of hydrogen bonds. The amino acids that

produce the interaction between them were the same, and this was the intuitive reason for the equivalent recognition of the antibody. The structural difference between AOH and AME was concentrated at the C9 site; however, in the docking model, the C9 sites of both AOH and AME were exposed outside the binding cavity and have no relevant force with the scFv, so the influence of structural differences on antibody recognition could be ignored, and this may be the most important factor of equivalent recognition.

As shown in **Table 1**, the mAb and pAb obtained by mixing AOH C3 and C7 hydroxyl-modified immunogens could only specifically recognize AOH (Ackermann, et al., 2011; Kong et al., 2017) and the mAb obtained by the AME C3 hydroxyl-modified immunogens could only recognize AME (Man et al., 2017). Wang et al., 2018 obtained mAbs that could specifically and equivalently recognize AOH and AME through mixed immunogens modified by C3, C7, and C9 hydroxyl groups of AOH. It is speculated that modifying the C3 and C7 hydroxyl will expose the structural difference, so the obtained antibodies could recognize certain compounds (Ackermann, et al., 2011). However, the structural difference between AOH and AME will be well-solved by modifying the C9 site; therefore, the modification of the C9 site may be the key point to the targeted preparation of antibodies that recognize AOH and AME equivalently. In the process of trying to dock ALT with the scFv model, it was shown that effective docking could not be formed, which was consistent with the conclusion that the antibody had no cross-reaction to ALT, which further demonstrated the reliability of the docking model. The failure to recognize ALT may be due to the addition of the methyl structure, which caused C3 hydroxyl to fail to enter the binding cavity. Thus, in the process of immunogen modification, exposing the common structure of a class of structural analogs and avoiding structural differences between the immunogen and target compounds may be a possible way to solve the difficulty of preparing broad-spectrum antibodies.

The results of molecular docking provided an explanation for the state of antigen–antibody interaction, but in fact, antigen–antibody interaction was a dynamic reaction process. Molecular dynamics is a technique for studying the physical motion of atoms or molecules in a complex system. It could predict the stability of the complex and is performed to analyze the stability of the protein–ligand complexes on a picosecond (ps) scale. The movement of the atoms in a chemical complex is discovered by keeping the temperature, volume, and pressure parameters constant for a set period of time (Bhardwaj et al., 2022). In this study, the dynamic interaction process of AOH–scFv and AME–scFv complexes in an aqueous solvent was studied. In 200 ps, the C9 site of AOH and AME was always exposed outside the binding cavity, which further confirmed the conclusion that the modification of the C9 site was the key point for preparing an antibody that recognizes AOH and AME equivalently.

Affinity Evolution

The scFv provides conditions for affinity evolution *in vitro*. Xie et al. (2020) explored the recognition mechanisms between amantadine and the scFv, and after virtual mutation and expression, the results indicated that the sensitivity of the scFv was improved by 3.9 times by mutating Gly 107 to Phe. In this study, there were 13 amino acids (Thr 30, Asp 31, Tyr 32, Gly 33, Trp 52, Gly 53, Val 98, Pro 99, Tyr 100, Asp 101, Leu 102, Arg 104, and Tyr 105 of VH) within a 3 Å radius of the binding site. As shown in **Supplementary Table S5**, the mutations of Thr 30, Asp 31, Pro 99, Leu 102, and Arg 104 almost will not affect the affinity of the scFv, whereas mutations of Gly 33 to Leu and Tyr 100 to Phe could improve the affinity of the scFv. In addition, the mutation of Tyr 32, Gly 33, Trp 52, Gly 53, Val 98, Tyr 100, Asp 101, and Tyr 105 could all reduce the affinity of the scFv. These results provided a scheme for further enhancing the affinity of the scFv for the detection of actual samples.

