

Evaluating and Enhancing Liquid-Phase Exfoliation of Graphitic Carbon Nitride with Molecular  
Dynamics and Machine Learning

by

Ehsan Shahini

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering  
University of Alberta

© Ehsan Shahini, 2024

## Abstract

Graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) has gained significant attention due to its versatile applications in photocatalysis, energy conversion, and environmental remediation. To harness its full potential, high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets are essential, which can be achieved through liquid-phase exfoliation (LPE). However, the efficiency of LPE is critically influenced by the choice of solvent, and understanding the molecular interactions involved remains a challenge. This thesis investigates the LPE of g-C<sub>3</sub>N<sub>4</sub> using molecular dynamics (MD) simulations and machine learning (ML) techniques, with the aim to provide a comprehensive understanding on how the solvent selection and nanosheet functionalization impact exfoliation efficiency.

First, MD simulations were employed to evaluate the free energy of exfoliation ( $\Delta G_{\text{exf}}$ ) for g-C<sub>3</sub>N<sub>4</sub> in various solvents. The findings revealed that solvents with higher magnitude of solvation free energy ( $\Delta G_{\text{sol}}$ ) tend to facilitate exfoliation by forming stable adsorption layers around the nanosheets. To measure the stability, a quantity called solvent mobility was introduced, which benefited subsequent investigations in this thesis. Building on these insights, an ML model was developed to predict  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  for a wider range of solvents, significantly reducing the need for extensive MD simulations. The ML model identified several promising solvents, including benzyl alcohol (Bn), and methanesulfonic acid (MSA) which were experimentally validated to enhance the dispersibility of g-C<sub>3</sub>N<sub>4</sub>.

The thesis further explored the use of binary solvent mixtures, uncovering the phenomenon of solvent dominance where one component governs the LPE performance. Detailed MD simulations demonstrated that in N-Methyl-2-Pyrrolidone (NMP):Cyclohexane mixture, the performance was dominated by NMP, leading to superior exfoliation efficiency. Conversely, in

methanol:dichloromethane mixtures, methanol dominated, resulting in poor LPE performance. These findings challenge the traditional surface tension-based criteria for solvent selection, highlighting the need to consider molecular interactions in the vicinity of the nanosheets. The solvent dominance motivated a novel strategy for designing binary solvent mixture with balanced LPE performance, cost and environmental friendliness.

Finally, the impact of chemical functionalization on LPE efficiency was examined. Functional groups such as  $\text{SO}_3\text{H}$  and  $\text{COOH}$  were found to significantly enhance exfoliation by reducing  $\Delta G_{\text{exf}}$  and promoting stable solvent-sheet interactions. The  $\text{SO}_3\text{H}$  group, in particular, exhibited the most substantial enhancement due to strong dipole-dipole interactions, decreasing solvent mobility and increasing interaction strength around the nanosheets.  $\text{NH}_2$  and  $\text{OH}$  groups also contributed positively, though to a lesser extent, while  $\text{CHO}$  hindered the process by increasing  $\Delta G_{\text{exf}}$  and disrupting solvent-solvent interactions. These varied outcomes underscore the complexity of solvent-functional group interactions and provide practical guidelines for selecting effective solvents for functionalized  $\text{g-C}_3\text{N}_4$ .

Overall, this dissertation advances the understanding of molecular mechanisms in LPE of  $\text{g-C}_3\text{N}_4$ , providing practical guidelines for solvent selection and functionalization to optimize nanosheet production. The integration of MD simulations and ML models offers a powerful framework for predicting and enhancing the exfoliation process, paving the way for scalable and efficient production of  $\text{g-C}_3\text{N}_4$  for various technological applications.

## Preface

Chapter 3 of this thesis has been published under the citation: Shahini, E.; Shankar, K.; Tang, T. Liquid-Phase Exfoliation of Graphitic Carbon Nitrides Studied by Molecular Dynamics Simulation. *J Colloid Interface Sci* 2023, 630, 900–910. My responsibilities included the conceptualization, methodology design, use of software, validation, formal analysis, investigation, original draft writing, and visualization. Dr. Karthik Shankar contributed through conceptualization and the review and editing of the manuscript. As the supervisor, Dr. Tian Tang was involved in conceptualization, methodology, formal analysis, resource provision, manuscript review and editing, supervision, project administration, and funding acquisition.

Chapter 4 has been published as Shahini, E.; Chaulagain, N.; Shankar, K.; Tang, T. Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning. *ACS Appl Mater Interfaces* 2023, 15 (46), 53786–53801. In this work, I handled the conceptualization, methodology design, use of software, validation, formal analysis, investigation, original draft writing, and visualization. Narendra Chaulagain conducted the experiment, prepared the original draft, and assisted in visualization. Dr. Karthik Shankar contributed to the conceptual framework and manuscript review and editing. Dr. Tian Tang, in her supervisory role, provided resources, oversaw the project, performed formal analysis, reviewed and edited the manuscript, and secured funding.

A version of Chapter 5 is being prepared for publication as Shahini, E.; Shankar, K.; Tang, T. Hierarchy in Binary Liquid Phase Exfoliation of Graphitic Carbon Nitride: Dissecting the Dominance of One Solvent. My role was to design and conduct molecular simulations, analyze data, and draft the manuscript. Dr. Karthik Shankar supported conceptualization and manuscript review and editing. Dr. Tian Tang, as my supervisor, was instrumental in analyzing simulation results, verifying the accuracy of the results, revising the manuscript, and managing financial support and computational resources.

A version of Chapter 6 is in preparation for publication as Shahini, E.; Shankar, K.; Tang, T. Molecular Dynamics Analysis on Liquid-Phase Exfoliation of Functionalized Graphitic Carbon Nitride. As the lead author, I was responsible for designing, conducting, and analyzing simulations, conceptualizing the research question, generating figures, and drafting and revising the



manuscript. Dr. Karthik Shankar aided in conceptualization and manuscript review and editing. Dr. Tian Tang, my supervisor, contributed by analyzing simulation results, revising the manuscript, acquiring financial support, and arranging computational resources.

## Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Tian Tang, for her invaluable guidance, support, and encouragement throughout my research journey. Her insightful feedback and unwavering patience have been instrumental in shaping my academic and professional growth. I am truly fortunate to have had the opportunity to work under her mentorship.

I would also like to extend my heartfelt thanks to Dr. Karthik Shankar and Dr. Ge Li for their expert advice and invaluable contributions to this work. Their insightful comments and guidance have significantly enriched this thesis. Special thanks go to Narendra Chaulagain for his collaboration and assistance in experimental aspects and his unwavering support.

My deepest appreciation goes to my mother and family for their unconditional love and support. Their encouragement and belief in my abilities have been a constant source of motivation throughout my academic journey.

Finally, I would like to thank my dear friends Fazel Rangriz, Cathy Lam, Mandi Falcon, and Lauren Silva for their intellectual support and camaraderie. Their friendship and stimulating discussions have been invaluable to me during this journey.

# Table of Contents

Abstract.....	ii
Preface.....	iv
Acknowledgments.....	vi
List of Tables .....	xi
List of Figures .....	xii
Chapter 1: Introduction .....	1
1.1. Experimental studies.....	3
1.2. Computational studies.....	6
1.3. Knowledge gaps and objectives.....	11
1.4. Thesis outline .....	12
1.5. References.....	14
Chapter 2: Simulation Methodology.....	21
2.1. Basic principle of molecular dynamics simulation.....	21
2.1.1. NVE ensemble .....	21
2.1.2. NVT ensemble .....	23
2.1.3. NPT ensemble.....	24
2.2. Forcefields.....	24
2.2.1. Bonded interactions .....	25
2.2.2. Non-bonded interactions.....	25
2.3. Umbrella sampling.....	27
2.4. Radial distribution function .....	28
2.5. Machine learning .....	29
2.5.1. Supervised learning.....	29
2.5.2. Feature selection .....	30
2.5.3. Model training and validation.....	31
2.6. References.....	31
Chapter 3: Liquid-phase Exfoliation of Graphitic Carbon Nitrides.....	34
3.1. Introduction.....	34
3.2. Computational details .....	36
3.2.1. Molecular models for g-C <sub>3</sub> N <sub>4</sub> and solvents .....	36
3.2.2. Simulated systems.....	37

3.2.3.	Simulation details.....	41
3.3	Results and discussion .....	41
3.3.1.	Selection of reaction coordinate.....	41
3.3.2.	PMF and free energy of exfoliation .....	42
3.3.3.	Mechanisms governing the free energy of exfoliation .....	44
3.3.4.	Discussion .....	50
3.4.	Conclusion .....	54
3.5.	Acknowledgement .....	55
3.6.	References .....	55
Chapter 4: Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning.....		65
4.1.	Introduction.....	65
4.2.	Methods.....	67
4.2.1.	g-C <sub>3</sub> N <sub>4</sub> and solvents .....	67
4.2.2.	Dataset.....	68
4.2.3.	ML framework and models.....	70
4.2.4.	MD details.....	71
4.3.	Results.....	78
4.3.1.	Quality of dataset .....	78
4.3.2.	Model comparison .....	79
4.3.3.	Descriptor selection .....	83
4.3.4.	Model finalization.....	84
4.4.	Discussion .....	85
4.5.	Limitations and future considerations.....	93
4.6.	Conclusion .....	94
4.7.	Acknowledgement .....	95
4.8.	References .....	95
Chapter 5: Hierarchy in Binary Liquid Phase Exfoliation of Graphitic Carbon Nitride: Dissecting the Dominance of One Solvent.....		105
5.1.	Introduction.....	105
5.2.	Methods.....	106
5.2.1.	Structure of simulated g-C <sub>3</sub> N <sub>4</sub> and solvents.....	106
5.2.2.	Initial screening of binary mixtures .....	107

5.2.3.	Simulation details.....	109
5.3.	Results and discussion .....	110
5.3.1.	ML prediction of $\Delta G_{\text{exf}}$ in binary mixtures .....	110
5.3.2.	Validation of solvent dominance with MD.....	112
5.3.3.	Dominance of NMP in NMP:CH mixture .....	115
5.3.4.	Dominance of MET in MET:DCM mixture .....	121
5.3.5.	Application: eco-friendly LPE with water-organic solvent mixtures .....	123
5.4.	Discussion .....	125
5.5.	Conclusion .....	126
5.6.	Acknowledgement .....	126
5.7.	References.....	126
Chapter 6: Molecular Dynamics Analysis on Liquid-Phase Exfoliation of Functionalized Graphitic Carbon Nitride .....		132
6.1.	Introduction.....	132
6.2.	Methods.....	134
6.2.1.	Structures of functionalized g-C <sub>3</sub> N <sub>4</sub> .....	134
6.3.	Results.....	137
6.3.1.	$\Delta G_{\text{exf}}$ comparison for different functionalizations .....	137
6.3.2.	Solvent mobility around nanosheets .....	138
6.3.3.	Functionalization impacts interactions in both periphery and interior regions... ..	141
6.3.4.	Functionalization impacts nanosheet dipole moment .....	144
6.4.	Discussion .....	148
6.5.	Conclusion .....	150
6.6.	Acknowledgement .....	151
6.7.	References.....	151
Chapter 7: Conclusion and Future Perspective.....		156
7.1.	Overall conclusions.....	156
7.2.	Future perspectives .....	158
7.2.1.	Enhancing realism in simulations .....	158
7.2.2.	Advancing computational and experimental integration .....	158
7.2.3.	Exploring solvent mixtures .....	158
7.2.4.	Exploring functional group synergies .....	159
7.2.5.	Extending methodologies to other 2D materials.....	159

Unified Bibliography .....	160
Appendix A.....	197
Appendix B .....	216
Appendix C .....	236
Appendix D.....	241

## List of Tables

Table 3.1. Summary of the simulated systems. ....	41
Table 4.1. Summary of MD simulations performed in this study. ....	74
Table 4.2. Statistics of descriptors and target properties used for ML modeling. ....	79
Table 6.1. Best five solvents based on $M_{ave}$ results for each functional group.....	150

## List of Figures

Figure 1.1 (a) Triazine structure; (b) heptazine structure, with building blocks highlighted in red.	2
Figure 1.2. (a) Images of a graphene bilayer with NMP molecules at various inter-sheet distances, (b) normalized density profiles of NMP relative to bulk as functions of the z coordinate which is perpendicular to the graphene surface, (c) potential of mean force (PMF) between two graphene sheets (red curve), the number of NMP molecules confined between the sheets (blue curve), and the interaction potential between the graphene sheets without NMP molecules (dashed curve), and (d) PMF per unit area ( $\Phi$ ) between two parallel graphene sheets in five solvents (NMP, DMF, GBL, DMSO, and water) as functions of the inter-sheet distance. [Reprinted with permission from ref. [41]. Copyright © 2010 American Chemical Society.]	7
Figure 1.3. (a, b) Schematic illustration of parallel (shearing) and perpendicular (pulling) exfoliation pathways for h-BN. (b) Definitions of the reaction coordinates $d_y$ and $d_z$ , representing the slipping (shearing) and vertical (pulling) distances between the h-BN nanosheets. PMF per unit area for the parallel (c) and perpendicular (d) exfoliation of an h-BN nanosheet from a solvent-intercalated bilayer. [Reprinted with permission from ref. [44]. Copyright © 2017 American Chemical Society.]	8
Figure 1.4. (a) PMF profiles for the exfoliation of graphene sheets in binary NMP/water mixtures with varying NMP concentrations (0%, 13%, 25%, 65%, 80%, and 100%). (b) Monolayer adsorption of NMP molecules on the graphene sheet as a function of separation distance in different NMP concentrations, illustrating the increase in adsorption with higher NMP content. Shaded regions represent standard deviations. (c) Snapshot of the simulation box for the 13 wt% NMP concentration, showing the distribution of NMP and water molecules around the exfoliated graphene sheets. NMP molecules are shown with cyan color. [Reprinted with permission from ref. [45]. Copyright © 2024 John Wiley & Sons.]	10
Figure 2.1. Flowchart illustrating the step-by-step process of updating particle positions and velocities in a MD simulation.	22
Figure 3.1. Molecular structure of the 21-unit g-C <sub>3</sub> N <sub>4</sub> used in this study. The cyan circle shows one heptazine core unit and the red circle shows a vacancy that exists around the N atoms. Molecular structures of simulated solvents, categorized into group 1 (b), group 2 (c) and group 3	



(d) according to their structural characteristics. The molar weight of each solvent is given in the parenthesis.....	37
Figure 3.2. Three RCs explored for the separation of two g-C <sub>3</sub> N <sub>4</sub> nanosheets: (a) the distance between the COMs of two sheets along the normal direction, (b) the distance between the edge atoms of two sheets in the normal direction, and (c) the distance between the COMs of two nanosheets in the lateral direction. For better illustration, the two sheets are colored differently. The COMs or edge atoms are shown with blue and red circles respectively for the two sheets. (d) PMF curves for separating two stacked g-C <sub>3</sub> N <sub>4</sub> nanosheets in DMF along RC II and RC III. Insets show representative snapshots at different stages of the separation. ....	39
Figure 3.3. (a) PMF curves for separating two g-C <sub>3</sub> N <sub>4</sub> nanosheets in all solvents investigated in this study. (b) $\Delta G_{\text{exf}}$ vs. $\Delta G_{\text{sol}}$ for solvents from the three groups. Error bars in $\Delta G_{\text{sol}}$ are obtained by splitting the data into 5 blocks, calculating the free energy difference over each block, and estimating the error from the average variance over the blocks [54]. ....	44
Figure 3.4. (a) For group 1 solvent: energy changes in the solvation process associated with cavity generation ( $\Delta E_1$ ) and sheet insertion ( $\Delta E_2$ ). The sum of the two terms ( $\Delta E = \Delta E_1 + \Delta E_2$ ) along with $\Delta H_{\text{sol}}$ are also shown. (b) For group 1 and 2 solvents: $r_p$ (location of the first peak in the RDF) vs. $\Delta G_{\text{sol}}$ (left axis) and average H-bond lifetime vs. $\Delta G_{\text{sol}}$ (right axis) for group 1 and 2 solvents. Inset shows the partial atomic charges (PAC) of the solvent oxygen. (c) For group 3 solvents: probability distribution of the angle between the normal of the sheet and the normal of the aromatic rings of the solvent in the first solvation layer. ....	47
Figure 3.5. (a) $\ln(M_{\text{ave}})$ versus $\Delta G_{\text{sol}}/kT$ for all the solvents studied in this work. (b) Comparison of $\Delta H_{\text{eq.(3.1)}}$ and $\Delta G_{\text{exf}}$ from this study. ....	51
Figure 4.1. Molecular structure of the g-C <sub>3</sub> N <sub>4</sub> nanosheet used for the simulations in this study. The nanosheet contains 21 units of heptazine rings and is saturated with H atoms on the periphery. ....	68
Figure 4.2. ML framework employed in this work. The flowchart outlines the systematic approach taken for dataset preparation, model training and validation, descriptor selection to reduce overfitting, and model finalization. ....	71
Figure 4.3. Illustration of computational methodology and free energy definitions. (a) Histograms from US simulations for two g-C <sub>3</sub> N <sub>4</sub> sheets in chloroform. (b) PMF curve and definition of $\Delta G_{\text{exf}}$ for two g-C <sub>3</sub> N <sub>4</sub> sheets in chloroform. Corresponding PMF curve in vacuum is shown as a comparison to demonstrate lowering of $\Delta G_{\text{exf}}$ by chloroform. (c) The change in solvation free	

energy (SFE) as the vdW and electrostatic interactions are turned on, via 21 steps, between a g-C<sub>3</sub>N<sub>4</sub> sheet and chloroform. The sum of the changes is the total solvation free energy for this solvent and its absolute value is denoted by  $\Delta G_{\text{sol}}$  in this work. .... 75

Figure 4.4. Comparison of solvent properties in this work and those in Coleman et al. [41] (containing a larger dataset). The probability distributions are compared for (a) density, (b) dielectric constant and (c) surface tension of the organic solvents in both works. The good agreement confirms that the organic solvents in our dataset constitute a sufficient sample. .... 78

Figure 4.5. Correlation between predicted and measured values of free energies for the six ML models. All descriptors are included in the ML modeling. The scatter plots for (a)  $\Delta G_{\text{exf}}$  and (b)  $\Delta G_{\text{sol}}$  provide a visual assessment of model accuracy, which is quantified by  $R^2$  and RMSE values indicated in each subplot. ETR stands out as the most accurate model for both  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ . .... 82

Figure 4.6. Refinement of the ETR model. (a) RMSE and  $R^2$  score with all descriptors and the optimal subset (listed in figure) identified from ASM. The arrows indicate the improvement of prediction accuracy after descriptor selection. (b) Learning curves (with  $\Delta G_{\text{exf}}$  being the target property) using 10-fold CV and  $R^2$  as the accuracy metric. The green and blue shadows show the standard deviation for training and validation accuracy, respectively. Both curves plateau after the training size reaches 20. .... 84

Figure 4.7. Histogram of predicted  $\Delta G_{\text{exf}}$  for an expanded set of organic solvents. The prediction is made using the ETR model with the optimal subset of descriptors, for the 153 organic solvents listed in Appendix B section B4. The distribution shows that most  $\Delta G_{\text{exf}}$  values are in the range of 50-60 kJ/mol·nm<sup>2</sup>. .... 87

Figure 4.8. Macroscopic images demonstrating the dispersibility of bulk carbon nitride in various solvents. For each solvent, two sets of images are shown which correspond to different time (0 min: immediately after sonication; 60 min: 1 hour after sonication with no disturbance). Mass concentration of g-C<sub>3</sub>N<sub>4</sub> for each test is indicated below the respective image. ACN shows considerably poorer dispersibility compared with other solvents. .... 88

Figure 4.9. Correlation between measured  $\Delta G_{\text{exf}}$  and key descriptors. Pair plots of  $\Delta G_{\text{exf}}$  vs. (a)  $\Delta H_{\text{vap}}$ , (b)  $\varepsilon$  and (c)  $\gamma$ . Pair plots of  $\Delta G_{\text{sol}}$  vs. (d)  $C_{\text{max}}$ , (e)  $\alpha P$  and (f)  $\gamma$ .  $\Delta G_{\text{sol}}$  and  $\Delta G_{\text{exf}}$  are measured (from MD) values for the 48 organic solvents in our dataset. Five solvents with the lowest  $\Delta G_{\text{exf}}$  and highest  $\Delta G_{\text{sol}}$  are highlighted with red color; their numbering is consistent with

Appendix B section B1, and their name are given in (a) and (d). The correlation is linear in (a) and (d), nonlinear in (b) and (e), and unclear in (c) and (f). .....	91
Figure 4.10. Guidance on solvent selection for LPE of g-C <sub>3</sub> N <sub>4</sub> gained from ML modeling. (a) Proposed ranges of descriptors (shown with pattern bars), based on the solvents with predicted $\Delta G_{\text{exf}} < 51 \text{ kJ/mol}\cdot\text{nm}^2$ . The maximum and minimum of each descriptor is also shown for the experimental values of 153 organic solvents listed in Appendix B section B4. (b) Heatmap of SHAP values for each descriptor in the ETR model. The color intensity represents the magnitude of the SHAP value, with red indicating a positive impact and blue indicating a negative impact on the predicted $\Delta G_{\text{exf}}$ . The descriptors are ordered by their impact on $\Delta G_{\text{exf}}$ , with decreasing significance from top to bottom. ....	92
Figure 5.1. Molecular structure of the 15-unit g-C <sub>3</sub> N <sub>4</sub> alongside four solvents used in US simulations. The molecular mass of each solvent is provided in parentheses. For the nanosheet, atoms highlighted in cyan are considered as periphery, while the remaining atoms are considered as interior. ....	107
Figure 5.2. Flowchart illustrating steps taken in this work to examine LPE of g-C <sub>3</sub> N <sub>4</sub> by binary solvents. ....	109
Figure 5.3. (a) Normalized histograms and (b) scatter plot of $\Delta G_{\text{exf}}$ for the binary solvent mixtures as well as the arithmetic average of the two solvents in each binary mixture. ....	112
Figure 5.4. PMF profiles for different solvents and their mixture in LPE of g-C <sub>3</sub> N <sub>4</sub> : (a) NMP, MET, and their mixture, (b) NMP, CH, and their mixture, and (c) MET, DCM, and their mixture. The definitions of reaction coordinate and $\Delta G_{\text{exf}}$ are shown in (a). ....	114
Figure 5.5. Comparison of change in solvation free energy (SFE) as vdW and electrostatic interactions are turned on in twenty steps; (a) NMP, CH, and their mixture; (b) MET, DCM, and their mixture. ....	115
Figure 5.6. $M_{\text{ave}}$ of NMP, CH, MET and DCM in the binary mixtures and their singular counterparts. “NMP & CH in Mix” considers both NMP and CH molecules in the first solvation layer; the same applies to “MET & DCM in Mix”. ....	116
Figure 5.7. (a) Schematic illustration of the average percentage of NMP, CH, MET and DCM molecules in the first three solvation layers of two stacked nanosheets. (b) The corresponding density profiles of NMP, CH, MET, and DCM in their corresponding mixtures. Here, "Distance" is measured from the midpoint between the two nanosheets along the direction perpendicular to	

the nanosheet plane. All solvent molecules in the systems are considered in the calculation. Each density profile is symmetric about Distance = 0 and only half of the profile is shown.....	117
Figure 5.8. Probability distribution for the orientation of solvent molecules in the first solvation layer relative to the nanosheets. ....	118
Figure 5.9. (a) RDF plots showing the distributions of aromatic rings in NMP and CH around nitrogen in the two stacked g-C <sub>3</sub> N <sub>4</sub> nanosheets, in both singular and binary mixture systems. The distribution of oxygen in NMP around nitrogen in g-C <sub>3</sub> N <sub>4</sub> nanosheets is also shown for the binary mixture system; the corresponding result for pure NMP is very similar and hence not shown. (b) RDF plots showing the distributions of oxygen in MET and chlorine in DCM around nitrogen atoms of the two stacked g-C <sub>3</sub> N <sub>4</sub> nanosheets, in both singular and binary systems. The top inset shows a snapshot of the first solvation layer at t = 20 ns, with MET and DCM molecules in red and green, respectively. The H-bond network between MET molecules is shown with blue lines in the bottom inset.....	120
Figure 5.10. (a) $\Delta G_{\text{exf}}$ for three organic solvents and their mixture with water. $\Delta G_{\text{exf}}$ for pure water is also shown for comparison. (b) RDF of IPA oxygen around nitrogen atoms of the nanosheet, in pure IPA and in IPA:water mixture. Inset show a snapshot of the first solvation layer at t = 20 ns where the IPA and water molecules are shown with green and red color, respectively.....	124
Figure 6.1. The molecular structure of (a) pristine g-C <sub>3</sub> N <sub>4</sub> and g-C <sub>3</sub> N <sub>4</sub> functionalized with (b) sulfonic, (c) carboxyl, (d) amine, (e) hydroxyl, and (f) aldehyde. The atoms shown with green color are defined to form the “periphery” of the nanosheet. The remaining atoms are defined to belong to the “interior”. ....	135
Figure 6.2. (a) The PMF curves of pristine g-C <sub>3</sub> N <sub>4</sub> and g-C <sub>3</sub> N <sub>4</sub> functionalized with different groups. (b) The relationship between $\Delta G_{\text{exf}}$ and $M_{\text{ave}}$ for 1L and 2L configurations of each type of nanosheet.....	138
Figure 6.3. $M_{\text{ave}}$ and number of solvent molecules in the periphery and interior regions of the first solvation layer, for both 1L and 2L configurations. ....	140
Figure 6.4. (a) Relationship between the H-bond lifetime and $M_{\text{ave}}$ in 1L and 2L configurations. Inset shows the number of H-bonds formed with DMF. (b) RDF curves of oxygen in DMF around COM of functional groups for both 1L and 2L configurations. (c) RDF curves of oxygen in DMF around interior atoms of the nanosheet for both 1L and 2L configurations. ....	143