CONCLUSION

In summary, this study expressed an scFv that could equivalently recognize AOH and AME and laid a foundation for the directed evolution of an antibody at the molecular level. It could promote the development of rapid detection technology for both AOH and AME in food. By analyzing the molecular structure and interaction processes, it was clarified that the equivalent recognition depended on the structural similarity of AOH and AME, and almost the same interaction between the amino acid and acting force was found. The structural difference was well-solved because different C9 sites were exposed outside the binding cavity, which made the common structure of AOH and AME fully exposed to the antibody. Hence, C9 was considered to be the decisive site to determine the equivalent recognition of the antibody in the process of hapten modification. Therefore, the exposure of common structures and avoiding re-formation of similar structures may be an important strategy for the targeted preparation of antibodies that recognize multiple hazards with similar structures, which could guide future hapten design and antibody improvement.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the **Supplementary Material**.

AUTHOR CONTRIBUTIONS

HJ guided the design of the experiment and the framework of the paper; ZW performed the part of the computer-aided analysis and the writing of the manuscript; LC and JW carried out the amplification and expression of antibody sequence; and PZ, ZR, HZ, and LZ assisted in the experiment and the writing of the manuscript.

FUNDING

This research is supported by the Ministry of Science and Technology of the People's Republic of China (2019YFC1604902).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2022.871659/full#supplementary-material

REFERENCES

Ackermann, Y., Curtui, V., Dietrich, R., Gross, M., Latif, H., Märtlbauer, E., et al. (2011). Widespread Occurrence of Low Levels of Alternariol in Apple and Tomato Products, as Determined by Comparative Immunochemical Assessment Using Monoclonal and Polyclonal Antibodies. *J. Agric. Food Chem.* 59 (12), 6360–6368. doi:10.1021/jf201516f

Akinyemi, M. O., Braun, D., Windisch, P., Warth, B., and Ezekiel, C. N. (2022). Assessment of Multiple Mycotoxins in Raw Milk of Three Different Animal Species in Nigeria. *Food Control* 131, 108258. doi:10.1016/j.foodcont.2021.108258

Bhardwaj, P., Biswas, G. P., Mahata, N., Ghanta, S., and Bhunia, B. (2022). Exploration of Binding Mechanism of Triclosan towards Cancer Markers Using Molecular Docking and Molecular Dynamics. *Chemosphere* 293, 133550. doi:10.1016/j.chemosphere.2022.133550

Bhat, R., Rai, R. V., and Karim, A. A. (2010). Mycotoxins in Food and Feed: Present Status and Future Concerns. *Compr. Rev. Food Sci. Food Saf.* 9 (1), 57–81. doi:10.1111/j.1541-4337.2009.00094.x

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S.-M., et al. (1988). Single-chain Antigen-Binding Proteins. *Science* 242 (4877), 423–426. doi:10.1126/science.3140379

Burkin, A. A., and Kononenko, G. P. (2011). Enzyme Immunassay of Alternariol for the Assessment of Risk of Agricultural Products Contamination. *Appl. Biochem. Microbiol.* 47 (1), 72–76. doi:10.1134/S0003683811010030

Cai, P., Wang, R., Ling, S., and Wang, S. (2021). A High Sensitive Platinum-Modified Colloidal Gold Immunoassay for Tenuazonic Acid Detection Based on Monoclonal igG. *Food Chem.* 360 (9), 130021. doi:10.1016/j.foodchem.2021.130021

EFSA (European Food Safety Authority) (2011). Scientific Opinion on the Risks for Animal and Public Health Related to the Presence of Alternaria Toxins in Feed and Food. *EFSA J.* 9 (10), 2407. doi:10.2903/j.efsa.2011.2407

Estiarte, N., Crespo-Sempere, A., Marín, S., Ramos, A. J., and Worobo, R. W. (2018). Stability of Alternariol and Alternariol Monomethyl Ether during Food Processing of Tomato Products. *Food Chem.* 245, 951–957. doi:10.1016/j.foodchem.2017.11.078

Gambacorta, L., El Darra, N., Fakhoury, R., Logrieco, A. F., and Solfrizzo, M. (2019). Incidence and Levels of Alternaria Mycotoxins in Spices and Herbs Produced Worldwide and Commercialized in lebanon. *Food Control* 106, 106724. doi:10.1016/j.foodcont.2019.106724

Garcia, R., Carreiro, E. P., Prates Ramalho, J. P., Mirão, J., Burke, A. J., Gomes da Silva, M. D. R., et al. (2018). A Magnetic Controllable Tool for the Selective Enrichment of Dimethoate from Olive Oil Samples: A Responsive Molecular Imprinting-Based Approach. *Food Chem.* 254, 309–316. doi:10.1016/j.foodchem.2018.02.003

Gruber-Dorninger, C., Novak, B., Nagl, V., and Berthiller, F. (2017). Emerging Mycotoxins: Beyond Traditionally Determined Food Contaminants. *J. Agric. Food Chem.* 65 (33), 7052–7070. doi:10.1021/acs.jafc.6b03413

Humphrey, W., Dalke, A., and Schulten, K. (1996). Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* 14 (1), 33–38. doi:10.1016/0263-7855(96)00018-5

Huston, J. S., Mudgett-Hunter, M., Tai, M.-S., McCartney, J., Warren, F., Haber, E., et al. (1991). [3] Protein Engineering of Single-Chain Fv Analogs and Fusion Proteins. *Methods Enzymol.* 203, 46–88. doi:10.1016/0076-6879(91)03005-2

Kong, D., Xie, Z., Liu, L., Song, S., Zheng, Q., and Kuang, H. (2017). Development of an Immunochromatographic Assay for the Detection of Alternariol in Cereal and Fruit Juice Samples. *Food Agric. Immunol.* 28, 1082–1093. doi:10.1080/09540105.2017.1326469

Korpimäki, T., Rosenberg, J., Virtanen, P., Karskela, T., Lamminmäki, U., Tuomola, M., et al. (2002). Improving Broad Specificity Hapten Recognition with Protein Engineering. *J. Agric. Food Chem.* 50 (15), 4194–4201. doi:10.1021/jf0200624

Krausová, M., Ayeni, K. I., Wisgrill, L., Ezekiel, C. N., Braun, D., and Warth, B. (2021). Trace Analysis of Emerging and Regulated Mycotoxins in Infant Stool by LC-MS/MS. *Anal. Bioanal. Chem.* doi:10.1007/s00216-021-03803-9

Li, Y., Zhang, X., Nie, J., Bacha, S. A. S., Yan, Z., and Gao, G. (2020). Occurrence and Co-Occurrence of Mycotoxins in Apple and Apple Products from china. *Food Control* 118, 107354. doi:10.1016/j.foodcont.2020.107354

Li, L., Wang, X., Hou, R., Wang, Y., Wang, X., Xie, C., et al. (2022). Single-chain Variable Fragment Antibody-Based Ic-ELISA for Rapid Detection of Macrolides in Porcine Muscle and Computational Simulation of its Interaction Mechanism. *Food Control* 133, 108571. doi:10.1016/j.foodcont.2021.108571

Liang, Y., Wang, Y., Wang, F., Li, J., Wang, C., Dong, J., et al. (2021). An Enhanced Open sandwich Immunoassay by Molecular Evolution for Noncompetitive Detection of Alternaria Mycotoxin Tenuazonic Acid. *Food Chem.* 361 (1), 130103. doi:10.1016/j.foodchem.2021.130103

Liu, Y. H., Jin, M. J., Gui, W. J., Cheng, J. L., Guo, Y. R., and Zhu, G. N. (2007). Hapten Design and Indirect Competitive Immunoassay for Parathion Determination: Correlation with Molecular Modeling and Principal Component Analysis. *Anal. Chim. Acta* 591 (2), 173–182. doi:10.1016/j.aca.2007.03.071

Liu, Y., Liu, S., Xu, C., Lin, M., Li, Y., Shen, C., et al. (2021). Epitopes Prediction for Microcystin-LR by Molecular Docking. *Ecotoxicol. Environ. Saf.* 227, 112925. doi:10.1016/j.ecoenv.2021.112925

Lu, T., and Chen, F. (2012). Multiwfn: a Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* 33 (5), 580–592. doi:10.1002/jcc.22885

Man, Y., Liang, G., Jia, F., Li, A., Fu, H., Wang, M., et al. (2017). Development of an Immunochromatographic Strip Test for the Rapid Detection of Alternariol Monomethyl Ether in Fruit. *Toxins* 9 (5), 152. doi:10.3390/toxins9050152