Figure 6.5. Probability distribution for the distance between nitrogen and oxygen atoms of different DMF molecules in the first solvation layer and bulk for (a) SO <sub>3</sub> H, (b) Pristine, and (c) CHO systems.....	147
Figure 6.6. (a) Probability distribution for the angle between DMF dipole moments and the nearest dipole moment in the periphery region of the nanosheet. (b) Top view of equilibrated nanosheets shown in (c) - DMF molecules are omitted for clarity. (c) Molecular configurations for 2L systems along with the first solvation layer for SO <sub>3</sub> H, pristine, and CHO. Inset in (a) shows an example for the angle definition. ....	148
Figure 6.7. (a) $\Delta G_{\text{exf}}$ values for 2L and 3L SO <sub>3</sub> H, Pristine, and CHO systems. (b) Histogram of $M_{\text{ave}}$ values for 50 common organic solvents tested with different functional groups in a 2L system. ....	149

# Chapter 1: Introduction

Nanotechnology has revolutionized material science. Advances in nanotechnology have catalyzed major progress in computing and electronics, yielding faster, smaller, and more portable systems capable of managing and storing vast amounts of information [1–4]. These advancements encompass transistors, magnetic random-access memory, and flash memory chips for smartphones [5]. As another example, nanoparticle copper suspensions serve as a safer, cost-effective, and reliable substitute for lead-based solder in electronics assembly [6]. Materials exhibit distinct physical properties at the nanoscale compared to their bulk counterparts. Certain nanoscale materials can spontaneously self-assemble into ordered structures [7]. Additionally, nanostructured materials boast a high surface-to-volume ratio, making them particularly suitable for interactions with other materials. This high surface-to-volume ratio is crucial since numerous reactions, both chemical and electrical, occur at interfaces and are heavily influenced by surface morphology and chemical composition [8]. Mastery over nanoscale interfacial properties is paramount in creating functional surfaces and nanomaterials. Manipulating nanoscale surface morphology and functionality demands precise control over electronic structures (e.g., molecular energy levels, dipole moments, polarizabilities, and optical properties) and intermolecular interactions [9].

A particular interest exists in exploring nanoscale properties to enhance energy efficiency [3]. Addressing the escalating global energy demand and environmental concerns caused by fossil fuel consumption has become a critical focus over recent decades [10]. Building on the unique properties of nanoscale materials, researchers are delving into the nanoscale realm to discover materials that promise to significantly boost the efficiency of contemporary energy sources and environmental protection [1,2,11]. Among these advancements, photocatalysis has emerged as a significant area of interdisciplinary interest [12–15]. Photocatalysis, driven by solar light, often employs semiconductors to facilitate catalytic reactions [16]. Notable examples include the splitting of water to produce hydrogen and oxygen [16], and the reduction of CO<sub>2</sub> to produce valuable hydrocarbons [17]. Developing a high-quality semiconductor photocatalyst to address energy shortages and environmental issues is a burgeoning research area. Among the promising

materials, graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) stands out as a next-generation photocatalyst, owing to its easy synthesis, attractive electronic band structure (normal band gap ~2.7 eV), high physicochemical stability, and abundance in nature [10,18]. Its two-dimensional structure, comprising carbon and nitrogen atoms arranged in a conjugated framework, imparts exceptional chemical stability and favorable electronic properties, which are crucial for photochemical applications. The structure of g-C<sub>3</sub>N<sub>4</sub> can be based on either triazine or tri-s-triazine (heptazine) units. Figure 1.1 illustrates these two structural types, highlighting their building blocks with a red circle. First-principle and experimental studies indicate that the heptazine configuration is more stable and therefore more likely under ambient conditions [19,20].

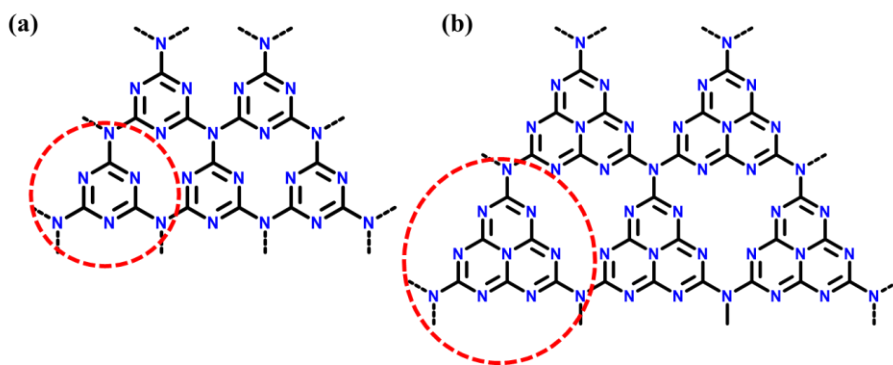


Figure 1.1 (a) Triazine structure; (b) heptazine structure, with building blocks highlighted in red.

Numerous studies have highlighted the advantages of ultrathin g-C<sub>3</sub>N<sub>4</sub> nanosheets, which offer higher specific surface areas and shorter charge carrier migration distances compared to bulk g-C<sub>3</sub>N<sub>4</sub> with polymeric layers [21]. Electrochemical impedance spectroscopy has shown that ultrathin g-C<sub>3</sub>N<sub>4</sub> nanosheets exhibit reduced electron transfer resistance [22]. Transient photocurrent measurements further revealed that these nanosheets produce a higher photocurrent. Additionally, photoluminescence studies indicated that ultrathin g-C<sub>3</sub>N<sub>4</sub> nanosheets have a lower recombination rate of photogenerated carriers than their bulk counterparts [23]. These findings collectively suggest that ultrathin g-C<sub>3</sub>N<sub>4</sub> nanosheets possess superior properties such as increased specific surface area, enhanced electron transport, and higher charge separation efficiency [21].

Exfoliation is a critical strategy to achieve the delamination of bulk g-C<sub>3</sub>N<sub>4</sub> into free-standing nanosheets [24]. g-C<sub>3</sub>N<sub>4</sub> nanosheets can be prepared by exfoliating bulk g-C<sub>3</sub>N<sub>4</sub>, akin to the exfoliation of two-dimensional (2D) graphene from bulk graphite. Different methods for

exfoliating g-C<sub>3</sub>N<sub>4</sub> include ultrasonication-assisted liquid-phase exfoliation (LPE), post-thermal oxidation etching, and combined sonication and thermal delamination. Among these, LPE is particularly advantageous due to its simplicity, scalability, and practical approach for producing high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets. LPE involves dispersing bulk g-C<sub>3</sub>N<sub>4</sub> in a liquid medium and applying ultrasonic energy to overcome the van der Waals forces between layers, leading to effective exfoliation. This method avoids the use of harsh chemicals and high temperatures, making it a more sustainable approach [25].

LPE of g-C<sub>3</sub>N<sub>4</sub> faces several limitations that can impact its effectiveness and scalability. One major challenge is the low exfoliation yield, which often results in a limited production of single or few-layer nanosheets [26]. The process is also time-consuming and energy-intensive, typically requiring prolonged ultrasonication or shear mixing [27]. The solvent type plays a crucial role in determining exfoliation efficiency, with certain solvents significantly enhancing the process while others hinder it [28]. Additionally, many effective solvents for LPE are expensive or toxic, raising concerns about cost-effectiveness and environmental impact [27]. The exfoliated nanosheets tend to restack or aggregate upon drying, reducing their effective surface area and compromising their performance in applications [29]. The process may also introduce defects or impurities, potentially affecting the material electronic and optical properties [30]. Given the complexity of interactions between solvents and g-C<sub>3</sub>N<sub>4</sub> nanosheets, both experimental and theoretical studies are crucial in elucidating the optimal conditions, especially the selection of solvents.

## 1.1. Experimental studies

The dispersion behavior of exfoliated g-C<sub>3</sub>N<sub>4</sub> in various solvents was evaluated by Ayán-Varela et al. [31] based on surface energy and Hildebrand/Hansen solubility parameters. These parameters, known as Hansen solubility parameters, include  $\delta D$  (dispersion forces),  $\delta P$  (polar forces), and  $\delta H$  (hydrogen bonding). Experimentally derived Hansen parameters for exfoliated g-C<sub>3</sub>N<sub>4</sub> yielded  $\delta D \approx 17.8 \text{ MPa}^{1/2}$ ,  $\delta P \approx 10.8 \text{ MPa}^{1/2}$ , and  $\delta H \approx 15.4 \text{ MPa}^{1/2}$ . The relatively high  $\delta H$  value highlighted the substantial role of H-bonding in the dispersion process, contrasting with other two-dimensional materials like graphene or transition metal dichalcogenides, where van der Waals forces dominate. This strong H-bonding was attributed to the high density of primary and



secondary amines in g-C<sub>3</sub>N<sub>4</sub>, resulting from incomplete condensation during synthesis. Solvents with high H-bonding parameters, such as diols and amine-containing compounds, were found to be particularly effective. Coleman et al. [32] proposed an equation to estimate the energy of exfoliation for graphite oxide as  $\frac{\Delta H_{mix}}{V_{mix}} = \frac{2}{T_{sheet}} (\sqrt{E_{sur}^{graphite}} - \sqrt{E_{sur}^{solvent}})^2 \phi$ , where  $\Delta H_{mix}$  is the enthalpy of mixing,  $T_{sheet}$  is the thickness of nanosheets,  $E_{sur}^i$  is the surface energy of each component, and  $\phi$  is the volume fraction of g-C<sub>3</sub>N<sub>4</sub>. They concluded that the exfoliation efficiency in a liquid is governed by  $\Delta H_{mix}$ . If one were to apply Coleman et al.'s theory to g-C<sub>3</sub>N<sub>4</sub>, then when the surface energies of bulk g-C<sub>3</sub>N<sub>4</sub> and the solvent are similar, the enthalpy of mixing would be low, facilitating successful exfoliation into 2D nanosheets. Water, with a surface energy of 102 mJ m<sup>-2</sup> [33], matches well with that of g-C<sub>3</sub>N<sub>4</sub> (115 mJ m<sup>-2</sup>), which would suggest water to be an effective dispersing solvent for g-C<sub>3</sub>N<sub>4</sub> nanosheets [34].

Zhang et al. [34] were the first to report the successful delamination of bulk layered g-C<sub>3</sub>N<sub>4</sub> into 2D nanosheets using a "green" liquid exfoliation method in water, while preserving the C–N chemical bonds within the g-C<sub>3</sub>N<sub>4</sub> plane. The resulting ultrathin nanosheets had a thickness of approximately 2.5 nm, consisting of around seven layers, and remained stably suspended in both acidic and alkaline environments for a week. However, the concentration of the resulting g-C<sub>3</sub>N<sub>4</sub> nanosheets in suspension was quite low, less than 0.15 mg mL<sup>-1</sup>. Yang et al. [35] conducted delamination experiments with five different solvents: water, ethanol, acetone, isopropanol (IPA), N-methyl-pyrrolidone (NMP). IPA and NMP outperformed water in terms of exfoliation efficiency, despite their much lower surface energies (~40 mJ m<sup>-2</sup>) compared to the calculated value for g-C<sub>3</sub>N<sub>4</sub>. These findings highlight the complexity of selecting an appropriate solvent for LPE and suggest that relying solely on surface energy matching is insufficient for determining the best solvent.

Building on this understanding, a promising approach involves using binary solvent mixtures. Shen et al. [36] investigated the efficacy of such mixtures by directly probing and matching the surface tension components for various 2D materials. Surface tension is divided into polar and dispersive components, and their study revealed that combining solvents with complementary polar and dispersive components significantly improved exfoliation outcomes. For instance, they found that a 1:1 mixture of IPA and water was highly effective for exfoliating graphene, WS<sub>2</sub>, h-BN, and MoSe<sub>2</sub>, while a 1:4 IPA/water mixture was optimal for Bi<sub>2</sub>Se<sub>3</sub> and SnS<sub>2</sub>.

Similarly, a 7:3 IPA/water mixture worked best for MoS<sub>2</sub>, and acetonitrile was the preferred solvent for TaS<sub>2</sub>. A similar approach was used for g-C<sub>3</sub>N<sub>4</sub>, where a mixed solvent method proved effective. Lin et al. [37] reported a versatile approach for the LPE of bulk g-C<sub>3</sub>N<sub>4</sub> to form monolayer g-C<sub>3</sub>N<sub>4</sub> nanosheets with tunable concentrations (0.1–3.0 mg mL<sup>-1</sup>). They examined different mixed solvents, including ethanol/water, IPA/water, and dimethylformamide (DMF)/water. In the ethanol/water system, reducing the volume ratio of ethanol led to a gradual increase in the suspension concentration. When the volume ratio of water reached 75%, a milky dispersion was achieved, with the highest concentration of g-C<sub>3</sub>N<sub>4</sub> at 3.0 mg mL<sup>-1</sup>. The Brunauer–Emmett–Teller (BET) surface area of the g-C<sub>3</sub>N<sub>4</sub> nanosheet was estimated to be about 59.4 m<sup>2</sup> g<sup>-1</sup>, which is approximately five times larger than that of bulk g-C<sub>3</sub>N<sub>4</sub> (12.5 m<sup>2</sup> g<sup>-1</sup>). Conversely, pure water resulted in a low g-C<sub>3</sub>N<sub>4</sub> concentration (about 0.5 mg mL<sup>-1</sup>), and pure ethanol produced an even lower concentration (about 0.2 mg mL<sup>-1</sup>). This demonstrates that the combination of two solvents can significantly enhance the formation of free-standing g-C<sub>3</sub>N<sub>4</sub> nanosheets with a high concentration.

Another effective method to enhance the dispersibility of g-C<sub>3</sub>N<sub>4</sub> in solvents is through functionalization. Kumru et al. [38] utilized a one-pot, visible light-induced grafting approach to improve the dispersibility of g-C<sub>3</sub>N<sub>4</sub> in both aqueous and organic media. They grafted functional groups, such as sulfonic acid (SO<sub>3</sub>H) and allylamine, onto the g-C<sub>3</sub>N<sub>4</sub> surface. This functionalization significantly enhanced hydrophilicity and dispersion stability, achieving up to 10% solid content in water and providing pH-dependent dispersibility. The SO<sub>3</sub>H groups increased the negative zeta potential, indicating better stability, while allylamine allowed for reversible dispersion in acidic (pH = 4) and basic (pH = 9) conditions.

The diverse experimental strategies explored for exfoliating g-C<sub>3</sub>N<sub>4</sub> highlight the complexity and multifaceted nature of achieving efficient dispersion and high-quality nanosheets. From leveraging Hansen solubility parameters to employing green liquid exfoliation methods, researchers have developed innovative approaches to enhance the LPE process. However, the variation in solvent performance across different studies underscores the need for a more nuanced understanding of solvent interactions. The use of binary solvent mixtures and functionalization techniques has shown significant promise, revealing that the combination of solvents and the introduction of specific functional groups can substantially improve exfoliation outcomes. These

findings suggest that a tailored approach, considering both the chemical nature of g-C<sub>3</sub>N<sub>4</sub> and the specific requirements of the application, is essential for optimizing liquid-phase exfoliation.

## 1.2. Computational studies

Computational modeling serves as a powerful tool for bridging the gap between experimental and theoretical understanding, particularly in the context of LPE. Firstly, it allows for the investigation of energy conversion dynamics, which is crucial for uncovering the underlying mechanisms of exfoliation processes [39]. Secondly, simulations can test different candidate molecules to evaluate their performance before their introduction into material synthesis [9]. Molecular dynamics (MD) simulations have been extensively used to model LPE for different layered nanomaterials such as graphene and boron nitride, providing molecular-level details of solvent-nanomaterial interactions, solvent ordering near surfaces, and intermolecular interactions in the immediate vicinity of such materials [40]. MD simulations can evaluate various dynamical properties that are not easily accessed through experimental techniques. Furthermore, MD can assess the thermodynamics of exfoliation and solvation processes using techniques like the potential of mean force (PMF) calculation and alchemical free energy perturbation (FEP) [40]. These computational approaches not only address the challenges associated with LPE but also predict new solvents and exfoliation protocols, enhancing our fundamental understanding of the underlying mechanisms.

One notable study by Shih et al. utilized MD simulations and the kinetic theory of colloid aggregation to understand the stabilization mechanism of graphene sheets in various organic solvents [41]. They examined solvents such as DMF, NMP, Dimethyl sulfoxide (DMSO), and  $\gamma$ -butyrolactone (GBL) to model the micromechanical processes during ultrasonication and agitation. By calculating the PMF between solvated graphene sheets, they discovered that the primary barrier preventing graphene aggregation was the last layer of confined solvent molecules between the sheets, resulting from the strong affinity of the solvent molecules to graphene (Figure 1.2a, b). This intercalation forms highly dense mono-to-few layers, which prevents the graphene sheets from aggregation. The stacked state corresponded to the free energy minimum in each solvent, but with varying degrees of stability (Figure 1.2c, d). As the vertical separation between the sheets increased and more solvent molecules intercalated, the stacked configuration became

less stable until a quasi-stable configuration was reached with a monolayer of solvent molecules between the sheets. Further separation led to multiple solvent layers, nearly nullifying the interlayer interactions. The solvents were ranked in terms of their ability to stabilize the dispersion as  $\text{NMP} \approx \text{DMSO} > \text{DMF} > \text{GBL} > \text{water}$ , which aligns with their common use in graphene dispersion (Figure 1.2d). Follow-up studies with phosphorene and  $\text{MoS}_2$  showed similar solvent intercalation within nanosheet gaps, indicating that solvent-induced stabilization favors exfoliation, though the ease of exfoliation varies with different solvents due to size and electrostatic effects [42,43].

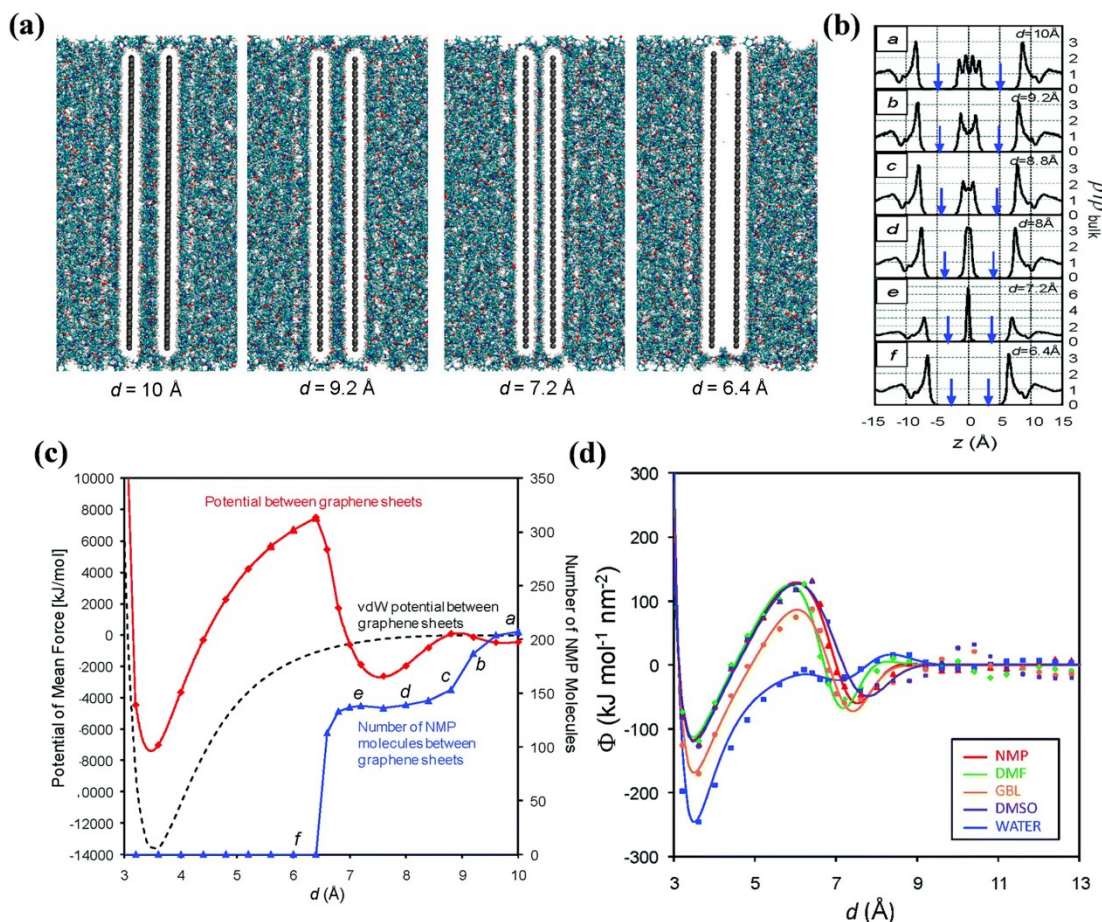


Figure 1.2. (a) Images of a graphene bilayer with NMP molecules at various inter-sheet distances, (b) normalized density profiles of NMP relative to bulk as functions of the  $z$  coordinate which is perpendicular to the graphene surface, (c) potential of mean force (PMF) between two graphene sheets (red curve), the number of NMP molecules confined between the sheets (blue curve), and the interaction potential between the graphene sheets without NMP molecules (dashed curve), and

(d) PMF per unit area ( $\Phi$ ) between two parallel graphene sheets in five solvents (NMP, DMF, GBL, DMSO, and water) as functions of the inter-sheet distance. [Reprinted with permission from ref. [41]. Copyright © 2010 American Chemical Society.]

Following the exploration of graphene, Mukhopadhyay and Datta [44] delved into the computational modeling of h-BN exfoliation using MD simulations. They investigated both parallel (shearing) and perpendicular (pulling) exfoliation pathways, calculating the PMF for each process (Figure 1.3a, b). The results showed that the free energy requirement for shearing is significantly lower than for pulling (Figure 1.3c, d). This is because, during shearing, most BN–BN interactions are retained, while in pulling, these interactions are drastically reduced as the layers move apart simultaneously. Polar solvents like DMSO, IPA, and water were found to be more effective in exfoliating h-BN compared to nonpolar solvents like dichloromethane (DCM) and dichloroethane (DCE), primarily due to stronger electrostatic interactions in addition to van der Waals forces. The PMF profiles indicated that polar solvents create a quasi-stable state with intercalated solvent layers, which aids in reducing the energy required for exfoliation.

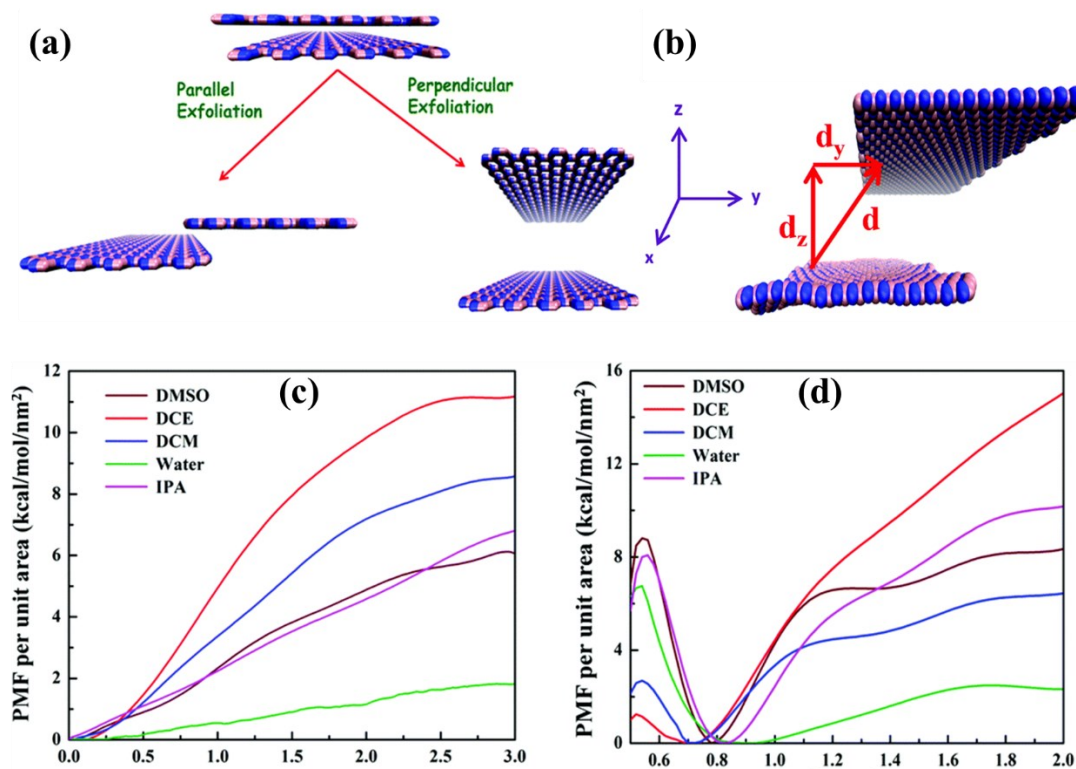


Figure 1.3. (a, b) Schematic illustration of parallel (shearing) and perpendicular (pulling) exfoliation pathways for h-BN. (b) Definitions of the reaction coordinates  $d_y$  and  $d_z$ , representing

the slipping (shearing) and vertical (pulling) distances between the h-BN nanosheets. PMF per unit area for the parallel (c) and perpendicular (d) exfoliation of an h-BN nanosheet from a solvent-intercalated bilayer. [Reprinted with permission from ref. [44]. Copyright © 2017 American Chemical Society.]

As previously discussed, a promising experimental strategy involves utilizing binary solvent mixtures, which have demonstrated significant improvements in the exfoliation and stabilization of graphene nanosheets. Gotzias and Lazarou utilized MD simulations to investigate the exfoliation of graphene in NMP/water mixtures. They computed the PMF for exfoliation, considering both shearing and pulling pathways. Similar to the work of Mukhopadhyay and Datta [44], the study found that the shearing pathway, where the graphene sheet is pulled parallel to its plane, is the preferred reaction coordinate, requiring less free energy compared to the pulling pathway. The results demonstrated that NMP substantially decreases the free energy barrier for exfoliation, emphasizing its superior solubilizing efficacy compared to water, as shown in Figure 1.4a. Remarkably, the optimal solvent mixture was found to be 70-80% NMP, which not only improved exfoliation efficiency but also addressed environmental concerns associated with using pure organic solvents. They also found as the graphene sheets exfoliate, they become coated with a dense monolayer of NMP, which protects the sheets from re-aggregation. Figure 1.4b illustrates the monolayer adsorption of NMP molecules on the graphene sheet, where the number of adsorbed NMP molecules increases with higher NMP concentrations. The snapshot at the bottom of Figure 1.4c visually demonstrates the distribution of NMP molecules around the exfoliating graphene.

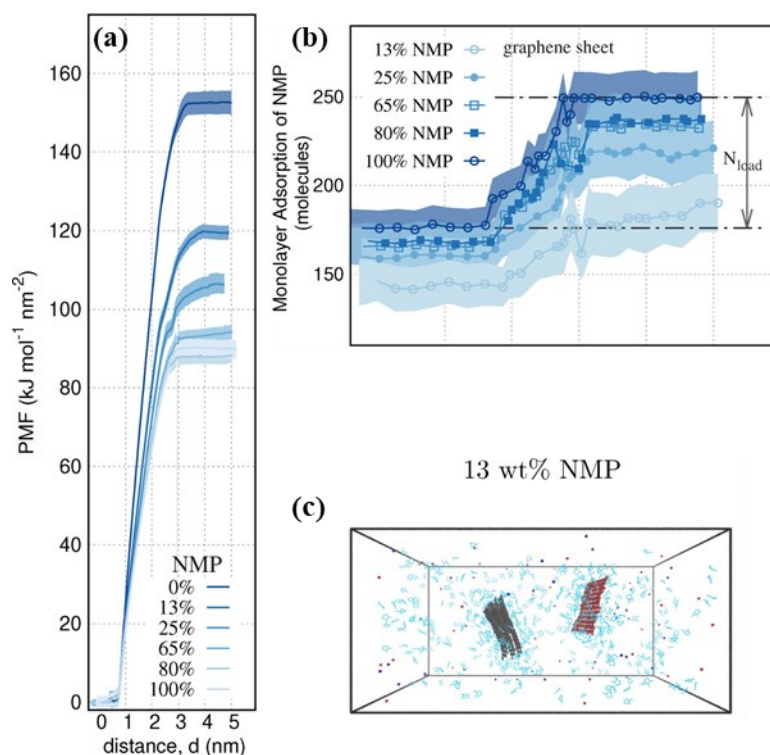


Figure 1.4. (a) PMF profiles for the exfoliation of graphene sheets in binary NMP/water mixtures with varying NMP concentrations (0%, 13%, 25%, 65%, 80%, and 100%). (b) Monolayer adsorption of NMP molecules on the graphene sheet as a function of separation distance in different NMP concentrations, illustrating the increase in adsorption with higher NMP content. Shaded regions represent standard deviations. (c) Snapshot of the simulation box for the 13 wt% NMP concentration, showing the distribution of NMP and water molecules around the exfoliated graphene sheets. NMP molecules are shown with cyan color. [Reprinted with permission from ref. [45]. Copyright © 2024 John Wiley & Sons.]

Among the computational studies available, the work by Zou et al. [46] is the first that used MD simulations to investigate the exfoliation mechanisms of  $\text{g-C}_3\text{N}_4$  nanosheets. Their approach involved a three-step process: first, the experimental preparation and initial exfoliation of bulk  $\text{g-C}_3\text{N}_4$  using various solvents; second, the utilization of MD simulations to analyze the interactions between the solvent molecules and  $\text{g-C}_3\text{N}_4$  nanosheets, identifying the energy barriers involved in the exfoliation process; and third, re-carrying out the exfoliation experimentally based on the simulation results to optimize the yield of exfoliated  $\text{g-C}_3\text{N}_4$  nanosheets. The simulations revealed

that solvent molecules intercalate into the g-C<sub>3</sub>N<sub>4</sub> layers, lowering the free energy barriers for exfoliation, thus enhancing the efficiency of the process. By examining the kinetic energy changes of the nanosheets, they introduced a method to estimate energy barriers for exfoliation. Their results demonstrated that the most suitable solvent mixtures for exfoliating g-C<sub>3</sub>N<sub>4</sub> were those with surface energies close to that of the g-C<sub>3</sub>N<sub>4</sub> nanosheets, with the mixture of water and IPA (mass ratio 1:2) predicted to be particularly effective. This computational prediction was experimentally validated, achieving a g-C<sub>3</sub>N<sub>4</sub> suspension concentration of 5.03 mg mL<sup>-1</sup>.

While the study by Zou et al. [46] provides insights into the exfoliation process of g-C<sub>3</sub>N<sub>4</sub>, several aspects need to be examined. Firstly, the reliance on surface energy matching as the primary criterion for solvent selection, although effective in their study, might not capture the complete complexity of solvent-nanosheet interactions. This approach assumes that surface energy is the dominant factor, potentially overlooking other critical properties such as solvent density, dielectric constant, and specific molecular interactions that could also influence exfoliation efficiency. Furthermore, the method introduced for estimating energy barriers based on kinetic energy changes requires further validation against more established techniques such as PMF calculations to ensure its robustness and accuracy.

### **1.3. Knowledge gaps and objectives**

The efficiency of producing high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets through LPE is critically dependent on various factors such as solvent selection, molecular interactions, and functionalization. LPE, a widely used method for synthesizing 2D g-C<sub>3</sub>N<sub>4</sub>, requires a detailed molecular-level understanding to optimize the process effectively. Current experimental techniques often fall short in providing such insights, making MD simulations a valuable tool for exploring the microscopic mechanisms of LPE. MD simulations offer a powerful approach to study the LPE process at a microscopic level, providing detailed information on the interactions between g-C<sub>3</sub>N<sub>4</sub> nanosheets and various solvents. However, these simulations are resource-intensive and time-consuming when applied to a wide range of solvents. Machine learning (ML) algorithms can complement MD simulations by predicting the free energy of exfoliation and solvation for numerous solvents based on a limited dataset. This combination of MD and ML



enables the efficient screening and selection of optimal solvents for LPE, addressing the limitations of purely experimental or computational methods.

Additionally, the performance of binary solvent mixtures in LPE introduces an additional layer of complexity that has not been sufficiently explored. Existing studies have not thoroughly investigated why some mixtures perform better than others or how the dominance of one solvent component affects the LPE process. This lack of understanding limits the ability to optimize solvent mixtures for improved exfoliation performance. While chemical functionalization of g-C<sub>3</sub>N<sub>4</sub> nanosheets has been shown to enhance solvent-sheet interactions and reduce the energy required for exfoliation, systematic studies investigating the effects of different functional groups on these processes are scarce. The literature lacks detailed MD simulations that analyze the specific interactions between functionalized g-C<sub>3</sub>N<sub>4</sub> and various solvents, which are crucial for optimizing functionalization strategies.

This thesis will extend the theoretical understanding of LPE for g-C<sub>3</sub>N<sub>4</sub> through the following objectives:

- (1) develop MD simulation frameworks to evaluate the free energy of exfoliation and solvation for g-C<sub>3</sub>N<sub>4</sub> in various solvents;
- (2) integrate ML algorithms to predict these free energies across a wide range of solvents, reducing the need for extensive simulations;
- (3) investigate the performance of binary solvent mixtures and their molecular interactions to optimize LPE efficiency;
- (4) explore the impact of chemical functionalization on the LPE efficiency of g-C<sub>3</sub>N<sub>4</sub> and provide guidelines for effective functionalization strategies.

Achieving these is expected to advance the theoretical and practical understanding of LPE, contributing to the development of efficient methods for producing high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets.

## 1.4. Thesis outline

**Chapter 2** presents the simulation methodology used throughout this research. This chapter includes a brief description of the MD simulation techniques, the all-atom forcefields

employed, potential of mean force calculations, and radial distribution functions which is a key analysis used to identify major molecular interactions.

**Chapter 3** investigates LPE of g-C<sub>3</sub>N<sub>4</sub> nanosheets in pure solvents using MD simulations. The chapter examines the molecular mechanisms of exfoliation, focusing on the evaluation of free energy in various solvents through PMF calculations. It analyzes how solvent properties affect exfoliation efficiency and develops guidelines for selecting effective solvents to optimize g-C<sub>3</sub>N<sub>4</sub> nanosheet production.

**Chapter 4** delves into the integration of ML with MD simulations to predict the free energy of exfoliation and solvation for g-C<sub>3</sub>N<sub>4</sub> in various solvents. It describes the development and validation of ML models, based on extensive MD simulation data, and demonstrates how these models can efficiently and accurately predict solvent performance, thereby reducing the need for resource-intensive simulations.

**Chapter 5** focuses on the dynamics of binary solvent mixtures in LPE of g-C<sub>3</sub>N<sub>4</sub>. The chapter explores how the exfoliation performance can be dominated by one solvent component. MD simulations and ML predictions are used to investigate the molecular interactions within binary mixtures, providing insights into optimizing these mixtures for improved LPE performance, and reduced cost and environmental impact.

**Chapter 6** examines the impact of chemical functionalization on the LPE efficiency of g-C<sub>3</sub>N<sub>4</sub> nanosheets. Various functional groups, such as sulfonic, carboxyl, amine, hydroxyl, and aldehyde, are studied using MD simulations. The chapter provides practical guidelines for selecting and designing functionalized g-C<sub>3</sub>N<sub>4</sub> nanosheets to enhance solvent-sheet interactions and reduce the free energy required for exfoliation.

**Chapter 7** summarizes the key findings, discusses the implications of the research, and suggests directions for future studies to further advance the field of g-C<sub>3</sub>N<sub>4</sub> nanosheet production and application.

## 1.5. References

- [1] M.H. Ahmadi, M. Ghazvini, M.A. Nazari, M.A. Ahmadi, F. Pourfayaz, G. Lorenzini, T. Ming, Renewable energy harvesting with the application of nanotechnology: A review, *Int J Energy Res* 43 (2019) 1387–1410. <https://doi.org/10.1002/ER.4282>.
- [2] G. Ali Mansoori, T.R. Bastami, A. Ahmadpour, Z. Eshaghi, ENVIRONMENTAL APPLICATION OF NANOTECHNOLOGY, (2008) 439–493. [https://doi.org/10.1142/9789812790248\\_0010](https://doi.org/10.1142/9789812790248_0010).
- [3] P.K. Sharma, S. Dorlikar, P. Rawat, V. Malik, N. Vats, M. Sharma, J.S. Rhyee, A.K. Kaushik, Nanotechnology and its application: a review, *Nanotechnology in Cancer Management: Precise Diagnostics toward Personalized Health Care* (2021) 1–33. <https://doi.org/10.1016/B978-0-12-818154-6.00010-X>.
- [4] M.L. Cohen, Nanotubes, nanoscience, and nanotechnology, *Materials Science and Engineering: C* 15 (2001) 1–11.
- [5] M.T. Bohr, Nanotechnology goals and challenges for electronic applications, *IEEE Trans Nanotechnol* 1 (2002) 56–62. <https://doi.org/10.1109/TNANO.2002.1005426>.
- [6] T. Kida, T. Oka, M. Nagano, Y. Ishiwata, X.G. Zheng, Synthesis and Application of Stable Copper Oxide Nanoparticle Suspensions for Nanoparticulate Film Fabrication, *Journal of the American Ceramic Society* 90 (2007) 107–110. <https://doi.org/10.1111/J.1551-2916.2006.01402.X>.
- [7] D.O. Hutchins, T. Weidner, J. Baio, B. Polishak, O. Acton, N. Cernetic, H. Ma, A.K.Y. Jen, Effects of self-assembled monolayer structural order, surface homogeneity and surface energy on pentacene morphology and thin film transistor device performance, *J Mater Chem C Mater* 1 (2013) 101–113. <https://doi.org/10.1039/c2tc00378c>.
- [8] R.N. Kostoff, R.G. Koytcheff, C.G.Y. Lau, Global nanotechnology research literature overview, *Technol Forecast Soc Change* 74 (2007) 1733–1747. <https://doi.org/10.1016/j.techfore.2007.04.004>.

- [9] J. Wen, W. Li, S. Chen, J. Ma, Simulations of molecular self-assembled monolayers on surfaces: Packing structures, formation processes and functions tuned by intermolecular and interfacial interactions, *Physical Chemistry Chemical Physics* 18 (2016) 22757–22771. <https://doi.org/10.1039/c6cp01049k>.
- [10] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.
- [11] C. George, M. Ammann, B. D’Anna, D.J. Donaldson, S.A. Nizkorodov, Heterogeneous Photochemistry in the Atmosphere, *Chem Rev* 115 (2015) 4218–4258. <https://doi.org/10.1021/cr500648z>.
- [12] M. Kapilashrami, Y. Zhang, Y.S. Liu, A. Hagfeldt, J. Guo, Probing the optical property and electronic structure of TiO<sub>2</sub>nanomaterials for renewable energy applications, *Chem Rev* 114 (2014) 9662–9707. <https://doi.org/10.1021/cr5000893>.
- [13] A.J. Bard, M.A. Fox, Artificial Photosynthesis: Solar Splitting of Water to Hydrogen and Oxygen, *Acc Chem Res* 28 (1995) 141–145. <https://doi.org/10.1021/ar00051a007>.
- [14] Z. Zou, J. Ye, K. Sayama, H. Arakawa, Direct splitting of water under visible light irradiation with an oxide semiconductor photocatalyst, *Nature* 414 (2001) 625–627. <https://doi.org/10.1038/414625a>.
- [15] P. Kumar, P. Kar, A.P. Manuel, S. Zeng, U.K. Thakur, K.M. Alam, Y. Zhang, R. Kisslinger, K. Cui, G.M. Bernard, V.K. Michaelis, K. Shankar, Noble Metal Free, Visible Light Driven Photocatalysis Using TiO<sub>2</sub> Nanotube Arrays Sensitized by P-Doped C<sub>3</sub>N<sub>4</sub> Quantum Dots, *Adv Opt Mater* 1901275 (2019) 1–15. <https://doi.org/10.1002/adom.201901275>.
- [16] Q. Wang, T. Hisatomi, Q. Jia, H. Tokudome, M. Zhong, C. Wang, Z. Pan, T. Takata, M. Nakabayashi, N. Shibata, Y. Li, I.D. Sharp, A. Kudo, T. Yamada, K. Domen, Scalable water splitting on particulate photocatalyst sheets with a solar-to-hydrogen energy conversion efficiency exceeding 1%, (2016). <https://doi.org/10.1038/NMAT4589>.

- [17] J. Wang, T. Heil, B. Zhu, C.W. Tung, J. Yu, H.M. Chen, M. Antonietti, S. Cao, A Single Cu-Center Containing Enzyme-Mimic Enabling Full Photosynthesis under CO<sub>2</sub>Reduction, *ACS Nano* 14 (2020) 8584–8593.  
<https://doi.org/10.1021/acsnano.0c02940>.
- [18] P. Kumar, E. Vahidzadeh, U.K. Thakur, P. Kar, K.M. Alam, A. Goswami, N. Mahdi, K. Cui, G.M. Bernard, V.K. Michaelis, K. Shankar, C<sub>3</sub>N<sub>5</sub>: A Low Bandgap Semiconductor Containing an Azo-Linked Carbon Nitride Framework for Photocatalytic, Photovoltaic and Adsorbent Applications, *J Am Chem Soc* 141 (2019) 5415–5436.  
<https://doi.org/10.1021/jacs.9b00144>.
- [19] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [20] Y. Xu, S.P. Gao, Band gap of C<sub>3</sub>N<sub>4</sub> in the GW approximation, *Int J Hydrogen Energy* 37 (2012) 11072–11080. <https://doi.org/10.1016/J.IJHYDENE.2012.04.138>.
- [21] S. Cao, J. Low, J. Yu, M. Jaroniec, Polymeric Photocatalysts Based on Graphitic Carbon Nitride, *Advanced Materials* 27 (2015) 2150–2176.  
<https://doi.org/10.1002/adma.201500033>.
- [22] J. Liu, Q. Huang, Z. Chen, L. Yao, L. Huang, C. Dong, L. Niu, Y. Zhang, Electrochemical impedance spectroscopy study on polysulfone/g-C<sub>3</sub>N<sub>4</sub> composite membrane during the separation process, (2020). <https://doi.org/10.5004/dwt.2020.24997>.
- [23] Y. Yuan, L. Zhang, J. Xing, M.I.B. Utama, X. Lu, K. Du, Y. Li, X. Hu, S. Wang, A. Genç, R. Dunin-Borkowski, J. Arbiol, Q. Xiong, High-yield synthesis and optical properties of g-C<sub>3</sub>N<sub>4</sub>, *Nanoscale* 7 (2015) 12343–12350.  
<https://doi.org/10.1039/c5nr02905h>.
- [24] Y. Zheng, L. Lin, B. Wang, X. Wang, Graphitic Carbon Nitride Polymers toward Sustainable Photoredox Catalysis, *Angewandte Chemie International Edition* 54 (2015) 12868–12884. <https://doi.org/10.1002/anie.201501788>.

- [25] A. Amiri, M. Naraghi, G. Ahmadi, M. Soleymaniha, M. Shanbedi, A review on liquid-phase exfoliation for scalable production of pure graphene, wrinkled, crumpled and functionalized graphene and challenges, *FlatChem* 8 (2018) 40–71.  
<https://doi.org/10.1016/J.FLATC.2018.03.004>.
- [26] F. Goni, A. Chemelli, F. Uhlig, High-Yield Production of Selected 2D Materials by Understanding Their Sonication-Assisted Liquid-Phase Exfoliation, *Nanomaterials* 2021, Vol. 11, Page 3253 11 (2021) 3253. <https://doi.org/10.3390/NANO11123253>.
- [27] Y. Xu, H. Cao, Y. Xue, B. Li, W. Cai, Liquid-Phase Exfoliation of Graphene: An Overview on Exfoliation Media, Techniques, and Challenges, *Nanomaterials* 8 (2018).  
<https://doi.org/10.3390/NANO8110942>.
- [28] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenis, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045.  
[https://doi.org/10.1021/ACSAMI.5B06974/SUPPL\\_FILE/AM5B06974\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.5B06974/SUPPL_FILE/AM5B06974_SI_001.PDF).
- [29] C. Backes, D. Hanlon, B.M. Szydłowska, A. Harvey, R.J. Smith, T.M. Higgins, J.N. Coleman, Preparation of Liquid-exfoliated Transition Metal Dichalcogenide Nanosheets with Controlled Size and Thickness: A State of the Art Protocol, *J Vis Exp* 2016 (2016) 54806. <https://doi.org/10.3791/54806>.
- [30] M. V. Bracamonte, G.I. Lacconi, S.E. Urreta, L.E.F. Foa Torres, On the nature of defects in liquid-phase exfoliated graphene, *Journal of Physical Chemistry C* 118 (2014) 15455–15459. [https://doi.org/10.1021/JP501930A/SUPPL\\_FILE/JP501930A\\_SI\\_001.PDF](https://doi.org/10.1021/JP501930A/SUPPL_FILE/JP501930A_SI_001.PDF).
- [31] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenis, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045.

[https://doi.org/10.1021/ACSAMI.5B06974/ASSET/IMAGES/LARGE/AM-2015-06974S\\_0012.JPEG](https://doi.org/10.1021/ACSAMI.5B06974/ASSET/IMAGES/LARGE/AM-2015-06974S_0012.JPEG).

- [32] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuvsen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571.  
<https://doi.org/10.1126/science.1194975>.
- [33] J. Zhang, Y. Chen, X. Wang, Two-dimensional covalent carbon nitride nanosheets: Synthesis, functionalization, and applications, *Energy Environ Sci* 8 (2015) 3092–3108.  
<https://doi.org/10.1039/c5ee01895a>.
- [34] X. Zhang, X. Xie, H. Wang, J. Zhang, B. Pan, Y. Xie, Enhanced photoresponsive ultrathin graphitic-phase C<sub>3</sub>N<sub>4</sub> nanosheets for bioimaging, *J Am Chem Soc* 135 (2013) 18–21.  
<https://doi.org/10.1021/ja308249k>.
- [35] S. Yang, Y. Gong, J. Zhang, L. Zhan, L. Ma, Z. Fang, R. Vajtai, X. Wang, P.M. Ajayan, Exfoliated graphitic carbon nitride nanosheets as efficient catalysts for hydrogen evolution under visible light, *Advanced Materials* 25 (2013) 2452–2456.  
<https://doi.org/10.1002/adma.201204453>.
- [36] J. Shen, Y. He, J. Wu, C. Gao, K. Keyshar, X. Zhang, Y. Yang, M. Ye, R. Vajtai, J. Lou, P.M. Ajayan, Liquid Phase Exfoliation of Two-Dimensional Materials by Directly Probing and Matching Surface Tension Components, *Nano Lett* 15 (2015) 5449–5454.  
[https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E\\_0005.JPEG](https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E_0005.JPEG).
- [37] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142.  
<https://doi.org/10.1016/j.apcatb.2014.07.053>.

- [38] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906.  
[https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL\\_FILE/LA7B02441\\_SI\\_007.AVI](https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL_FILE/LA7B02441_SI_007.AVI).
- [39] C.Y. Wong, B.L. Cotts, H. Wu, N.S. Ginsberg, Exciton dynamics reveal aggregates with intermolecular order at hidden interfaces in solution-cast organic semiconducting films, *Nat Commun* 6 (2015) 1–7. <https://doi.org/10.1038/ncomms6946>.
- [40] T.K. Mukhopadhyay, A. Datta, Disentangling the liquid phase exfoliation of two-dimensional materials: an “in silico” perspective, *Physical Chemistry Chemical Physics* 22 (2020) 22157–22179. <https://doi.org/10.1039/D0CP03128C>.
- [41] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648.  
<https://doi.org/10.1021/ja1064284>.
- [42] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268.  
<https://doi.org/10.1021/acsnano.5b02683>.
- [43] V. Sresht, A. Govind Rajan, E. Bordes, M.S. Strano, A.A.H. Pádua, D. Blankschtein, Quantitative Modeling of MoS<sub>2</sub>-Solvent Interfaces: Predicting Contact Angles and Exfoliation Performance using Molecular Dynamics, *Journal of Physical Chemistry C* 121 (2017) 9022–9031.  
[https://doi.org/10.1021/ACS.JPCC.7B00484/SUPPL\\_FILE/JP7B00484\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCC.7B00484/SUPPL_FILE/JP7B00484_SI_001.PDF).
- [44] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.



- [45] A. Gotzias, Y.G. Lazarou, Graphene Exfoliation in Binary NMP/Water Mixtures by Molecular Dynamics Simulations, *Chempluschem* 89 (2024) e202300758. <https://doi.org/10.1002/CPLU.202300758>.
- [46] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight from molecular dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954. <https://doi.org/10.1007/s10934-021-01047-7>.

# Chapter 2: Simulation Methodology

## 2.1. Basic principle of molecular dynamics simulation

MD simulation is grounded in Newton's second law of motion,  $\mathbf{F} = m\mathbf{a}$ , where  $\mathbf{F}$  denotes the force acting on a particle,  $m$  is the mass of the particle, and  $\mathbf{a}$  is the acceleration [1]. For conservative forces, the force can be expressed as the negative gradient of a potential energy function,  $U$ , leading to  $\mathbf{F} = -\nabla U$ . This relationship is pivotal for classical MD simulations where intricate intermolecular potentials are employed [1]. In an MD simulation, Newton's second law is numerically integrated over time to update the positions and velocities of all particles in the system.

Various statistical ensembles can be employed in MD simulations depending on the physical conditions and properties one wishes to study [2]. Each ensemble is characterized by different constraints on the system's thermodynamic variables. The most commonly used ensembles in MD simulations are the microcanonical (NVE), canonical (NVT), and isothermal-isobaric (NPT) ensembles [2]. These ensembles derive from the fundamental principles of statistical mechanics and provide different ways to sample the phase space.

### 2.1.1. NVE ensemble

The microcanonical ensemble, also known as the NVE ensemble, is characterized by a constant number of particles, volume, and energy. In this ensemble, the total energy of the system is conserved, and the system is isolated with no exchange of energy or matter with the surroundings. The fundamental equation for the microcanonical ensemble is the conservation of total energy [3]:

$$E = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U \quad (2.1)$$

where  $\mathbf{p}_i = m_i\mathbf{v}_i$  is the linear momentum of particle  $i$ ,  $m_i$  and  $\mathbf{v}_i$  being the mass and velocity of particle  $i$  respectively,  $U$  is the system's total potential energy, and  $N$  is the total number of particles.

The NVE ensemble is useful for studying the intrinsic dynamics of isolated systems without external influences [3]. Because the total energy is conserved, second time derivative of Eq. (2.1) results in

$$\frac{d\mathbf{v}_i}{dt} = \mathbf{a}_i = \frac{\mathbf{F}_i}{m_i} = -\frac{1}{m_i} \nabla_{\mathbf{r}_i} U \quad (2.2)$$

where  $\mathbf{F}_i = -\nabla_{\mathbf{r}_i} U$  is the force on particle  $i$  and  $\nabla_{\mathbf{r}_i}$  indicates the gradient taken with respect to the position  $\mathbf{r}_i$  of particle  $i$ .

The initial conditions for position ( $\mathbf{r}_i$ ) and velocity ( $\mathbf{v}_i$ ) of all particles are set, typically with the linear momentum of the center of mass (COM) and the system angular momentum initialized to zero [4]. In this thesis, the leap-frog algorithm [5] is employed to solve the equations of motion, described by the following equations:

$$\mathbf{v}_i \left( t + \frac{\Delta t}{2} \right) = \mathbf{v}_i \left( t - \frac{\Delta t}{2} \right) + \frac{\Delta t}{m} \mathbf{F}_i(t) \quad (2.3)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i \left( t + \frac{\Delta t}{2} \right) \quad (2.4)$$

Here,  $t$  denotes the simulation time and  $\Delta t$  is the integration time step. By integrating these equations over all timesteps, we can predict how particles move within the simulation. Figure 2.1 visually represents the sequential process of updating particle positions and velocities. To accurately simulate bulk properties and avoid edge effects, periodic boundary conditions (PBC) are employed [1]. PBCs ensure that a particle leaving one side of the simulation box re-enters from the opposite side, maintaining a continuous and infinite system environment.

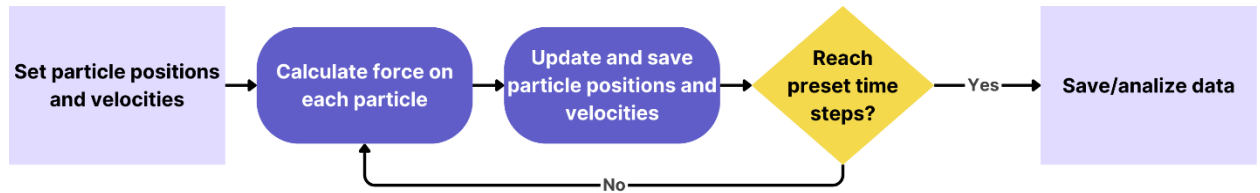


Figure 2.1. Flowchart illustrating the step-by-step process of updating particle positions and velocities in a MD simulation.

### 2.1.2. NVT ensemble

The canonical ensemble, or NVT ensemble, maintains a constant number of particles, volume, and temperature. In this ensemble, the system can exchange energy with a thermal reservoir to maintain a constant temperature. The probability distribution of states in the canonical ensemble is given by the Boltzmann factor [3]:

$$P(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta H(\mathbf{r}, \mathbf{p})}}{Z} \quad (2.5)$$

where  $\beta = \frac{1}{k_B T}$  (with  $k_B$  being the Boltzmann constant and  $T$  being the absolute temperature),  $\mathbf{r}$  and  $\mathbf{p}$  contains the positions and momenta of all the particles respectively,  $H(\mathbf{r}, \mathbf{p})$  is the Hamiltonian of the system, and  $Z$  is the partition function, defined as [3]:

$$Z = \sum_{states} e^{-\beta H(\mathbf{r}, \mathbf{p})} \quad (2.6)$$

The canonical ensemble is particularly useful for simulating systems at a fixed temperature, such as biological macromolecules in a thermal bath [2].

Since direct integration of Eq. (2.2), (2.3) and (2.4) corresponds to simulation in the NVE ensemble, special algorithms are required to maintain the constant temperature in the NVT ensemble. These algorithms are called thermostats [4]. The Nosé-Hoover thermostat introduces additional generalized coordinate and momentum, and employs Lagrangian mechanics to establish extra equations of motion so that the probability distribution Eq. (2.5) is achieved [1]. This method is rigorous yet time-consuming due to the added coupled equations of motion [1]. The velocity-rescale (v-rescale) thermostat adjusts the velocities of all particles proportionally to reach the desired kinetic energy which is related to temperature [6]. This method is fast but less accurate as it does not ensure the satisfaction of the probability distribution Eq. (2.5). Thus, it is mainly used during the setup phase to quickly achieve the target temperature [6]. In this thesis, the v-rescale thermostat was utilized for the equilibration phase, while the Nosé-Hoover thermostat was employed during the production run to maintain accurate temperature control.

### 2.1.3. NPT ensemble

The isothermal-isobaric ensemble, or NPT ensemble, is characterized by a constant number of particles, pressure, and temperature. This ensemble allows the system to exchange both energy and volume with the surroundings, making it ideal for studying phase transitions and material properties under constant pressure and temperature conditions [3]. The probability distribution for the NPT ensemble is given by [3]:

$$P(\mathbf{r}, \mathbf{p}, V) = \frac{e^{-\beta(H(\mathbf{r}, \mathbf{p}) + pV)}}{\Delta} \quad (2.7)$$

where  $\Delta$  is the isothermal-isobaric partition function:

$$\Delta = \int dV \sum_{states} e^{-\beta(H(\mathbf{r}, \mathbf{p}) + pV)} \quad (2.8)$$

Barostats are algorithms used to control the pressure within MD simulations, ensuring the system accurately reflects the desired pressure conditions [1]. The stochastic cell rescaling (C-rescale) algorithm, introduced by Bernetti and Bussi, is a first-order barostat designed to sample correct volume fluctuations by incorporating a stochastic term [7]. C-rescale operates by adjusting the simulation cell dimensions at each timestep based on the pressure difference, with an added random noise component [7]. This ensures that the volume fluctuations align with those expected from the NPT ensemble, making it suitable for both equilibration and production phases of MD simulations [7]. The frequently used Berendsen [3] barostat, while efficient at equilibrating systems, often result in inaccurate volume fluctuations. We used C-rescale for both equilibration and production run in our simulations.

## 2.2. Forcefields

Forcefields are essential for accurately modeling the interactions between particles. The OPLS-AA (Optimized Potentials for Liquid Simulations All Atom) forcefield was chosen for this thesis due to its proven versatility and accuracy in modeling a wide range of molecular systems [8]. Specifically, OPLS-AA is highly effective for simulating organic molecules and liquids,

making it an ideal choice for studying the interactions between g-C<sub>3</sub>N<sub>4</sub> and various organic solvents [8].

### 2.2.1. Bonded interactions

Bonded interactions in the OPLS-AA forcefield encompass forces arising from bond stretching, angle bending, and dihedral torsions [8]. These interactions are described by specific potential energy functions that depend on the relative positions of atoms. Bond stretching interactions are typically modeled using a harmonic potential, which describes how the energy increases quadratically as the bond length deviates from its equilibrium value [8]:

$$E_{bond} = \frac{1}{2}k_b(r - r_0)^2 \quad (2.9)$$

where  $k_b$  is the bond force constant,  $r$  is the current bond length, and  $r_0$  is the equilibrium bond length. Angle bending is also represented by a harmonic potential [8]:

$$E_{angle} = \frac{1}{2}k_\theta(\theta - \theta_0)^2 \quad (2.10)$$

In this equation,  $k_\theta$  is the angle force constant,  $\theta$  is the bond angle, and  $\theta_0$  is the equilibrium bond angle. Dihedral or torsional interactions are modeled using a periodic potential, capturing the energy changes due to rotations around bonds [8]:

$$E_{dihedral} = \sum_n \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad (2.11)$$

where  $V_n$  is the barrier height or dihedral force constant,  $n$  is the periodicity of the dihedral, indicating how many times the potential repeats as the dihedral angle completes a full 360° rotation,  $\phi$  is the dihedral angle, and  $\gamma$  is the phase angle, representing the shift in the cosine function.

### 2.2.2. Non-bonded interactions

Non-bonded interactions in the OPLS-AA forcefield describe the forces between atoms that are not directly bonded, including van der Waals and electrostatic interactions [3]. Van der

Waals (vdW) interactions are modeled using a Lennard-Jones potential, which combines both attractive and repulsive components [1]:

$$E_{LJ} = 4\epsilon' \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (2.12)$$

In this equation,  $\epsilon'$  is the depth of the potential well,  $\sigma$  is the distance at which the inter-particle potential is zero, and  $r$  is the distance between particles. Electrostatic interactions are described by Coulomb's law, representing the potential energy between charged particles [1]:

$$E_{coulomb} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \quad (2.13)$$

here,  $q_i$  and  $q_j$  are the charges of particles  $i$  and  $j$ ,  $r_{ij}$  is the distance between them, and  $\epsilon_0$  is the permittivity of free space.

Efficient computation of non-bonded interactions, particularly electrostatics, is crucial in MD simulations due to their slow convergence [3]. A common cutoff distance for non-bonded interactions, including vdW and short-range electrostatic interactions, is typically around 1.2 nm [2]. Unlike vdW interactions, the effects of electrostatic interactions do not become negligible beyond the cutoff distance [2]. To address long-range electrostatic interactions, the Particle Mesh Ewald (PME) method is employed [9], which splits the electrostatic calculation into short-range and long-range components, allowing for faster convergence than traditional methods [9]. The PME method involves grouping and averaging charges beyond the cutoff distance by assigning them to cells in a mesh. Charges are interpolated over the cells to form a charge density, which is then transformed into the reciprocal space using a Fast Fourier Transform (FFT) algorithm [10]. Electrostatic calculations are performed in the reciprocal space and subsequently converted back into the real space [9]. This approach significantly reduces the computational cost while maintaining high accuracy. By employing PME, we balance the need for accurate electrostatic calculations with computational efficiency, ensuring that long-range interactions are properly accounted for [9].

## 2.3. Umbrella sampling

Umbrella sampling (US) is a powerful technique used in MD simulations to enhance the sampling of rare events and to calculate the PMF [11]. The PMF provides insights into the free energy profile along a reaction coordinate, which is crucial for understanding various molecular processes [11]. In US, a series of simulations are performed with additional potential terms (biasing potentials) applied to the system to restrain it to specific regions of the phase space [11]. These potentials help in overcoming energy barriers and allow the system to sample configurations that would be otherwise rarely visited [11]. The biasing potential  $U_{bias}(x)$  is typically a harmonic potential [11]:

$$U_{bias}(x) = \frac{1}{2}k(x - x_0)^2 \quad (2.14)$$

where  $k$  is the force constant,  $x$  is the reaction coordinate, and  $x_0$  is the center of the bias. Each biased simulation is referred to as a "window," and multiple windows are set up along the reaction coordinate to cover the entire range of interest.

The PMF  $W(x)$  along a reaction coordinate  $x$  can be extracted from the biased simulations using the US data. The PMF is related to the unbiased probability distribution  $P(x)$  by [11]:

$$W(x) = -k_B T \ln P(x) + C \quad (2.15)$$

where  $T$  is the temperature, and  $C$  is an arbitrary constant. The unbiased probability distribution  $P(x)$  is obtained by combining the histograms from all the windows and correcting them for the biasing potentials. This is typically done using the Weighted Histogram Analysis Method (WHAM) [12]. WHAM is a technique used to combine data from multiple biased simulations to reconstruct the unbiased free energy profile. WHAM solves for the unbiased probabilities by iteratively adjusting the weights of the histograms obtained from each window until they converge to a consistent set of probabilities [12]. The WHAM equations are given by [12]:

$$P(x) = \frac{\sum_{i=1}^M n_i P_i(x) e^{\beta f_i}}{\sum_{i=1}^M n_i e^{\beta(f_i - U_i(x))}} \quad (2.16)$$



$$e^{-\beta f_i} = \sum_x P(x) e^{-\beta U_i(x)} \quad (2.17)$$

where  $P_i(x)$  is the probability distribution from window  $i$ ,  $U_i(x)$  is the biasing potential in window  $i$ ,  $f_i$  is the free energy offset for window  $i$ ,  $n_i$  is the number of samples in window  $i$ , and  $M$  is the total number of windows. WHAM iteratively solves these equations to obtain  $f_i$  and the unbiased probability distribution  $P(x)$ , which is then used to understand the free energy landscape of the system.

## 2.4. Radial distribution function

The radial distribution function (RDF),  $g(r)$ , is a key tool for characterizing the structure of a system at the atomic level [1]. It describes how the density of atoms varies as a function of distance from a reference atom, providing insights into the local structure and spatial correlations within the system. The RDF is defined as [1]:

$$g(r) = \frac{\langle \rho(r) \rangle}{\rho_0} \quad (2.18)$$

where  $\langle \rho(r) \rangle$  is the average local density of atoms at a distance  $r$  from the reference atom, and  $\rho_0$  is the bulk density of the same type of atoms. To calculate the RDF, the distances between pairs of atoms are binned into a histogram, which is then normalized by the volume of the spherical shell and the total number of reference atoms [1]:

$$g(r) = \frac{1}{N} \left( \frac{V}{4\pi r^2 \Delta r} \right) \sum_{i=1}^N \sum_{j \neq i}^N \delta(r - r_{ij}) \quad (2.19)$$

here,  $\Delta r$  is the bin width used to create the histogram,  $r_{ij}$  is the distance between atoms  $i$  and  $j$ , and  $\delta$  is the Dirac delta function, which ensures that only pairs of atoms separated by a distance  $r$  contribute to the RDF.

The RDF provides information on the likelihood of finding a particle at a distance  $r$  from a reference particle [3]. Peaks in  $g(r)$  indicate preferred distances corresponding to coordination

shells, while a value of  $g(r) \approx 1$  at large  $r$  signifies a random distribution of particles, as expected in a homogeneous system [3].

## 2.5. Machine learning

Machine learning (ML) has emerged as a transformative tool in computational materials science, offering the ability to uncover complex patterns and relationships within large datasets. Traditional methods in materials science, particularly those involving MD simulations, can be computationally intensive, often requiring significant time and resources [13]. ML provides an alternative by enabling rapid predictions of material properties, system behaviors, and potential outcomes based on existing data, without the need for full-scale simulations each time [13]. In the context of computational materials science, ML leverages algorithms that can build predictive models from historical simulation or experimental data by analyzing relationships between input features (such as structural, electronic, or chemical properties) and desired outputs (such as material performance or free energy calculations) [14], allowing for identification of trends that may not be easily evident through classical theoretical approaches.

ML models in materials science typically employ supervised learning techniques, where the model is trained on a set of known data to make predictions for unseen cases. These models can handle a wide variety of tasks, ranging from regression problems (e.g., predicting material properties like dielectric constants or exfoliation energies) to classification problems (e.g., categorizing materials based on their potential applications) [14]. Furthermore, ML role in computational materials science is not limited to prediction. It also facilitates the optimization of simulations by guiding the selection of parameters for MD simulations, or even suggesting which materials or chemical compositions are most likely to exhibit desired properties.

### 2.5.1. Supervised learning

Supervised learning is one of the most commonly used approaches in ML, particularly suited for tasks where the relationships between input features and known outputs need to be mapped [14]. In this framework, a model is trained on labeled data, where each input is paired with its corresponding output, allowing the model to learn the underlying patterns or dependencies within the dataset [13]. Once trained, the model can predict outputs for new, unseen inputs based

on these learned relationships. The following provides an overview of the specific models employed in this work.

Extra-trees regressor (ETR) and random forest regression (RFR) are ensemble ML techniques. RFR [15,16] consists of an ensemble (forest) of decision trees. Each tree has a series of decision nodes from which the tree splits into branches until it reaches its end (“the leaf”, i.e., the prediction). Each decision node depends on whether a certain input characteristic is greater than a specific value. Each tree in RFR is trained using a subset of the entire training set, and by averaging the predictions of individual trees a prediction is formed. ETR is an extension of RFR and it trains each tree using the entire set of training data. In order to split the nodes, ETR chooses a random cut point rather than a locally optimal cut point as used in RFR, which helps reduce overfitting. In gradient boosting regression (GBR), a new weak learner (a model that performs slightly better than random chance) is added to the previous model at each step such that the new model gives rise to smaller collective error. The weak learners are usually decision trees because of their strong empirical performance. Like GBR, AdaBoost regressor (ABR) is one of the boosting algorithms that automatically adjust its parameters based on the performance in the current iteration [17]. In ABR, each new model will weigh the data such that data with less accurate prediction are assigned more weight. By formulating linear regression using probability distributors, Bayesian ridge regression (BRR) is a method suitable for insufficient or poorly distributed data [18]. Rather than estimated as a single value, the output is drawn from a probability distribution.

### **2.5.2. Feature selection**

Feature selection is a critical step in ML, particularly for supervised learning models [19]. It involves identifying the most relevant input variables (features) that significantly contribute to the model predictions while discarding those that are redundant or less relevant. By focusing on the most informative features, feature selection enhances model accuracy, reduces overfitting, and decreases computational complexity [19]. In materials science, features can include various properties such as molecular structure, surface tension, or dielectric constant, which influence material behavior. For this work, we employed the all subset model (ASM) approach for feature selection [20]. ASM evaluates all possible combinations of features to determine which set provides the best predictive performance. This exhaustive method ensures that the model captures

the most important features while minimizing the inclusion of irrelevant ones, leading to more reliable predictions.

### **2.5.3. Model training and validation**

Model training and validation are essential for developing ML models that generalize well to new data. To ensure robust performance, k-fold cross-validation is often used. This technique divides the dataset into k subsets, or “folds,” training the model on k-1 folds while using the remaining fold for validation. The process is repeated k times, with each fold being used as the validation set once. This method provides a comprehensive assessment of model performance by averaging the results across all folds, reducing the bias that may result from a single training-test split.

In addition to cross-validation, hyperparameter tuning is key to optimizing the model. Instead of relying on default settings, hyperparameters—such as the number of trees in a random forest or the learning rate in gradient boosting—are fine-tuned to enhance model performance. Grid search is a common method used for this, where different combinations of hyperparameters are systematically tested, and the combination that yields the best cross-validation results is selected [21]. Together, these methods ensure that the model is both well-validated and optimized for predictive accuracy.

## **2.6. References**

- [1] C.L. Brooks, Computer simulation of liquids, *J Solution Chem* 18 (1989) 99–99.  
<https://doi.org/10.1007/BF00646086>.
- [2] H. Kamberaj, *Molecular Dynamics Simulations in Statistical Physics: Theory and Applications*, Springer International Publishing, Cham, 2020. <https://doi.org/10.1007/978-3-030-35702-3>.
- [3] M.E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, 2023.
- [4] D. Frenkel, B. Smit, *Understanding Molecular Simulation: Monte Carlo Simulations in Various Ensembles*, Academic Press, London (2002) 111–137.  
<https://www.sciencedirect.com/science/article/pii/B9780122673511500079> (accessed August 5, 2024).

- [5] W.F. Van Gunsteren, H.J.C. Berendsen, A Leap-frog Algorithm for Stochastic Dynamics, *Mol Simul* 1 (1988) 173–185. <https://doi.org/10.1080/08927028808080941>.
- [6] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *Journal of Chemical Physics* 126 (2007). <https://doi.org/10.1063/1.2408420>/186581.
- [7] M. Bernetti, G. Bussi, Pressure control using stochastic cell rescaling, *Journal of Chemical Physics* 153 (2020). <https://doi.org/10.1063/5.0020514>/199610.
- [8] W. Damm, A. Frontera, J. Tirado-Rives, W.L. Jorgensen, OPLS all-atom force field for carbohydrates, *J Comput Chem* 18 (1997) 1955–1970. [https://doi.org/10.1002/\(SICI\)1096-987X\(199712\)18:16<1955::AID-JCC1>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1096-987X(199712)18:16<1955::AID-JCC1>3.0.CO;2-L).
- [9] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J Chem Phys* 103 (1995) 8577–8593. <https://doi.org/10.1063/1.470117>.
- [10] R.P. Futrelle, D.J. McGinty, Calculation of spectra and correlation functions from molecular dynamics data using the fast Fourier transform, *Chem Phys Lett* 12 (1971) 285–287. [https://doi.org/10.1016/0009-2614\(71\)85065-0](https://doi.org/10.1016/0009-2614(71)85065-0).
- [11] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [12] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [13] N.E. Jackson, B.M. Savoie, A. Statt, M.A. Webb, Introduction to Machine Learning for Molecular Simulation, *J Chem Theory Comput* 19 (2023) 4335–4337. <https://doi.org/10.1021/ACS.JCTC.3C00735>.
- [14] E. Prašnikar, M. Ljubič, A. Perdih, J. Borišek, Machine learning heralding a new development phase in molecular dynamics simulations, *Artificial Intelligence Review* 2024 57:4 57 (2024) 1–36. <https://doi.org/10.1007/S10462-024-10731-4>.

- [15] L. Breiman, Random Forests, Machine Learning 2001 45:1 45 (2001) 5–32.  
<https://doi.org/10.1023/A:1010933404324>.
- [16] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Machine Learning 2006 63:1 63 (2006) 3–42. <https://doi.org/10.1007/S10994-006-6226-1>.
- [17] H.D.- ICML, undefined 1997, Improving regressors using boosting techniques, Citeseer (n.d.).  
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.21.5683&rep=rep1&type=pdf>  
(accessed August 8, 2022).
- [18] M.T.-J. of machine learning research, undefined 2001, Sparse Bayesian learning and the relevance vector machine, Jmlr.Org (n.d.).  
<https://www.jmlr.org/papers/volume1/tipping01a/tipping01a.pdf?ref=https://githubhelp.com>  
(accessed August 8, 2022).
- [19] W. Bin How, B. Wang, W. Chu, A. Tkatchenko, O. V. Prezhdo, Significance of the Chemical Environment of an Element in Nonadiabatic Molecular Dynamics: Feature Selection and Dimensionality Reduction with Machine Learning, Journal of Physical Chemistry Letters 12 (2021) 12026–12032.  
[https://doi.org/10.1021/ACS.JPCLETT.1C03469/SUPPL\\_FILE/JZ1C03469\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCLETT.1C03469/SUPPL_FILE/JZ1C03469_SI_001.PDF).
- [20] M.A. Hall, L.A. Smith, Practical feature subset selection for machine learning, Computer Science Volume 20 No 1 (1998) 181–191.  
<https://researchcommons.waikato.ac.nz/handle/10289/1512> (accessed August 13, 2022).
- [21] D.M. Belete, M.D. Huchaiah, Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results, International Journal of Computers and Applications 44 (2022) 875–886. <https://doi.org/10.1080/1206212X.2021.1974663>.

# Chapter 3: Liquid-phase Exfoliation of Graphitic Carbon Nitrides

## 3.1. Introduction

Two-dimensional g-C<sub>3</sub>N<sub>4</sub> nanosheets have been explored as a promising candidate in many applications such as photocatalysis [1–4], membranes [5], sensing [6], imaging [7], and energy conversion [8,9]. Importantly, the basic structural unit of g-C<sub>3</sub>N<sub>4</sub> is a  $\pi$ -conjugated heptazine or triazine, which exhibits special electronic and photocatalytic performances [10]. In addition, a weaker photoluminescence intensity was observed in ultrathin g-C<sub>3</sub>N<sub>4</sub> nanosheets compared to the layered bulk material [3], which corresponds to a lower inter-sheet recombination rate of photogenerated carriers. Therefore, the superiority of ultrathin g-C<sub>3</sub>N<sub>4</sub> sheets results from a large specific surface area, enhanced electron transport ability, and high charge separation efficiency [11]. LPE, as an important functional modification strategy, can be used to delaminate layered g-C<sub>3</sub>N<sub>4</sub> bulk material to obtain free-standing nanosheets [12–18]. In this process, bulk materials are placed into certain solvents followed by ultrasonic treatments [19]. Despite some successes [20–27], it remains a challenge to identify solvents that enable effective exfoliation and stable dispersion of the g-C<sub>3</sub>N<sub>4</sub> sheets.

A number of studies have investigated the delamination of bulk g-C<sub>3</sub>N<sub>4</sub> in different organic solvents. The selection of solvents was mostly arbitrary, based on trial-and-error experimentation. For example, Yang et al. reported LPE of bulk g-C<sub>3</sub>N<sub>4</sub> using various organic solvents as the dispersion medium, including but not limited to IPA, NMP, acetone (ACE), and ethanol [28]. It was found that NMP was a promising solvent, which could stabilize the individual nanosheets. In comparison, precipitation of g-C<sub>3</sub>N<sub>4</sub> was observed after 2 days when using ethanol and ACE. Lin et al. reported a mixed solvent approach for the LPE of bulk g-C<sub>3</sub>N<sub>4</sub>, to form monolayer g-C<sub>3</sub>N<sub>4</sub> nanosheets with tunable concentrations (0.1–3 mg mL<sup>-1</sup>) [29]. Different binary solvents, namely ethanol/H<sub>2</sub>O, IPA/H<sub>2</sub>O, and DMF/H<sub>2</sub>O were examined. For the ethanol/H<sub>2</sub>O system, when the volume ratio of H<sub>2</sub>O was 75%, a milky dispersion was obtained, reaching the maximum g-C<sub>3</sub>N<sub>4</sub> concentration of 3 mg mL<sup>-1</sup>.

To the best of our knowledge, attempts to determine a quantitative relationship between the solvent properties and the performance of the exfoliation are completely absent for g-C<sub>3</sub>N<sub>4</sub>, while there has been some work on another 2D material, graphene. Hernandez and co-workers [21] reported that the Hildebrand solubility parameters, Hansen solubility parameters, and surface tensions may be used to identify solvents that are effective in dispersing graphene. It was discussed that in order to minimize the enthalpic cost of mixing, efficient solvents should have a surface tension close to that of graphene. In another study, Coleman et al. [30] stated that the concentration of dispersed graphene nanosheets is maximized when the energy of exfoliation is minimized. An equation was proposed to calculate the enthalpy of mixing  $\Delta H_{mix}$  as an estimate for the energy of exfoliation,

$$\frac{\Delta H_{mix}}{V_{mix}} = \frac{2}{T_{sheet}} \left( \sqrt{E_{sur}^{graphite}} - \sqrt{E_{sur}^{solvent}} \right)^2 \phi \quad (3.1)$$

where  $V_{mix}$  is the volume of the solution mixture,  $T_{sheet}$  is the thickness of a nanosheet,  $E_{sur}^i$  is the surface energy of component  $i$ , and  $\phi$  is the volume fraction of graphene in the solution.

The above semiempirical criteria have provided some general guidelines for solvent selection, however, they do not allow for molecular-level design of novel solvents capable of effectively dispersing 2D nanosheets, especially g-C<sub>3</sub>N<sub>4</sub>. For example, an effective solvent should have interactions with g-C<sub>3</sub>N<sub>4</sub> that are strong enough to compensate for the vdW attraction between the g-C<sub>3</sub>N<sub>4</sub> sheets. It is unclear whether these interactions can be captured by solubility parameters or surface tension alone. In this regard, little is known about the molecular details of the interactions between g-C<sub>3</sub>N<sub>4</sub> and solvent molecules, and the correlation of these interactions with the structural properties of the solvents. Hence, developing a molecular-level understanding of the interactions between g-C<sub>3</sub>N<sub>4</sub> and common solvents can provide fundamental insights that help the design of exfoliation strategies to obtain stable g-C<sub>3</sub>N<sub>4</sub> solutions.

MD is a powerful numerical tool that can shed light on interactions beyond the accessibility of current experimental techniques. Zou et al. [31] presented the first MD study on the exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets, where an external force was added to imitate the sonication process. The energy barrier for exfoliation was estimated by the variation of kinetic energies in MD simulations, which was a very rough approximation. There was a lack of accurate evaluation of the free energy



required in the exfoliation process, as well as elucidation on the relationship between the free energy and solvent properties. In this work, we present a systematic MD study on the LPE of g-C<sub>3</sub>N<sub>4</sub> in nine different solvents with varying molecular structures. Through PMF calculations, we have for the first time quantified the free energy of exfoliation between two g-C<sub>3</sub>N<sub>4</sub> sheets, which can be used as a metric to evaluate the performance of g-C<sub>3</sub>N<sub>4</sub> exfoliation in those solvents. Mechanistic explanations are provided on the relationship between the free energy of exfoliation and molecular features of the solvent. Results from this study shed light on how to make molecular structure-based selection of solvents to improve g-C<sub>3</sub>N<sub>4</sub> exfoliation and dispersion in the liquid-phase.

## 3.2. Computational details

### 3.2.1. Molecular models for g-C<sub>3</sub>N<sub>4</sub> and solvents

The heptazine-based g-C<sub>3</sub>N<sub>4</sub> nanosheet was considered for this study (Figure 3.1a). This structure consists of 6-membered rings with sp<sup>2</sup> bonds between C and N atoms, and intrinsically contains vacancies (or voids) surrounded by nitrogen atoms as shown by the red circle in Figure 3.1a. The presence of these vacancies and the rotation at the N-C bonds that connects two building blocks (cyan circle in Figure 3.1a highlights one building block) make the growth of a large-sized layer difficult [20]. In this work, we considered an equilateral triangular sheet to be representative of a g-C<sub>3</sub>N<sub>4</sub> nanosheet, consistent with several previous studies [32–36], with a side length of 4.5 nm. This nanosheet consists of 21 heptazine core units, with the periphery saturated with 21 H atoms. The justification for choosing this size is provided in Appendix A, section A1.

For the study of LPE, nine solvents commonly used in the experiments [27,29,37] were considered as the medium. These solvents were categorized into three groups based on their structural character and functional groups. The first group (Figure 3.1b) featured polar O-H (hydroxyl) groups and included water and two alcohols: methanol (MET) and 1,4-butanediol (BD). The second group (Figure 3.1c) had a carbonyl (O=C) group which can be an amide as in formamide (FRM) and DMF, or a ketone as in ACE. The third group (Figure 3.1d) contained an aromatic structure: a benzene ring in chlorobenzene (CB), or a heterocycle in tetrahydrofuran (THF) and NMP.

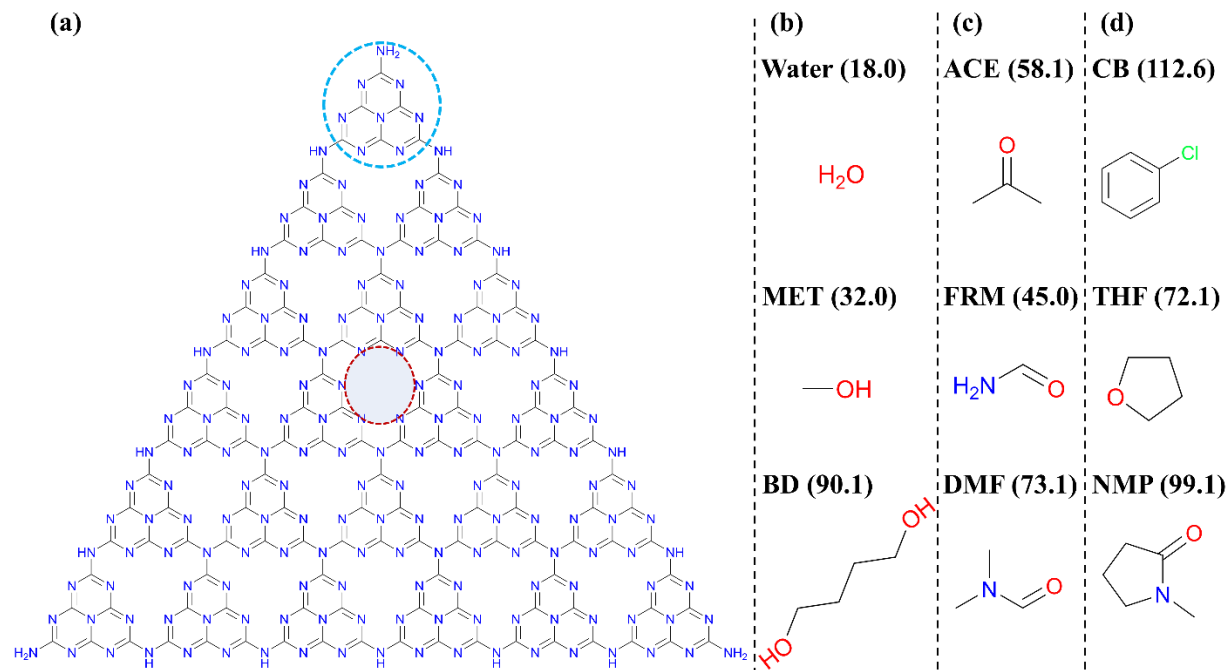


Figure 3.1. Molecular structure of the 21-unit g-C<sub>3</sub>N<sub>4</sub> used in this study. The cyan circle shows one heptazine core unit and the red circle shows a vacancy that exists around the N atoms. Molecular structures of simulated solvents, categorized into group 1 (b), group 2 (c) and group 3 (d) according to their structural characteristics. The molar weight of each solvent is given in the parenthesis.

The OPLS-AA force field [38] was used for both the solute (g-C<sub>3</sub>N<sub>4</sub>) and the solvents. The force field parameters were generated from LigParGen and PolyParGen servers [39,40] unless otherwise specified. The Charge Model 5 (CM5) with a scaling factor of 1.20 [41] was used for the partial atomic charges. To validate the force field parameters and partial atomic charges, key properties of g-C<sub>3</sub>N<sub>4</sub> such as bond length/angle distributions, surface energy and hydration free energy were compared with results from density functional theory (DFT) calculations. Validation was also performed for the solvents, by comparing their density, dielectric constant, and surface tension with experimental results. Details of the validation are shown in Appendix A, Section A2.

### 3.2.2. Simulated systems

US [42] was used to calculate the PMF, which is defined as the potential whose negative gradient with respect to a pre-defined reaction coordinate (RC) corresponds to the average force

in the direction of the RC. The weighted histogram analysis method (WHAM) [43] was used to extract the PMF curve from a set of US simulations. It is non-trivial to define a RC that best describes the exfoliation of the g-C<sub>3</sub>N<sub>4</sub> nanosheets. Two adhered 2D sheets can be separated via different paths, which may be associated with different free energy requirements [44–48]. For example, it was shown that for graphene sheets in water the PMF to separate them in the direction perpendicular to the sheets (i.e., normal direction) was lower compared to that in the direction parallel to the sheets (i.e., lateral direction) [45]. However, for boron nitride (BN) nanosheets in several organic solvents, it was reported that the lateral direction had lower PMF [46].

In this work, we first performed a set of simulations to determine the most probable path for the exfoliation of g-C<sub>3</sub>N<sub>4</sub>. For this purpose, three different RCs were explored with DMF being the solvent. RC I was defined as the distance between the COMs of two sheets along the normal direction (Figure 3.2a); PMF associated with this RC corresponded to the process of separating the two sheets in the normal direction. RC II was the distance between the edge atoms of two sheets in the normal direction (Figure 3.2b), resembling the peeling of the top sheet from the bottom one. Finally, RC III was the distance between the COMs of two sheets in the lateral direction (Figure 3.2c), to capture the process of sliding the top layer over the bottom one. RC I ranged from 0.3 to 1.3 nm, beyond which the two sheets no longer interacted with each other, while RC II and III ranged from 0.3 to 4.0 nm. For each chosen RC, US simulations were performed by applying a harmonic biasing potential (force constant 5000 kJ/mol.nm<sup>2</sup>) to restrain the RC at each US window. To complement the PMF calculations, a pulling simulation was carried out along each of the three RCs in DMF. In these simulations, the bottom sheet was fixed while the center of the mass (COM) of the top sheet (RC I and RC III) or edge atom (RC II) of the top sheet was attached to a spring of stiffness 1000 kJ/mol.nm<sup>2</sup> and pulled away at a speed of 0.005 nm/ps. PMFs and forces obtained from the three RCs were compared to determine the one with the lowest resistance to separation, i.e., the most probable path for exfoliation. Then this RC was employed to calculate the PMFs for all the other solvents.

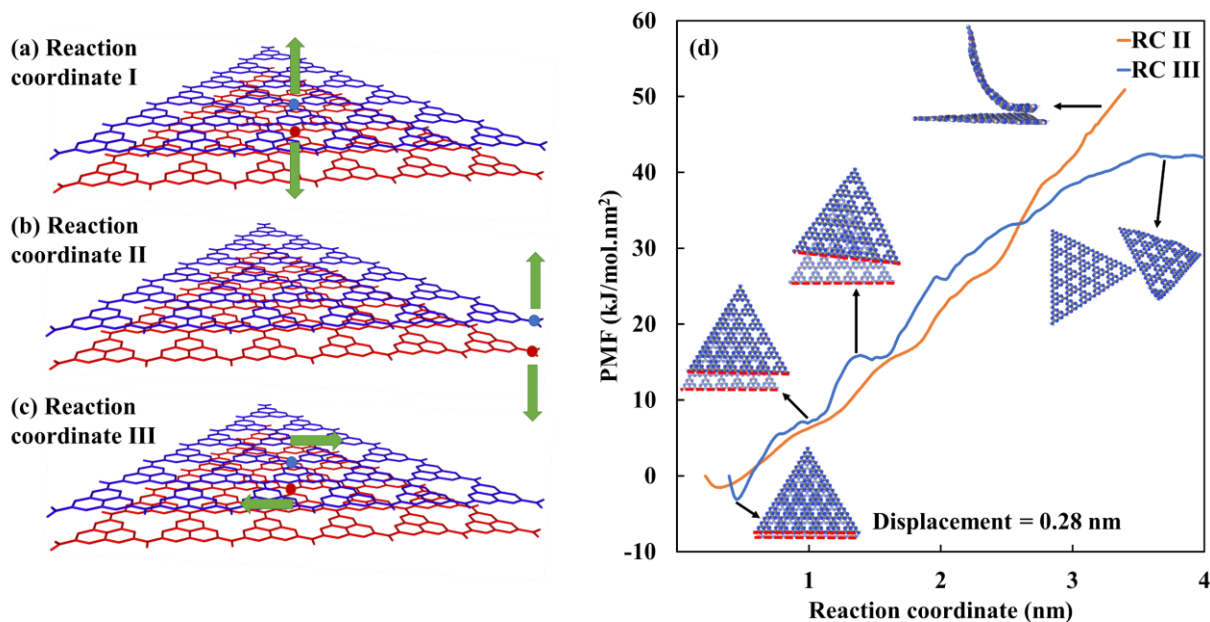


Figure 3.2. Three RCs explored for the separation of two g-C<sub>3</sub>N<sub>4</sub> nanosheets: (a) the distance between the COMs of two sheets along the normal direction, (b) the distance between the edge atoms of two sheets in the normal direction, and (c) the distance between the COMs of two nanosheets in the lateral direction. For better illustration, the two sheets are colored differently. The COMs or edge atoms are shown with blue and red circles respectively for the two sheets. (d) PMF curves for separating two stacked g-C<sub>3</sub>N<sub>4</sub> nanosheets in DMF along RC II and RC III. Insets show representative snapshots at different stages of the separation.

In addition to the PMF calculations, a set of simulations were performed to determine the solvation free energy (SFE, denoted by  $\Delta G_{\text{sol}}$ ) of a single g-C<sub>3</sub>N<sub>4</sub> sheet in different solvents. Bennett Acceptance Ratio (BAR) method [49] was used to gradually couple a sheet to its equilibrium solvation environment and the average  $\langle dH(\lambda)/d\lambda \rangle$  was evaluated, where  $H$  is the Hamiltonian and  $\lambda$  is the coupling parameter. Twenty-one states were defined in this procedure where the first state corresponded to no interaction ( $\lambda = 0$ ) between the g-C<sub>3</sub>N<sub>4</sub> sheet and the solvent. In the next 10 states the vdW interaction between the g-C<sub>3</sub>N<sub>4</sub> sheet and the solvent was increased with the step of  $\Delta\lambda = 0.1$ . For the last 10 states the electrostatic interaction was turned on by the step of  $\Delta\lambda = 0.1$  reaching the fully interacting state between the sheet and solvent. Consequently, the  $\Delta G_{\text{sol}}$  between  $\lambda = 0$  and  $\lambda = 1$  can be calculated from:

$$\Delta G_{sol} = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{dH}{d\lambda} \right\rangle_{\lambda} d\lambda \quad (3.2)$$

Finally, the fully interacting systems from the end of the  $\Delta G_{sol}$  simulations were further equilibrated for 60 ns. Afterward, the g-C<sub>3</sub>N<sub>4</sub> sheet was removed from each system, the solvent was re-equilibrated for 60 ns and the g-C<sub>3</sub>N<sub>4</sub> sheet was equilibrated in vacuum for 10 ns. This set of simulations allowed us to calculate the solvation enthalpy ( $\Delta H_{sol}$ ) from

$$\Delta H_{sol} = H_{solution} - H_{solvent} - H_{solute} \quad (3.3)$$

where  $H_{solution}$  is the enthalpy of the solution after 60 ns equilibration,  $H_{solvent}$  is the enthalpy of the solvent alone after 60 ns re-equilibration, and  $H_{solute}$  is the enthalpy of the sheet alone after 10 ns equilibration in vacuum. Each enthalpy term was calculated based on  $H = E + pV$  where  $E$  is the internal energy,  $p$  is pressure and  $V$  is volume. The solvation entropy was then estimated by

$$T\Delta S_{sol} = \Delta H_{sol} - \Delta G_{sol} \quad (3.4)$$

where  $T = 300$  K is the simulation temperature. A summary of the simulated systems is given in Table 3.1. These simulations generated trajectories with a total length of 5  $\mu$ s.

Table 3.1. Summary of the simulated systems.

Simulation type	# g-C <sub>3</sub> N <sub>4</sub> sheets	Solvent	Size of the box (nm*nm*nm)	# simulation windows	Time of production run in each simulation window (ns)
Pulling	2	DMF	9*10*9	1	1
PMF with RC I	2	DMF	8*8*10	40	10
PMF with RC II	2	DMF	8*8*10	30	10
PMF with RC III	2	Water, MET, BD, ACE, FRM, DMF, CB, THF, NMP	9*10*9	40	10
$\Delta G_{\text{sol}}$ calculation	1	Water, MET, BD, ACE, FRM, DMF, CB, THF, NMP	7*7*4	21	1
$\Delta H_{\text{sol}}$ and $\Delta S_{\text{sol}}$ calculation	0 and 1	Water, MET, BD, ACE, FRM, DMF, CB, THF, NMP	5.8*5.8*5.8	1	60

### 3.2.3. Simulation details

All simulations were performed using the GROMACS package [50] (version 2021.2). Prior to each production run, energy minimization of the system was carried out using a steepest descent algorithm. The system was then equilibrated in the NPT ensemble for 200 ps with a 2 fs timestep. The temperature and pressure were controlled at 300 K and 1 bar by the Berendsen thermostat and barostat, with a coupling time constant of 0.2 ps and 5 ps respectively. A production run was subsequently performed with a 2 fs timestep, where the bond lengths involving a hydrogen atom were constrained using the LINCS algorithm. A temperature of 300 K was maintained by means of a velocity-rescale algorithm with a time constant of 0.1 ps. An isotropic pressure of 1 bar was set by using a C-rescale scheme with a coupling constant of 1.0 ps. Periodic boundary condition was employed in all directions, and long-range electrostatic interaction was evaluated using the particle-mesh Ewald summation. The cutoff radius for the nonbonded interactions was set to 1.2 nm. The length of each production simulation is given in Table 3.1.

## 3.3. Results and discussion

### 3.3.1. Selection of reaction coordinate

To determine the most probable path for exfoliation, pulling simulations for two stacked g-C<sub>3</sub>N<sub>4</sub> sheets in DMF were first performed for all three RCs. The force-displacement curves are

shown in Appendix A, Section A3. Along RC I a large force ( $\sim 4300 \text{ kJ/mol}\cdot\text{nm}$ ) is required, which occurs at around 0.45 nm, in order to overcome the  $\pi$ - $\pi$  interaction [51] between the two stacked sheets. After this peak, the force decreases rapidly because of the diminishing interactions. For RC II and III, the force is steady and below  $\sim 1000 \text{ kJ/mol}\cdot\text{nm}$ . The results suggest that exfoliation via RC I is significantly more difficult compared with RC II and III as demonstrated by the more than 4-fold larger force required at the initial stage. Subsequently, we focused on RC II and III and generated the associated PMF curves in Figure 3.2d. Here the dimension of the PMF is energy per unit area (normalization done using the area of one nanosheet), and the PMF value of the first data point was set to zero in both curves. At the early stage ( $\text{RC} < 2.5 \text{ nm}$ ), the two PMFs are comparable. However, as the RC further increases, the two curves show the sign of separating from each other. When  $\text{RC} > 3.0 \text{ nm}$ , the curve for RC III has leveled off while the one for RC II curve is still climbing. The difference between the two curves can be understood by recognizing that while the separation of the two nanosheets, and hence the loss of interaction between them, is gradual at the beginning for both RCs, in the final stage detaching the top sheet from the bottom one along RC II (inset of Figure 3.2d) requires a much larger amount of energy. These results suggest that among the three RCs, RC III requires the least amount of energy to completely separate two stacked g-C<sub>3</sub>N<sub>4</sub> sheets, and therefore exfoliation in the experiments is more probable to occur along this RC. The same observation is also confirmed numerically in other solvents (CB and NMP, see Appendix A, section A3), and for other 2D materials such as BN [46]. For this reason, PMF calculations were performed along RC III for all 9 solvents and compared to assess their ability to disperse g-C<sub>3</sub>N<sub>4</sub> sheets.

### 3.3.2. PMF and free energy of exfoliation

The PMF curve in DMF associated with RC III (blue curve in Figure 3.2d) is examined with details here. As the COM distance ( $D$ ) between the two sheets increases, the PMF shows an initial decreasing trend, characteristic of a repulsive interaction between them. The curve reaches the global minimum at  $D = 0.45 \text{ nm}$  where the average force between the two sheets is zero. At the global minimum, the sheets are stacked with a lateral shift of 0.28 nm (inset of Figure 3.2d) compared to the initial configuration where they are in complete overlap. This configuration has been predicted for both g-C<sub>3</sub>N<sub>4</sub> [52] and graphene like C<sub>3</sub>N [53] bilayers by DFT calculations and corresponds to displaced  $\pi$ - $\pi$  stacking. After the global minimum, the PMF exhibits an overall increasing trend, which represents attraction between the two sheets. There are small fluctuations

in the curve, and the local minima and maxima are caused by the triangle patterns in the g-C<sub>3</sub>N<sub>4</sub> structure. Specifically, the heptazine core units from the two sheets prefer to be overlapping with each other, and the same applies to the void regions. Occurrence of such a configuration tends to result in a local minimum in the PMF. For example, the inset of Figure 3.2d at  $D = 1.15$  nm shows the structure of two nanosheets that have ten of their heptazine core units overlapped. On the contrary, at  $D = 1.25$  nm (inset of Figure 3.2d), the heptazine core units are not aligned, leading to a local maximum in the PMF curve. The PMF curve reaches a plateau around  $D = 3.50$  nm, indicating negligible interaction between the sheets beyond this COM separation. The difference in PMF value between the global minimum and the plateau represents the free energy required to separate two stacked sheets, via the mode of sliding, from their equilibrium distance until they are no longer interacting. This difference is therefore defined as the free energy of exfoliation ( $\Delta G_{\text{exf}}$ , positive throughout this work).

The PMF curves for all 9 simulated solvents are presented in Figure 3.3a, with  $\Delta G_{\text{exf}}$  marked for DMF as an example. Validity of the  $\Delta G_{\text{exf}}$  calculation is demonstrated by good overlaps between the US histograms as well as reproducibility confirmed by five independent set of US simulations performed for DMF (Appendix A, Section A3). Quantitatively,  $\Delta G_{\text{exf}}$  measures the level of difficulty to exfoliate one sheet from another; lower  $\Delta G_{\text{exf}}$  is preferred and exfoliation is expected to be easier with less external efforts (e.g. ultrasonication time and intensity) required. The ranking of  $\Delta G_{\text{exf}}$  follows: water > MET > FRM > ACE > BD > THF > CB > DMF > NMP, which suggests that exfoliation of g-C<sub>3</sub>N<sub>4</sub> is easiest in DMF and NMP, and hardest in water.



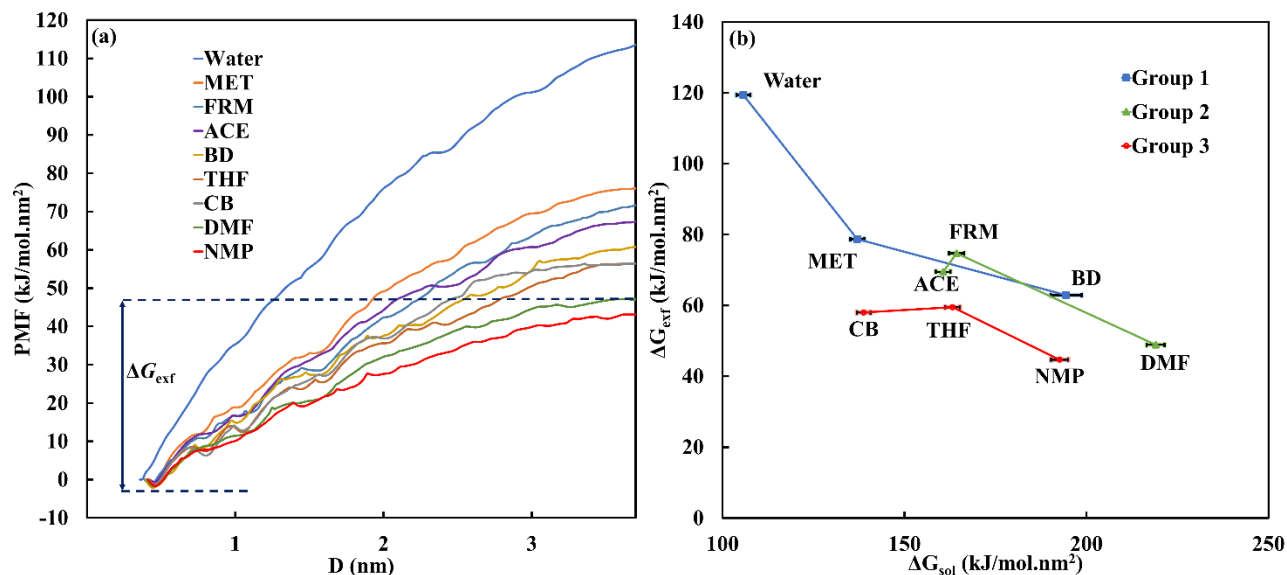


Figure 3.3. (a) PMF curves for separating two g-C<sub>3</sub>N<sub>4</sub> nanosheets in all solvents investigated in this study. (b)  $\Delta G_{\text{exf}}$  vs.  $\Delta G_{\text{sol}}$  for solvents from the three groups. Error bars in  $\Delta G_{\text{sol}}$  are obtained by splitting the data into 5 blocks, calculating the free energy difference over each block, and estimating the error from the average variance over the blocks [54].

### 3.3.3. Mechanisms governing the free energy of exfoliation

While the PMF calculations have allowed us to evaluate and rank the free energy of exfoliation, mechanisms behind such ranking require more investigation. Fundamentally, stronger attractive interaction between g-C<sub>3</sub>N<sub>4</sub> and the solvent should lead to better solubility and a lower value of  $\Delta G_{\text{exf}}$ , and this was demonstrated for graphene oxide nanosheets [55]. Quantitatively, the solubility can be evaluated by calculating the SFE ( $\Delta G_{\text{sol}}$ ) of a single sheet, which is the energy required to create a unit area of contact between the g-C<sub>3</sub>N<sub>4</sub> sheet and the solvent.  $\Delta G_{\text{sol}}$  is negative for all solvents, indicating attractive interaction between the sheet and the solvent (for the ease of discussion, in this thesis the solvation free energy is denoted by  $\Delta G_{\text{sol}}$  and refers to its absolute value).  $\Delta G_{\text{exf}}$  is plotted against  $\Delta G_{\text{sol}}$  in Figure 3.3b for all three groups of the solvents. The standard deviation in  $\Delta G_{\text{sol}}$  is small; hence hereafter we will refer only to the average values. In general, a negative correlation between  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  is observed within each group, consistent with our expectation that better solubility would lead to easier exfoliation. The three groups are

examined in more detail to identify how the structural features of the solvent molecules affect  $\Delta G_{\text{sol}}$ , and hence  $\Delta G_{\text{exf}}$ .

### 3.3.3.1. Group 1 (solvents with hydroxyl group):

In this group the ranking for  $\Delta G_{\text{sol}}$  follows  $\text{BD} > \text{MET} > \text{water}$  which suggests BD as a better liquid for the exfoliation of g-C<sub>3</sub>N<sub>4</sub>. Interestingly, the size and molecular weight ( $M_w$ ) of the solvents in this group follow the same trend,  $\text{BD} (90.1) > \text{MET} (32.0) > \text{water} (18.0)$ . A few analyses were done to further investigate the relationship between  $\Delta G_{\text{sol}}$  and structural features of the solvent molecules. Because all three solvents in this group contain hydroxyl groups capable of forming hydrogen bonds (H-bonds), an analysis was first performed to calculate the number of H-bonds dissociated and generated during the solvation. The results are provided in Appendix A, Section A4, which show no direct correlation with  $\Delta G_{\text{sol}}$ . In fact, larger size of MET and BD molecules has limited their capability to orient themselves in order to form close contact with the interior of the sheet. The solvent-sheet interaction is therefore suspected to be more impacted by non-specific electrostatic and vdW forces. In addition, entropy may play an important role in solvation [56]. In the following,  $\Delta G_{\text{sol}}$  is separated into its enthalpic and entropic components to gain more insights into the governing contribution.

$\Delta G_{\text{sol}}$  and its partition into the enthalpic ( $\Delta H_{\text{sol}}$ ) and entropic ( $T\Delta S_{\text{sol}}$ ) terms are given in Appendix A, Section A5. Both  $\Delta H_{\text{sol}}$  and  $T\Delta S_{\text{sol}}$  are negative, suggesting that solvation of the nanosheet in the three solvents is favored by enthalpy while opposed entropically. In addition, the magnitude of  $\Delta H_{\text{sol}}$  is more than double that of  $T\Delta S_{\text{sol}}$ , and the magnitudes of  $\Delta H_{\text{sol}}$  and  $\Delta G_{\text{sol}}$  follow the same order of  $\text{BD} > \text{MET} > \text{water}$ . The solvation of the nanosheet is therefore an enthalpy-driven process, larger reduction in enthalpy indicates higher sheet-solvent affinity, and leads to more favorable exfoliation [44,46,57,58].

Solvation of a g-C<sub>3</sub>N<sub>4</sub> nanosheet involves multiple steps with different energy implications. First, a cavity must be created in the solvent to accommodate the sheet. This step is both entropically ( $\Delta S < 0$ ) and enthalpically ( $\Delta H > 0$ ) unfavorable, as it increases the order of the solvent and reduces solvent-solvent interaction. Stronger interaction among solvent molecules leads to a greater enthalpic penalty for cavity formation. Next, the sheet enters the cavity, and the resulting solvent-sheet interaction is enthalpically favorable ( $\Delta H < 0$ ). Finally, as the solute is mixed into the solvent, there is an entropy gain ( $\Delta S > 0$ ). Since the solvation of g-C<sub>3</sub>N<sub>4</sub> nanosheet is enthalpy

driven, we further calculated the energy changes during these steps. The energy change associated with creating a cavity in the solvent, denoted by  $\Delta E_1$ , was estimated from the difference in non-bonded interactions (vdW and electrostatic) among the solvent molecules from two 60 ns simulations (last set in Table 3.1): one with the presence of a sheet and one without the sheet. Similarly, the energy change due to the insertion of a sheet into a pre-existing cavity, denoted by  $\Delta E_2$ , was estimated from the non-bonded interactions between the solvent and the sheet after 60 ns equilibration of the sheet in the solvent. The total energy change in the solvation process then can be estimated as

$$\Delta E = \Delta E_1 + \Delta E_2 \quad (3.5)$$

Figure 3.4a shows  $\Delta E_1$ ,  $\Delta E_2$ ,  $\Delta E$  along with  $\Delta H_{\text{sol}}$  from Appendix A, Section A5. Statistics are based on the last 30 ns of the two 60 ns simulations described above.  $\Delta E$  and  $\Delta H_{\text{sol}}$  only differ by a small amount suggesting that non-bonded interactions are the main contribution to the enthalpy of solvation. From Figure 3.4a, the penalty for disrupting solvent-solvent interaction ( $\Delta E_1$ , positive) follows order of water > BD > MET. Meanwhile, attractive solvent-sheet interaction ( $\Delta E_2$ , negative) follows water > MET > BD. Solvating the nanosheet in water involves both higher penalty for disrupting solvent-solvent interaction and weaker solvent-sheet attraction, consistent with its smallest  $\Delta G_{\text{sol}}$  and highest  $\Delta G_{\text{exf}}$ . While  $\Delta E_1$  is higher in BD than in MET, it is compensated by the stronger attraction between BD and the sheet, making BD a better medium for g-C<sub>3</sub>N<sub>4</sub> solvation.

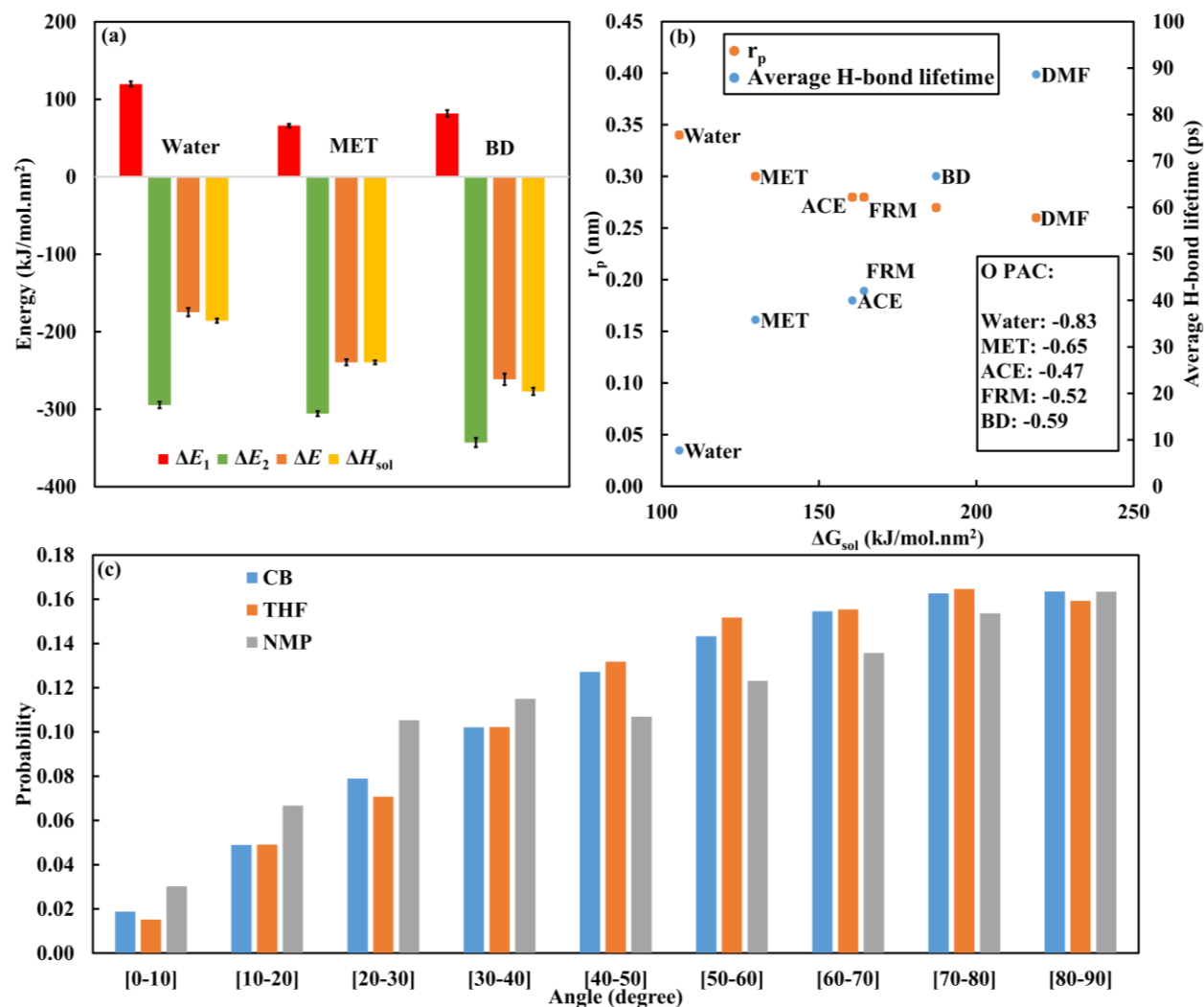


Figure 3.4. (a) For group 1 solvent: energy changes in the solvation process associated with cavity generation ( $\Delta E_1$ ) and sheet insertion ( $\Delta E_2$ ). The sum of the two terms ( $\Delta E = \Delta E_1 + \Delta E_2$ ) along with  $\Delta H_{sol}$  are also shown. (b) For group 1 and 2 solvents:  $r_p$  (location of the first peak in the RDF) vs.  $\Delta G_{sol}$  (left axis) and average H-bond lifetime vs.  $\Delta G_{sol}$  (right axis) for group 1 and 2 solvents. Inset shows the partial atomic charges (PAC) of the solvent oxygen. (c) For group 3 solvents: probability distribution of the angle between the normal of the sheet and the normal of the aromatic rings of the solvent in the first solvation layer.

Now that it is shown that the exfoliation process is enthalpy-driven and influenced by both solvent-solvent and solvent-sheet interactions, the question remains as to why  $\Delta H_{sol}$  follows the order of BD > MET > water, exactly the same as the molecular weight of the solvent molecules?

To investigate this, the solvent molecules in the first solvation layer of the sheet were identified at 30 ns of the 60 ns simulation, and tracked for the next nanosecond. Here the first solvation layer is defined as solvent molecules whose COM is within 0.5 nm of the sheet, based on the solvent distribution around the sheet (Appendix A, Section A4). Simulation snapshots in Appendix A Section A4 shows that as the solvent size increases, the molecules in the first solvation layer tend to adhere more stably to the sheets. For instance, 80% of the initially adhering BD molecules remain in the first solvation layer while the corresponding number is only 5% for water. Similar observations are made in the US simulations for two nanosheets (Appendix A, Section A4), where after the first 1 ns 82%, 17%, and 5% of the solvent molecules remain in the first solvation layer, respectively for BD, MET, and water. Water behaves much more dynamically, and there are frequent exchanges between water molecules on the interface and those in the bulk. In contrast, the larger MET and especially BD molecules can maintain their positions around and even between the sheets (Appendix A, Section A4). Their lower mobility near the sheet and less exchange with the bulk are consistent with the higher value of  $|\Delta H_{\text{sol}}|$ , suggesting more stably established vdW and electrostatic interaction between the sheet and the solvent.

### 3.3.3.2. Group 2 (solvents with carbonyl group)

Similar to group 1, solvent molecules in group 2 are also linear molecules with polar groups (albeit different) capable of forming H-bonds. It is therefore not surprising that the result of  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  for these two groups show similar behaviors, and data in Figure 3.3b for all solvents in these two groups form a single class that can be clearly distinguished from group 3. Analysis in Appendix A Section A5 shows that solvation of the g-C<sub>3</sub>N<sub>4</sub> nanosheet in group 2 solvents is also driven by enthalpy. The  $T\Delta S_{\text{sol}}$  ( $< 0$ ) term is in the same range for all three solvents;  $\Delta H_{\text{sol}}$  ( $< 0$ ) for ACE and FRM is close, while the magnitude of  $\Delta H_{\text{sol}}$  is much larger for DMF. Consequently, DMF also has the highest  $\Delta G_{\text{sol}}$  and lowest  $\Delta G_{\text{exf}}$ . Considering data from groups 1 and 2 together in Figure 3.3, DMF stands out as the solvent with the best potential to exfoliate g-C<sub>3</sub>N<sub>4</sub> nanosheets.

Unlike group 1, the ranking of  $\Delta G_{\text{sol}}$  in group 2 does not follow the order of the  $M_{\text{W}}$  of the solvent molecules. In particular, FRM has a smaller  $M_{\text{W}}$  (45.0) than ACE (58.1), yet the two solvents have similar  $\Delta G_{\text{sol}}$ . Considering groups 1 and 2 together, BD has larger  $M_{\text{W}}$  than DMF, yet its  $\Delta G_{\text{sol}}$  is lower. Therefore, it appears that when the solvent molecules possess a carbonyl group, their interaction with the sheet is not directly correlated with their size. The radial

distribution function (RDF,  $g(r)$ ) of the solvent oxygen around N atoms of the sheet is shown in Appendix A Section A6 for all solvents in groups 1 and 2, based on the last 30 ns of the 60 ns simulations. The location of the first peak ( $r_p$ ) in each curve corresponds to the accumulation of the solvent oxygen near the N atoms of the sheet in the first hydration layer. A negative correlation is observed between  $r_p$  and  $\Delta G_{\text{sol}}$  (Figure 3.4b). In other words, solvents that can orient their oxygen closer to the surface of the sheet tend to have a higher  $\Delta G_{\text{sol}}$ . Among the solvents from groups 1 and 2, DMF has the lowest value of  $r_p$ . One possible reason could be the partial atomic charge of the solvent oxygen, which is shown in inset of Figure 3.4b. The charge is most negative for water oxygen and least negative for DMF. Since the nitrogen atoms on the g-C<sub>3</sub>N<sub>4</sub> nanosheet also carry a negative partial charge (-0.35), less negative charge of the solvent oxygen can alleviate their local repulsion with the nitrogen, allowing other attractive interactions (vdW, electrostatic force between atoms with opposing partial atomic charges) to be more stably established.

In Section 3.3.3.1, it is discussed that  $\Delta G_{\text{sol}}$  does not have a direct correlation with the number of H-bonds broken and formed during the solvation; rather it is related to the mobility of the solvent molecules around the sheet. This motivates us to investigate the persistence of the solvent around the sheet by calculating the lifetime of the H-bonds between them, through the following autocorrelation function [59]:

$$C(\tau) = \langle s_i(t) s_i(t + \tau) \rangle \quad (3.6)$$

Here  $s_i(t)$  indicates whether H-bond  $i$  is present ( $s_i = 1$ ) or absent ( $s_i = 0$ ) at time  $t$ , and  $\langle \bullet \rangle$  performs an average over all time  $t$  and all H-bonds  $i$  between the solvent and the sheet. The integral  $\int_0^\infty C(\tau) d\tau$  gives a rough estimate of the average H-bond lifetime, which is shown in Figure 3.4b for group 1 and 2 solvents. Interestingly, the average H-bond lifetime follows an almost linear relationship with  $\Delta G_{\text{sol}}$ . Consistent with the visual observations in Appendix A Section A4 this result confirms that higher magnitude of  $\Delta G_{\text{sol}}$  is associated with less mobility and more stable adherence of the solvent molecules around the sheet.

### 3.3.3.3. Group 3 (solvents with aromatic structures):

Group 3 solvents have a distinct structure compared with group 1 and 2 solvents, containing a ring-like, instead of linear, structure. Their  $\Delta G_{\text{exf}}$  vs.  $\Delta G_{\text{sol}}$  data are separated from the other two

groups in Figure 3.3b, and within similar range of  $\Delta G_{\text{sol}}$  this group tends to have lower  $\Delta G_{\text{exf}}$ . Energetics analysis in Appendix A Section A5 shows that the solvation of the g-C<sub>3</sub>N<sub>4</sub> nanosheet in group 3 solvent is still enthalpy-driven. Since these solvents contain aromatic structures, it is of interest to study potential  $\pi$ - $\pi$  interactions between solvent molecules and the sheet. Previous DFT simulations investigated the adsorption of ionic liquid on BN nanosheets [60], and reported the presence of  $\pi$ - $\pi$ , CH- $\pi$ , and anion- $\pi$  interactions.  $\pi$ - $\pi$  interactions between two aromatic rings are typically manifested by a small separation (less than 5.0 Å [61]) between them [62]. Appendix A Section A6 provides the RDFs of the center of the aromatic ring around the carbon and nitrogen atoms on the sheet (data based on last 30 ns of the 60 ns simulations from the last set of simulations in Table 3.1). The curves corresponding to different solvents almost overlap, demonstrating similar distribution of the aromatic ring around the sheet. The location of the first RDF peak is 4.2 Å for carbon and 3.8 Å for nitrogen, suggesting that the aromatic centers are positioned slightly closer to the nitrogen atoms. To measure the relative orientation of the solvent aromatic rings with respect to the sheet, Figure 3.4c shows the probability distribution of the angle between the normal of the aromatic rings in the solvents and the normal of the solvated sheet. Here the solvent molecules considered are in the first solvation layer and the probability distribution is generated based on last 30 ns of the 60 ns simulation. For all three solvents most of the molecules have an angle  $> 80^\circ$ , corresponding to nearly perpendicular orientation relative to the sheets. NMP, however, has a higher probability of acquiring smaller angles ( $< 10^\circ$ ) than the other two solvents. This suggests that while all solvents experience T-shaped  $\pi$ - $\pi$  interactions [63] with the sheet, NMP benefits from having more offset-stacked (parallel-displaced) [64]. This conformation in turn promotes the solvent-sheet interactions and makes NMP a better medium for LPE (highest  $\Delta G_{\text{sol}}$  and lowest  $\Delta G_{\text{exf}}$  in this group).

### 3.3.4. Discussion

#### 3.3.4.1. Mobility of solvent molecules around nanosheet

In sections 3.3.3.1 and 3.3.3.2, an interesting correlation is observed between  $\Delta G_{\text{sol}}$  and the mobility of the linear solvent molecules around the sheet. Here the analysis is extended to solvents in all three groups. For each 60 ns simulation where a single nanosheet was equilibrated in a solvent, the solvent molecules in the first solvation layer at  $t = 25$  ns were tracked till the end of the simulation. The fraction of these molecules that departed from the first solvation layer ( $M$ ) was

calculated as a function of time and presented in Appendix A Section A7.  $M$  starts from zero and increases with time, reaching a plateau at  $t = 50$  ns for all solvents. The average from the last 10 ns (50 to 60 ns), denoted by  $M_{ave}$ , is shown in Figure 3.5a by plotting  $\ln(M_{ave})$  against  $\Delta G_{sol}/kT$ , where  $k$  is the Boltzmann constant. All data fall near a straight line, suggesting the following relationship between  $M_{ave}$  and  $\Delta G_{sol}$ :

$$M_{ave} = A e^{\frac{-a\Delta G_{sol}}{kT}} \quad (3.7)$$

where  $A$  and  $a > 0$  are constants. This is a relationship that resembles the Arrhenius equation [65], with  $\Delta G_{sol}$  serving the role of an activation energy that is required to drive the molecules to depart from the first solvation layer and diffuse into the bulk.

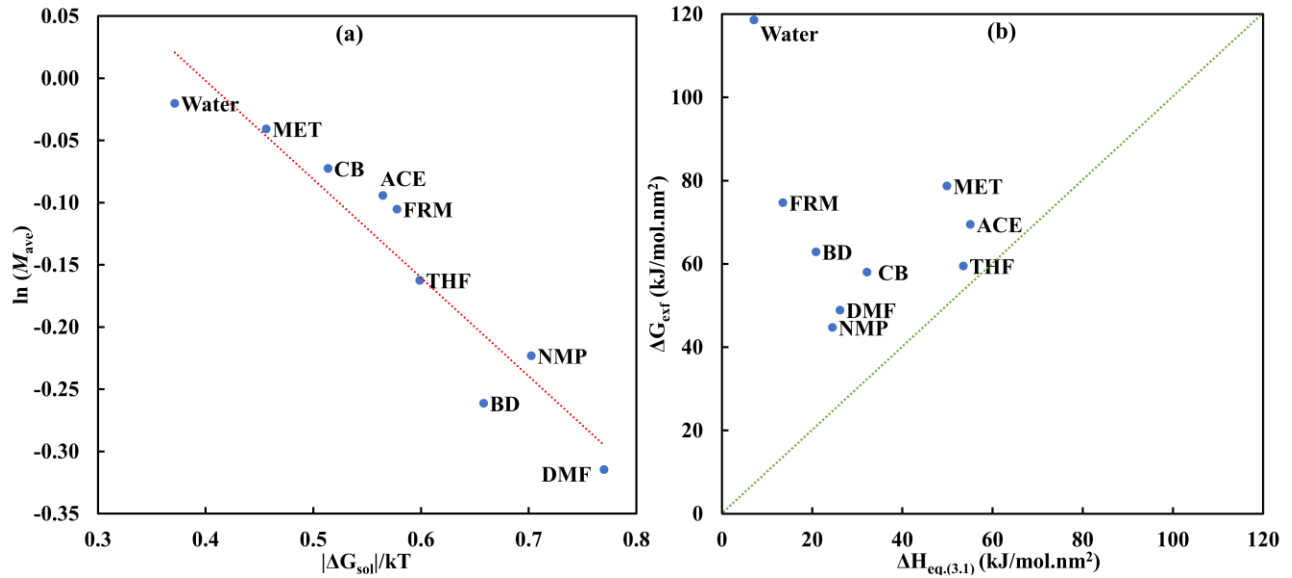


Figure 3.5. (a)  $\ln(M_{ave})$  versus  $\frac{\Delta G_{sol}}{kT}$  for all the solvents studied in this work. (b) Comparison of  $\Delta H_{eq.(3.1)}$  and  $\Delta G_{ext}$  from this study.

### 3.3.4.2. Implications for LPE

Previous studies [29] have reported that the concentration of the as-obtained g-C<sub>3</sub>N<sub>4</sub> nanosheets in suspension was extremely low (less than 3 mg/mL). Fabricating graphene-like g-C<sub>3</sub>N<sub>4</sub> with a single atomic layer in a relatively high concentration remains a strenuous task since the g-C<sub>3</sub>N<sub>4</sub> nanosheets tend to aggregate due to their high surface energy [66]. Hence, a key goal



of the studies on LPE is to find the most suitable dispersant–solvent system that can generate uniform and stable dispersion of g-C<sub>3</sub>N<sub>4</sub> sheets with high concentration. Smaller values of  $\Delta G_{\text{exf}}$  correspond to less amount of energy required to separate aggregated sheets, as well as less tendency for the sheets to aggregate [57]. It is therefore expected that the concentration of dispersed g-C<sub>3</sub>N<sub>4</sub> sheets can reach higher values in the solvents with lower  $\Delta G_{\text{exf}}$ , which has been verified experimentally for graphene solvated in ionic solvents [67]. Our results showed that NMP, DMF, CB, and THF have the lowest  $\Delta G_{\text{exf}}$ , therefore they are predicted to be better media for LPE. This is in line with the experiments by Yang et al. where g-C<sub>3</sub>N<sub>4</sub> nanosheets were found to be more stably dispersed in NMP and IPA among five solvents (NMP, IPA, water, ACE, and ethanol) [28]. It should be noted that NMP has a high boiling point and is challenging to be removed in order to obtain exfoliated g-C<sub>3</sub>N<sub>4</sub> nanosheets. The evaporation process is slow and aggregation of exfoliated nanosheets may occur during the process [28]. This provides quantitative and molecular-level support for why DMF is widely used as an effective exfoliation medium in experimental studies [28,68–72].

In drastic contrast to NMP and DMF, water has a very large  $\Delta G_{\text{exf}}$ , 52% higher than the second poorest solvent (MET) in this study.  $\Delta G_{\text{sol}}$  of water is also the lowest, about 23% smaller than the next in line (MET). This suggests low affinity of water molecules to g-C<sub>3</sub>N<sub>4</sub> sheets, which is manifested through its lowest average H-bond lifetime, furthest distribution from the sheet (largest  $r_p$  value), and highest molecular mobility in the adsorption layer. Based on the data, water is not a suitable candidate for the LPE of g-C<sub>3</sub>N<sub>4</sub>. The predicted poor performance of water is in line with experimental studies where a relatively low concentration of dispersed g-C<sub>3</sub>N<sub>4</sub> sheets was found in water [29,31]. Recognizing the advantages of water being a solvent (e.g., safe, abundant, easily accessible), surface modifications of g-C<sub>3</sub>N<sub>4</sub> sheets may be considered to enhance their interactions with water and increase their aqueous dispersibility.

Sresht et al. investigated the LPE of phosphorene sheets using MD simulations [44]. It was discussed that the performance of a solvent depended on its molecular shape, and solvents with a planar structure, such as NMP and dimethyl sulfoxide, behave like molecular wedges that can intercalate more efficiently [44]. In agreement with this, our data show that for similar  $\Delta G_{\text{sol}}$ , group 3 solvents with planar structures tend to have smaller  $\Delta G_{\text{exf}}$ . More interestingly, our results have repeatedly shown a correlation between  $\Delta G_{\text{exf}}$  and the mobility of solvent molecules around the

nanosheet. For example, for DMF, NMP, and BD, up to 30% of the solvents in the first solvation layer lingered near the sheet even after 35 ns (Appendix A Section A7). Those solvent molecules essentially formed an adsorption layer on the nanosheet, which could shield the sheet-sheet interactions, potentially hindering the aggregation and allowing stable dispersion of individual sheets. Based on our observations, several suggestions can be made that might promote the stability of the adsorption layer. Firstly, highly negative partial atomic charges (typically related to high polarity) should be avoided, so as to reduce the repulsion with the electronegative nitrogen on the sheet. Secondly, solvents with larger size (higher  $M_w$ ) present an advantage. Finally, solvents with an aromatic structure are preferred, especially those that are able to form more parallel  $\pi$ - $\pi$  stacked conformations.

Energetics analysis in this work has demonstrated enthalpy as the main driving force in the solvation of g-C<sub>3</sub>N<sub>4</sub>. While the dominance of enthalpy has been reported in the literature for the solvation of long-chain polymers [73,74], to our knowledge this is the first time it is reported for 2D materials. Coleman et al. [22] proposed a first-order estimation, eq. (3.1), for the enthalpy of mixing of graphene as an approximation for the free energy of exfoliation. To examine the validity of this approximation for g-C<sub>3</sub>N<sub>4</sub>, we used  $\phi = \frac{V_{sheet}}{V_{mix}}$  to rewrite eq. (3.1) into

$$\Delta H_{eq(3.1)} \equiv \frac{\Delta H_{mix} T_{sheet}}{V_{sheet}} = 2 \left( \sqrt{E_{sur}^{sheet}} - \sqrt{E_{sur}^{solvent}} \right)^2$$

which represents the enthalpy of mixing per unit area of the sheet.  $\Delta H_{eq. (3.1)}$  was calculated using the surface energy values from Appendix A Tables A2 and A3, respectively for  $E_{sur}^{sheet}$  and  $E_{sur}^{solvent}$ , and compared with  $\Delta G_{exf}$  from our work (Figure 3.5b). For all of the solvents,  $\Delta H_{eq.(2.1)}$  underestimates  $\Delta G_{exf}$  and the discrepancy is particularly large for solvents with high surface tension (such as water and FRM). The ranking of  $\Delta H_{eq.(2.1)}$  follows: ACE > THF > MET > CB > DMF > NMP > BD > FRM > water. If  $\Delta H_{eq.(2.1)}$  were to be used to select the solvent for LPE, then water would be chosen as the best solvent among these nine solvents. This is in contradiction with our earlier discussion on the performance of water in LPE. Therefore, eq. (3.1) is a poor estimation for the free energy of exfoliation for g-C<sub>3</sub>N<sub>4</sub> and should not be used as a selection criterion.

### 3.4. Conclusion

In this study, MD simulations are performed along with extensive potential of mean force (PMF) calculations, to evaluate the free energy of exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets in nine common solvents. The solvation free energy of a single g-C<sub>3</sub>N<sub>4</sub> nanosheet is also calculated and compared with the free energy of exfoliation of two g-C<sub>3</sub>N<sub>4</sub> nanosheets in the same solvent. The effect of structural properties of solvents on the free energy cost of the exfoliation process is analyzed. Our results show that similar to other 2D materials such as BN [46] and graphene [47], the most probable path for the exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets is in the parallel (shear) direction. Based on PMF calculations, the performance of the nine solvents in exfoliating or dispersing g-C<sub>3</sub>N<sub>4</sub> sheets is ranked, with the best solvents being NMP, DMF, CB, and THF. This provides quantitative and molecular-level support for why NMP and DMF are widely used as exfoliation media in experimental studies [28,68–72]. A high correlation is found between the free energy of exfoliation of two sheets and the solvation free energy of a single sheet, with higher magnitude of solvation free energy corresponding to lower free energy of exfoliation. Regardless of the molecular structure of the solvents, the solvation of g-C<sub>3</sub>N<sub>4</sub> sheet is driven primarily by enthalpy. Analysis of the first solvation layer shows that solvents with higher magnitude of solvation free energy tend to be less mobile in this layer, and in some cases a stable adsorption layer is formed around the sheet. Additionally, our results show that the first-order estimation for the free energy of exfoliation proposed by Coleman et al. [22] is insufficient and can lead to large errors especially for water.

To our best knowledge, this is the first atomistic-level study that determines a quantitative relationship between solvent properties and the performance of liquid phase exfoliation. Not only have the simulations been able to provide explanations for solvent selection in some experiments [28,68–72], they have also generated critical insights into the underlying mechanisms which are not accessible by experiments. Also, through several novel analyses, we have for the first time revealed key factors that govern the efficacy of the solvent. This has allowed us to propose a comprehensive set of principles for the selection and design of effective solvents in liquid phase exfoliation of g-C<sub>3</sub>N<sub>4</sub>, including less negative partial atomic charges, higher molecular weight and the presence of aromatic structures.

Several future directions are recognized. First, in the literature there are two main molecular structures proposed for g-C<sub>3</sub>N<sub>4</sub> nanosheets: heptazine based and *s*-triazine based [26]. Here, we only considered the former and it is of interest to investigate the latter structure. It is also worth studying whether the geometry of the model (e.g., triangle in this work vs. other shapes) affects the free energy results. Second, there exist many other solvents with structural features different from those studied in this work. Extending the present framework to a large number of solvents may allow for a more general model to be established that can predict the free energy of exfoliation. Lastly, future work can be conducted to relate the free energy of exfoliation to critical concentrations of delaminated g-C<sub>3</sub>N<sub>4</sub> in a solvent that lead to aggregation and precipitation [29].

### 3.5. Acknowledgement

Compute Canada is gratefully acknowledged for providing the computing resources and technical support. TT acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Grant numbers: RGPIN-2018-04281, RGPAS-2018-522655). KS thanks NSERC for financial support (Grant number: RGPIN-2020-04620).

### 3.6. References

- [1] J. Fu, J. Yu, C. Jiang, B. Cheng, g-C<sub>3</sub>N<sub>4</sub>-Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.
- [2] X. Ma, J. Hu, H. He, S. Dong, C. Huang, X. Chen, New Understanding on Enhanced Photocatalytic Activity of g-C<sub>3</sub>N<sub>4</sub>/BiPO<sub>4</sub> Heterojunctions by Effective Interfacial Coupling, *ACS Appl Nano Mater* 1 (2018) 5507–5515. <https://doi.org/10.1021/acsanm.8b01012>.
- [3] Y. Yuan, L. Zhang, J. Xing, M.I.B. Utama, X. Lu, K. Du, Y. Li, X. Hu, S. Wang, A. Genç, R. Dunin-Borkowski, J. Arbiol, Q. Xiong, High-yield synthesis and optical properties of g-C<sub>3</sub>N<sub>4</sub>, *Nanoscale* 7 (2015) 12343–12350. <https://doi.org/10.1039/c5nr02905h>.
- [4] J. Xu, L. Zhang, R. Shi, Y. Zhu, Chemical exfoliation of graphitic carbon nitride for efficient heterogeneous photocatalysis, *J Mater Chem A Mater* 1 (2013) 14766–14772. <https://doi.org/10.1039/C3TA13188B>.

- [5] Y. Liu, D. Xie, M. Song, L. Jiang, G. Fu, L. Liu, J. Li, Water desalination across multilayer graphitic carbon nitride membrane: Insights from non-equilibrium molecular dynamics simulations, *Carbon N Y* 140 (2018) 131–138. <https://doi.org/10.1016/j.carbon.2018.08.043>.
- [6] A. Alaghmandfard, K. Ghandi, A Comprehensive Review of Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)–Metal Oxide-Based Nanocomposites: Potential for Photocatalysis and Sensing, *Nanomaterials* 2022, Vol. 12, Page 294 12 (2022) 294. <https://doi.org/10.3390/NANO12020294>.
- [7] Y. Dong, Q. Wang, H. Wu, Y. Chen, C.H. Lu, Y. Chi, H.H. Yang, Graphitic Carbon Nitride Materials: Sensing, Imaging and Therapy, *Small* 12 (2016) 5376–5393. <https://doi.org/10.1002/SMLL.201602056>.
- [8] S. Patnaik, D.P. Sahoo, K. Parida, An overview on Ag modified g-C<sub>3</sub>N<sub>4</sub> based nanostructured materials for energy and environmental applications, *Renewable and Sustainable Energy Reviews* 82 (2018) 1297–1312. <https://doi.org/10.1016/J.RSER.2017.09.026>.
- [9] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [10] X. Dong, F. Cheng, Recent development in exfoliated two-dimensional g-C<sub>3</sub>N<sub>4</sub> nanosheets for photocatalytic applications, *J Mater Chem A Mater* 3 (2015) 23642–23652. <https://doi.org/10.1039/c5ta07374j>.
- [11] S. Cao, J. Low, J. Yu, M. Jaroniec, Polymeric Photocatalysts Based on Graphitic Carbon Nitride, *Advanced Materials* 27 (2015) 2150–2176. <https://doi.org/10.1002/adma.201500033>.
- [12] Y. Zheng, L. Lin, B. Wang, X. Wang, Graphitic Carbon Nitride Polymers toward Sustainable Photoredox Catalysis, *Angewandte Chemie International Edition* 54 (2015) 12868–12884. <https://doi.org/10.1002/anie.201501788>.
- [13] K.M. Alam, C.E. Jensen, P. Kumar, R.W. Hooper, G.M. Bernard, A. Patidar, A.P. Manuel, N. Amer, A. Palmgren, D.N. Purschke, N. Chaulagain, J. Garcia, P.S. Kirwin, L.C.T. Shoute, K. Cui, S. Gusarov, A.E. Kobryn, V.K. Michaelis, F.A. Hegmann, K. Shankar, Photocatalytic Mechanism Control and Study of Carrier Dynamics in CdS@C<sub>3</sub>N<sub>5</sub>Core-Shell Nanowires, *ACS Appl Mater Interfaces* 13 (2021) 47418–47439.

[https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550\\_0011.JPEG](https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550_0011.JPEG)

.

- [14] S. Angizi, M. Ali Akbar, M. Darestani-Farahani, al -, S. Dhanraj Nehate, S. Sundaresh, R. Peale, P. Kumar, S. Mulmi, D. Laishram, K.M. Alam, U.K. Thakur, V. Thangadurai, K. Shankar, Water-splitting photoelectrodes consisting of heterojunctions of carbon nitride with a p-type low bandgap double perovskite oxide, *Nanotechnology* 32 (2021) 485407. <https://doi.org/10.1088/1361-6528/ABEDEC>.
- [15] S. Cao, H. Chen, F. Jiang, Z. Hu, X. Wang, Construction of Acetaldehyde-Modified g-C<sub>3</sub>N<sub>4</sub> Ultrathin Nanosheets via Ethylene Glycol-Assisted Liquid Exfoliation for Selective Fluorescence Sensing of Ag<sup>+</sup>, *ACS Appl Mater Interfaces* 10 (2018) 44624–44633. [https://doi.org/10.1021/ACSAMI.8B15501/ASSET/IMAGES/LARGE/AM-2018-15501V\\_0002.JPEG](https://doi.org/10.1021/ACSAMI.8B15501/ASSET/IMAGES/LARGE/AM-2018-15501V_0002.JPEG).
- [16] B. Abreu, B. Almeida, P. Ferreira, R. M. F. Fernandes, D.M. Fernandes, E.F. Marques, A critical assessment of the role of ionic surfactants in the exfoliation and stabilization of 2D nanosheets: The case of the transition metal dichalcogenides MoS<sub>2</sub>, WS<sub>2</sub> and MoSe<sub>2</sub>, *J Colloid Interface Sci* 626 (2022) 167–177. <https://doi.org/10.1016/J.JCIS.2022.06.097>.
- [17] H. Zhao, H. Wu, J. Wu, J. Li, Y. Wang, Y. Zhang, H. Liu, Preparation of MoS<sub>2</sub>/WS<sub>2</sub> nanosheets by liquid phase exfoliation with assistance of epigallocatechin gallate and study as an additive for high-performance lithium-sulfur batteries, *J Colloid Interface Sci* 552 (2019) 554–562. <https://doi.org/10.1016/J.JCIS.2019.05.080>.
- [18] Q. Wan, H. Wang, S. Li, J. Wang, Efficient liquid-phase exfoliation of few-layer graphene in aqueous 1, 1, 3, 3-tetramethylurea solution, *J Colloid Interface Sci* 526 (2018) 167–173. <https://doi.org/10.1016/J.JCIS.2018.04.110>.
- [19] C. Huo, Z. Yan, X. Song, H. Zeng, 2D materials via liquid exfoliation: a review on fabrication and applications, *Sci Bull (Beijing)* 60 (2015) 1994–2008. <https://doi.org/10.1007/S11434-015-0936-3>.

- [20] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon N Y* 141 (2019) 580–607. <https://doi.org/10.1016/J.CARBON.2018.09.082>.
- [21] Y. Hernandez, M. Lotya, D. Rickard, S.D. Bergin, J.N. Coleman, Measurement of multicomponent solubility parameters for graphene facilitates solvent discovery, *Langmuir* 26 (2010) 3208–3213. <https://doi.org/10.1021/la903188a>.
- [22] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. v. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuvsen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/SCIENCE.1194975>.
- [23] L. Ma, H. Fan, J. Wang, Y. Zhao, H. Tian, G. Dong, Water-assisted ions in situ intercalation for porous polymeric graphitic carbon nitride nanosheets with superior photocatalytic hydrogen evolution performance, *Appl Catal B* 190 (2016) 93–102. <https://doi.org/10.1016/j.apcatb.2016.03.002>.
- [24] C. Wu, S. Xue, Z. Qin, M. Nazari, G. Yang, S. Yue, T. Tong, H. Ghasemi, F.C.R. Hernandez, S. Xue, D. Zhang, H. Wang, Z.M. Wang, S. Pu, J. Bao, Making g-C<sub>3</sub>N<sub>4</sub> ultra-thin nanosheets active for photocatalytic overall water splitting, *Appl Catal B* 282 (2021) 119557. <https://doi.org/10.1016/j.apcatb.2020.119557>.
- [25] S.P. Pattnaik, A. Behera, S. Martha, R. Acharya, K. Parida, Facile synthesis of exfoliated graphitic carbon nitride for photocatalytic degradation of ciprofloxacin under solar irradiation, *J Mater Sci* 54 (2019) 5726–5742. <https://doi.org/10.1007/s10853-018-03266-x>.
- [26] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.

- [27] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenís, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045. <https://doi.org/10.1021/acsami.5b06974>.
- [28] S. Yang, Y. Gong, J. Zhang, L. Zhan, L. Ma, Z. Fang, R. Vajtai, X. Wang, P.M. Ajayan, Exfoliated graphitic carbon nitride nanosheets as efficient catalysts for hydrogen evolution under visible light, *Advanced Materials* 25 (2013) 2452–2456. <https://doi.org/10.1002/adma.201204453>.
- [29] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [30] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/science.1194975>.
- [31] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight from molecular dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954. <https://doi.org/10.1007/s10934-021-01047-7>.
- [32] Y. Zhang, C. Shen, X. Lu, X. Mu, P. Song, Effects of defects in g-C<sub>3</sub>N<sub>4</sub> on excited-state charge distribution and transfer: Potential for improved photocatalysis, *Spectrochim Acta A Mol Biomol Spectrosc* 227 (2020) 117687. <https://doi.org/10.1016/J.SAA.2019.117687>.
- [33] X. Zhou, M. Zhu, L. Kang, Single-Atom X/g-C<sub>3</sub>N<sub>4</sub> (X = Au<sup>1</sup>, Pd<sup>1</sup>, and Ru<sup>1</sup>) Catalysts for Acetylene Hydrochlorination: A Density Functional Theory Study, *Catalysts* 2019, Vol. 9, Page 808 9 (2019) 808. <https://doi.org/10.3390/CATAL9100808>.



- [34] X. Chen, R. Hu, DFT-based study of single transition metal atom doped g-C<sub>3</sub>N<sub>4</sub> as alternative oxygen reduction reaction catalysts, *Int J Hydrogen Energy* 44 (2019) 15409–15416. <https://doi.org/10.1016/J.IJHYDENE.2019.04.057>.
- [35] Z. Bonakchi, A. Nakhaei Pour, S. Soheili, Molecular simulation of methane on various g-C<sub>3</sub>N<sub>4</sub> isomers: collision, adsorption, desorption, and diffusion studies, *Journal of the Iranian Chemical Society* 19 (2022) 3649–3657. <https://doi.org/10.1007/S13738-022-02562-3/TABLES/2>.
- [36] L. Kang, M. Zhu, Y. Zhao, A DFT Study of Acetylene Hydrogenation Catalyzed by S-Doped Pd 1 /g-C 3 N 4, (n.d.). <https://doi.org/10.3390/catal9110887>.
- [37] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906. <https://doi.org/10.1021/acs.langmuir.7b02441>.
- [38] W.L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, *Proc Natl Acad Sci U S A* 102 (2005) 6665–6670. <https://doi.org/10.1073/pnas.0408037102>.
- [39] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017). <https://doi.org/10.1093/nar/gkx312>.
- [40] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a. <https://doi.org/10.2477/jccjie.2018-0034>.
- [41] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [42] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).

- [43] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [44] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268. <https://doi.org/10.1021/acsnano.5b02683>.
- [45] A. Gotzias, Binding Free Energy Calculations of Bilayer Graphenes Using Molecular Dynamics, *J Chem Inf Model* 61 (2021) 1164–1171. <https://doi.org/10.1021/acs.jcim.1c00043>.
- [46] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.
- [47] C. Fu, X. Yang, Molecular simulation of interfacial mechanics for solvent exfoliation of graphene from graphite, *Carbon N Y* 55 (2013) 350–360. <https://doi.org/10.1016/j.carbon.2012.12.083>.
- [48] G. Zhou, P. Rajak, S. Susarla, P.M. Ajayan, R.K. Kalia, A. Nakano, P. Vashishta, Molecular Simulation of MoS<sub>2</sub> Exfoliation, *Sci Rep* 8 (2018) 1–9. <https://doi.org/10.1038/s41598-018-35008-z>.
- [49] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).
- [50] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1–2 (2015) 19–25. <https://doi.org/10.1016/J.SOFTX.2015.06.001>.
- [51] V. Georgakilas, J.N. Tiwari, K.C. Kemp, J.A. Perman, A.B. Bourlinos, K.S. Kim, R. Zboril, Noncovalent Functionalization of Graphene and Graphene Oxide for Energy Materials, Biosensing, Catalytic, and Biomedical Applications, *Chem Rev* 116 (2016) 5464–5519. <https://doi.org/10.1021/ACS.CHEMREV.5B00620>.

- [52] A.M. Silva, M.I. Rojas, Electric and structural properties of polymeric graphite carbon nitride (g-C<sub>3</sub>N<sub>4</sub>): A Density Functional Theory study, *Comput Theor Chem* (2016).  
<https://doi.org/10.1016/j.comptc.2016.11.004>.
- [53] B. Mortazavi, Ultra high stiffness and thermal conductivity of graphene like C<sub>3</sub>N, *Carbon N Y* 118 (2017) 25–34. <https://doi.org/10.1016/J.CARBON.2017.03.029>.
- [54] D. Wu, D.A. Kofke, Phase-space overlap measures. II. Design and implementation of staging methods for free-energy calculations, *J Chem Phys* 123 (2005) 084109.  
<https://doi.org/10.1063/1.2011391>.
- [55] H.K. Choi, Y. Oh, H. Jung, H. Hong, B.C. Ku, N.H. You, Y.K. Kim, E.S. Shin, J. Yu, Influences of carboxyl functionalization of intercalators on exfoliation of graphite oxide: A molecular dynamics simulation, *Physical Chemistry Chemical Physics* 20 (2018) 28616–28622.  
<https://doi.org/10.1039/c8cp05436c>.
- [56] H.A. Yu, M. Karplus, A thermodynamic analysis of solvation, *J Chem Phys* 89 (1998) 2366.  
<https://doi.org/10.1063/1.455080>.
- [57] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648.  
<https://doi.org/10.1021/ja1064284>.
- [58] R. Biswas, Molecular dynamics studies on the exfoliation of graphene in room temperature ionic liquids, *J Mol Liq* 337 (2021) 116592. <https://doi.org/10.1016/j.molliq.2021.116592>.
- [59] A. Luzar, D. Chandler, Hydrogen-bond kinetics in liquid water, *Nature* 1996 379:6560 379 (1996) 55–57. <https://doi.org/10.1038/379055a0>.
- [60] M. Shakourian-Fard, G. Kamath, Z. Jamshidi, Trends in Physisorption of Ionic Liquids on Boron-Nitride Sheets, (2014). <https://doi.org/10.1021/jp506277n>.
- [61] T. Chen, M. Li, J. Liu,  $\pi$ - $\pi$  Stacking Interaction: A Nondestructive and Facile Means in Material Engineering for Bioapplications, *Cryst Growth Des* 18 (2018) 2765–2783.  
[https://doi.org/10.1021/ACS.CGD.7B01503/ASSET/IMAGES/LARGE/CG-2017-01503A\\_0004.JPEG](https://doi.org/10.1021/ACS.CGD.7B01503/ASSET/IMAGES/LARGE/CG-2017-01503A_0004.JPEG).

- [62] C.A. Hunter, J.K.M. Sanders, The nature of  $\pi$ - $\pi$  interactions, *J Am Chem Soc* 112 (2002) 5525–5534. <https://doi.org/10.1021/JA00170A016>.
- [63] A.L. Ringer, M.O. Sinnokrot, R.P. Lively, C.D. Sherrill, The effect of multiple substituents on sandwich and T-shaped  $\pi$ - $\pi$  interactions, *Chemistry* 12 (2006) 3821–3828. <https://doi.org/10.1002/CHEM.200501316>.
- [64] M. J. Rashkin, M. L. Waters, Unexpected Substituent Effects in Offset  $\pi$ - $\pi$  Stacked Interactions in Water, *J Am Chem Soc* 124 (2002) 1860–1861. <https://doi.org/10.1021/ja016508z>.
- [65] M. Peleg, M.D. Normand, M.G. Corradini, The Arrhenius Equation Revisited, <https://doi.org/10.1080/10408398.2012.667460> 52 (2012) 830–851. <https://doi.org/10.1080/10408398.2012.667460>.
- [66] L. Sun, Y. Qi, C.J. Jia, Z. Jin, W. Fan, Enhanced visible-light photocatalytic activity of g-C<sub>3</sub>N<sub>4</sub>/Zn<sub>2</sub>GeO<sub>4</sub> heterojunctions with effective interfaces based on band match, *Nanoscale* 6 (2014) 2649–2659. <https://doi.org/10.1039/C3NR06104C>.
- [67] E. Bordes, B. Morcos, D. Bourgoigne, J.M. Andanson, P.O. Bussière, C.C. Santini, A. Benayad, M.C. Gomes, A.A.H. Pádua, Dispersion and stabilization of exfoliated graphene in ionic liquids, *Front Chem* 7 (2019) 223. <https://doi.org/10.3389/FCHEM.2019.00223/BIBTEX>.
- [68] Y.J. Yuan, Z. Shen, S. Wu, Y. Su, L. Pei, Z. Ji, M. Ding, W. Bai, Y. Chen, Z.T. Yu, Z. Zou, Liquid exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets to construct 2D-2D MoS<sub>2</sub>/g-C<sub>3</sub>N<sub>4</sub> photocatalyst for enhanced photocatalytic H<sub>2</sub> production activity, *Appl Catal B* 246 (2019) 120–128. <https://doi.org/10.1016/j.apcatb.2019.01.043>.
- [69] Q. Hao, Y. Song, H. Ji, Z. Mo, X. She, J. Deng, T. Muhmood, X. Wu, S. Yuan, H. Xu, H. Li, Surface N modified 2D g-C<sub>3</sub>N<sub>4</sub> nanosheets derived from DMF for photocatalytic H<sub>2</sub> evolution, *Appl Surf Sci* 459 (2018) 845–852. <https://doi.org/10.1016/j.apsusc.2018.07.154>.
- [70] Y. Wang, L. Li, Y. Wei, J. Xue, H. Chen, L. Ding, J. Caro, H. Wang, Water Transport with Ultralow Friction through Partially Exfoliated g-C<sub>3</sub>N<sub>4</sub> Nanosheet Membranes with Self-Supporting Spacers, *Angewandte Chemie - International Edition* 56 (2017) 8974–8980. <https://doi.org/10.1002/anie.201701288>.

- [71] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [72] L. Ma, H. Fan, M. Li, H. Tian, J. Fang, G. Dong, A simple melamine-assisted exfoliation of polymeric graphitic carbon nitrides for highly efficient hydrogen production from water under visible light, *J Mater Chem A Mater* 3 (2015) 22404–22412. <https://doi.org/10.1039/C5TA05850C>.
- [73] K. Koschek, V. Durmaz, O. Krylova, M. Wieczorek, S. Gupta, M. Richter, A. Bujotzek, C. Fischer, R. Haag, C. Freund, M. Weber, J. Rademann, Peptide–polymer ligands for a tandem WW-domain, an adaptive multivalent protein–protein interaction: lessons on the thermodynamic fitness of flexible ligands, *Beilstein Journal of Organic Chemistry* 11:93 11 (2015) 837–847. <https://doi.org/10.3762/BJOC.11.93>.
- [74] F. Jiménez-Ángeles, H.-K. Kwon, K. Sadman, T. Wu, K. R. Shull, M. Olvera de la Cruz, Self-Assembly of Charge-Containing Copolymers at the Liquid–Liquid Interface, *ACS Cent Sci* 5 (2019) 688–699. <https://doi.org/10.1021/acscentsci.9b00084>.

# Chapter 4: Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning

## 4.1. Introduction

Graphitic carbon nitride, g-C<sub>3</sub>N<sub>4</sub>, has garnered significant attention owing to its unique properties and broad spectrum of potential applications. Its outstanding chemical stability, metal-free nature, and visible-light-responsive photocatalytic activity make it a promising material in fields such as energy conversion [1], environmental remediation [2], advanced membrane fabrication [3], chemical sensing [4], bio-imaging [5], and even in emerging areas like electronics and optoelectronics [6,7]. The critical bottleneck on enabling g-C<sub>3</sub>N<sub>4</sub> for their applications is the lack of industrial-scale methods to produce high-quality monolayer nanosheets. Many efforts have been spent on developing low-cost and large-scale procedures for high-quality g-C<sub>3</sub>N<sub>4</sub> production. Compared with other methods, LPE has a relatively high yield and low cost [8–10]. This method separates layered g-C<sub>3</sub>N<sub>4</sub> bulk material into monolayered nanosheets in a liquid medium by means of an external force such as sonication. For the LPE to be successful, the solvent and externally applied force must overcome interlayer interactions such as  $\pi$ - $\pi$  interaction and hydrogen bonding between neighboring sheets.

It has been widely reported both experimentally and computationally that different solvents have different performance in the LPE of 2D materials, due to their distinct physical properties and interactions with the nanosheets [11–20]. Two quantities can be used to measure the ability of a solvent to exfoliate 2D nanosheets. Free energy of exfoliation,  $\Delta G_{\text{exf}}$ , is defined as the free energy required to separate a unit area of two stacked sheets from their equilibrium distance [14,15,21]. Lower  $\Delta G_{\text{exf}}$  corresponds to less energy cost during LPE as well as easier stabilization of separated nanosheets after LPE. Solvation free energy of a nanosheet is the free energy change by bringing a unit area of the nanosheet from the gas phase into the solution. Solvents used for LPE typically produce negative  $\Delta G_{\text{sol}}$ , and higher absolute value of this quantity indicates better dispersibility. While the solvation free energy has been investigated in drug delivery [22], organic synthesis [23],

and solvation of nanoparticles [24], its quantification for 2D materials is much scarcer. In fact, to our best knowledge there had been no reports on the value of  $\Delta G_{\text{exf}}$  or  $\Delta G_{\text{sol}}$  for g-C<sub>3</sub>N<sub>4</sub> nanosheets until our recent work [21].

The evaluation of  $\Delta G_{\text{exf}}$  for g-C<sub>3</sub>N<sub>4</sub> involves the separation of two stacked nanosheets, which is experimentally difficult due to the required molecular resolution. Determining  $\Delta G_{\text{sol}}$  from experiments are possible but setting up its measurement is complicated and the results are prone to noises at ambient conditions [25,26]. MD simulation provides an alternative for capturing the molecular details and accurately calculating  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ , but doing so for every single solvent is extensively time and resource consuming. It is therefore of interest to combine MD data with other predictive tools to identify the key properties of solvents that govern  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ , and study their quantitative influence.

ML methods are now widely used to predict material properties such as band gap, enthalpy of formation and transition-state properties [27,28]. In the field of 2D materials, Saito et al. applied a deep-learning-based method for quality filtering of mechanically exfoliated 2D crystals [29]. Siriwardane et al. [30] developed energy-structure correlation using DFT and ML to reveal the exfoliation energy of MAB (where M is a transition metal, A is a group 13–16 element, and B is boron) phases. Also based on DFT data, Wan et al. [31] used an ML method to accurately predict the exfoliation energies for various 2D materials. It should be noted that the data in Siriwardane et al. [30] and Wan et al. [31] were from DFT simulations which did not involve any solvents. In this work, we utilized MD simulations to create a dataset for  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  of g-C<sub>3</sub>N<sub>4</sub> in 49 solvents (water and 48 organic solvents), totaling 31  $\mu\text{s}$  of simulation time. ML methods are then employed to find the correlation between  $\Delta G_{\text{exf}}$  (and  $\Delta G_{\text{sol}}$ ) and descriptors that correspond to physical properties of the solvent. This procedure allows us to identify the most significant descriptors and highlight the guiding principles in the design of solvents for LPE of g-C<sub>3</sub>N<sub>4</sub>. By implementing the most accurate ML model on a new dataset, several solvents are proposed to be potentially effective candidates, including one that has not been attempted in the LPE of 2D material. In tandem with these efforts, experimental dispersibility tests are conducted on four solvents to validate the predictions of our ML model.

In solving the critical bottleneck in the industrial-scale production of high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets, our approach enables significant practical advances in this field. It not only streamlines the process of solvent selection for LPE, but also provides a foundation for more efficient

production of g-C<sub>3</sub>N<sub>4</sub> for a variety of applications, showcasing the potential of machine learning in accelerating materials design and production. The experimental results largely corroborate the model predictions, thereby demonstrating the practical utility of our approach in guiding solvent selection for LPE.

## **4.2. Methods**

### **4.2.1. g-C<sub>3</sub>N<sub>4</sub> and solvents**

g-C<sub>3</sub>N<sub>4</sub> is composed of heptazine or triazine rings as its basic units [32,33]. The heptazine moiety is more common and therefore used in this study [34]. In agreement with previous studies [35,36], an equilateral triangular sheet was chosen as the molecular model for g-C<sub>3</sub>N<sub>4</sub> nanosheet (Figure 4.1). This structure has 21 heptazine units and a side length of 4.5 nm. Nanosheets larger than this size were found to not alter the  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  values [21]. The PolyParGen server [37] was used to generate the parameters compatible with the OPLS-AA force field [38], along with the Charge Model 5 (with a scaling factor of 1.20) [39] for the PACs. For the liquid phase, it is vital to choose solvents that exhibit a wide range of structural and physical properties. In this study, solvents were selected based on two main criteria: (i) extensively used in the literature on LPE of 2D layered materials [40]; and (ii) covering a large range of properties such as density, surface tension, heat of vaporization, dielectric constant, compressibility, thermal expansion coefficient, and molecular weight. Chemical structures of the 49 solvents used in this work are shown in the Appendix B, section B1. For more details on the OPLS-AA force field parameters of the solvents, refer to Coleman et al. [41].



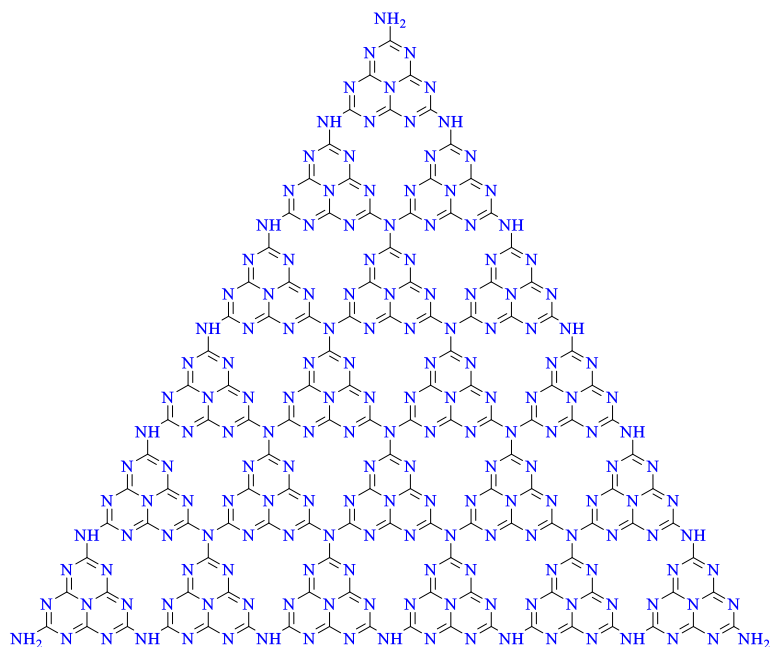


Figure 4.1. Molecular structure of the g-C<sub>3</sub>N<sub>4</sub> nanosheet used for the simulations in this study. The nanosheet contains 21 units of heptazine rings and is saturated with H atoms on the periphery.

#### 4.2.2. Dataset

The objective of the ML modeling is to predict  $\Delta G_{\text{exf}}$  in the LPE of g-C<sub>3</sub>N<sub>4</sub> nanosheets as well as  $\Delta G_{\text{sol}}$ ; these two quantities are considered as the target properties. A significant step in the development of a ML model is the design of appropriate descriptors, which in the context of computational chemistry are mathematical representations of the chemical information of a molecule or properties of a material [42]. In the literature, there are more than 5000 descriptors generated from various methods [42]. Among them, we selected a subset that is potentially relevant to  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ , based on insights from our previous work [21]. These descriptors are categorized into two types. The first type is at the macroscopic level, defined for the bulk solvent. Eight descriptors were considered in this category: heat of vaporization ( $\Delta H_{\text{vap}}$ ), heat capacity under constant pressure ( $C_p$ ), dielectric constant ( $\epsilon$ ), surface tension ( $\gamma$ ), isothermal compressibility ( $\kappa_T$ ), thermal expansion coefficient ( $\alpha_p$ ), density ( $\rho$ ), and finally  $\Delta G_{\text{sol}}$  when  $\Delta G_{\text{exf}}$  is considered as the target property in the ML modeling.

The second type of descriptors are based on the structure of a solvent molecule. They include  $M_w$ , count of hydrogen bond donors (DHBC), count of hydrogen bond acceptors (AHBC),

heavy atom count (HAC), and aromaticity (Arm). In a solvent molecule, DHBC and AHBC respectively count the total number of atoms that are capable of serving as hydrogen donors and acceptors, while HAC is the number of non-hydrogen atoms. Arm = 1 for a solvent that contains an aromatic hydrocarbon; otherwise Arm = 0. Also, Jämbeck et al. concluded that PACs can significantly affect the value of  $\Delta G_{\text{sol}}$  [43]. Therefore, two additional descriptors were introduced: the mean absolute value of PAC of heavy (non-hydrogen) atoms,  $C_{\text{mean}}$ ; and the maximum absolute value of PAC of heavy atoms,  $C_{\text{max}}$ . Mathematically,  $C_{\text{mean}}$  and  $C_{\text{max}}$  are calculated by

$$C_{\text{mean}} = \frac{\sum_{i=1}^N |C_i|}{N} \quad (4.1)$$

$$C_{\text{max}} = \max(|C_1|, |C_2|, \dots, |C_N|) \quad (4.2)$$

where  $C_i$  are the PAC of the  $i^{\text{th}}$  heavy atoms in a solvent molecule, and  $N$  is the total number of heavy atoms. Here, the absolute values of PACs were considered because both positive and negative PACs in a solvent molecule signify its degree of polarity. In fact, carbon and nitrogen atoms in g-C<sub>3</sub>N<sub>4</sub> nanosheets have positive and negative PACs, respectively, and hence can interact with partially charged solvent atoms.

While type 2 descriptors are defined based on the chemical structure of a solvent molecule and hence can be directly calculated, determination of type 1 descriptors needs more consideration. These macroscopic descriptors can be obtained from both experiments and simulations, and there can be some minor difference between the values from the two approaches [41]. Because the objective of this work is to find a relationship between the descriptors and target properties, and the latter ( $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ ) were determined from MD simulations, to be consistent type 1 descriptors were also evaluated from simulations (details in section 4.2.4.2). It is expected that once a well-performed model is established, it can be used to predict  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  from experimental values of the descriptors.

Our initial dataset consisted of the target properties and both types of descriptors for 49 solvents (water and 48 organic solvents). Using the Tukey's method [44], water was identified as the only outlier among the 49 solvents (see details in Appendix B, section B2). It was therefore eliminated from subsequent ML modeling.

### 4.2.3. ML framework and models

The ML framework for predicting  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  is shown in Figure 4.2. MD simulations were first performed to prepare the dataset by calculating the target properties ( $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ ) as well as type 1 descriptors. The second step involved training, tuning, validation and selection of ML models. Six different supervised ML algorithms were used, namely extra-trees regressor (ETR) [45], random forest regression (RFR) [45], gradient boosting regression (GBR) [46], AdaBoost regressor (ABR) [47], Bayesian ridge regression (BRR) [48], and linear regression (LR). These algorithms were chosen based on an initial screening of 50 ML algorithms, which identified ETR, RFR, GBR and ABR as the best-performing nonlinear algorithms as well as BRR and LR as the best-performing linear algorithms. A brief description of these models is given in Chapter 2, section 2.5. Each ML method has hyper-parameters that must be specified in order to customize the model to the dataset. In addition to general heuristics or rules of thumb, a better approach for configuring the hyper-parameters is to objectively test different hyper-parameter values and determine a subset that results in best model performance on a given dataset. Such hyper-parameter optimization or tuning (called “grid search”) was employed in this work, by defining a grid of hyper-parameter values as the search space and evaluating the model performance at every grid point. The entire dataset was used as training data for hyper-parameter tuning.

After hyper-parameter tuning, each ML model was validated with 10-fold cross validation (CV). The entire dataset was randomly divided into 10 subsets or folds, with 9 used for training and 1 for validation. The process was repeated 10 times by alternating the validation fold. This allows the model to be evaluated using the full dataset via resampling, while maximizing the total number of points used for testing (beneficial for reducing overfitting) [49]. Two indices were used to gauge the performance of the ML models: squared correlation coefficient ( $R^2$  score) and root mean square error (RMSE).

The best performing model was adopted for subsequent descriptor selection. Several methods can be applied to reduce the number of descriptors and select the ones that have strongest correlation with the target properties, such as forward selection [50], backward elimination [50], and all subset model (ASM) [51]. If the number of data and descriptors is not very large (as in our case), it is appropriate to use ASM which generates all possible combinations of descriptors (from a single descriptor to the combination of all descriptors). While guaranteeing that the best subset of descriptors is found, ASM can be very demanding computationally. For example, for  $n$

descriptors  $2^{n-1}$  combinations need to be tested. This means for the prediction of  $\Delta G_{\text{exf}}$  (with 15 descriptors) there are 16384 possible subsets. To make the computation more realistic, an upper bound (= 6 in this study) was