Orina, A. S., Gavrilova, O. P., Gogina, N. N., Gannibal, P. B., and Gagkaeva, T. Y. (2021). Natural Occurrence of Alternaria Fungi and Associated Mycotoxins in Small-Grain Cereals from the Urals and West Siberia Regions of Russia. *Toxins* 13 (10), 681. doi:10.3390/toxins13100681

Quílez-Alburquerque, J., Descalzo, A. B., Moreno-Bondi, M. C., and Orellana, G. (2021). Luminescent Molecularly Imprinted Polymer Nanocomposites for Emission Intensity and Lifetime Rapid Sensing of Tenuazonic Acid Mycotoxin. *Polymer* 230, 124041. doi:10.1016/j.polymer.2021.124041

Singh, G., Velasquez, L., Brady, B., Koerner, T., Huet, A.-C., and Delahaut, P. (2018). Development of an Indirect Competitive Elisa for Analysis of Alternariol in Bread and Bran Samples. *Food Anal. Methods* 11, 1444–1450. doi:10.1007/s12161-017-1126-5

Solhaug, A., Torgersen, M. L., Holme, J. A., Lagadic-Gossmann, D., and Eriksen, G. S. (2014). Autophagy and Senescence, Stress Responses Induced by the DNA-Damaging Mycotoxin Alternariol. *Toxicology* 326, 119–129. doi:10.1016/j.tox.2014.10.009

Wang, J., Peng, T., Zhang, X., Yao, K., Ke, Y., Shao, B., et al. (2018). A Novel Hapten and Monoclonal Antibody-Based Indirect Competitive ELISA for Simultaneous Analysis of Alternariol and Alternariol Monomethyl Ether in Wheat. *Food Control* 94, 65–70. doi:10.1016/j.foodcont.2018.06.027

Wang, X., Chen, Q., Sun, Z., Wang, Y., Su, B., Zhang, C., et al. (2020). Nanobody Affinity Improvement: Directed Evolution of the Anti-ochratoxin a Single Domain Antibody. *Int. J. Biol. Macromol.* 151, 312–321. doi:10.1016/j.ijbiomac.2020.02.180

Wang, S., Gao, H., Wei, Z., Zhou, J., Ren, S., He, J., et al. (2022). Shortened and Multivalent Aptamers for Ultrasensitive and Rapid Detection of Alternariol in Wheat Using Optical Waveguide Sensors. *Biosens. Bioelectron.* 196, 113702. doi:10.1016/j.bios.2021.113702

Wen, K., Bai, Y., Wei, Y., Li, C., Shen, J., and Wang, Z. (2020). Influence of Small Molecular Property on Antibody Response. *J. Agric. Food Chem.* 68, 10944–10950. doi:10.1021/acs.jafc.0c04333

Xie, S., Wang, J., Yu, X., Peng, T., Yao, K., Wang, S., et al. (2020). Site-directed Mutations of Anti-amantadine scFv Antibody by Molecular Dynamics Simulation: Prediction and Validation. *J. Mol. Model.* 26 (3), 49. doi:10.1007/s00894-020-4286-y

Xu, Z.-L., Dong, J.-X., Wang, H., Li, Z.-F., Beier, R. C., Jiang, Y.-M., et al. (2012). Production and Characterization of a Single-Chain Variable Fragment Linked Alkaline Phosphatase Fusion Protein for Detection of O,o-Diethyl Organophosphorus Pesticides in a One-step Enzyme-Linked Immunosorbent Assay. *J. Agric. Food Chem.* 60 (20), 5076–5083. doi:10.1021/jf300570q

Zhang, F., Liu, B., Zhang, Y., Wang, J., Lu, Y., Deng, J., et al. (2019). Application of CdTe/CdS/ZnS Quantum Dot in Immunoassay for Aflatoxin B1 and Molecular Modeling of Antibody Recognition. *Anal. Chim. Acta* 1047, 139–149. doi:10.1016/j.aca.2018.09.058

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership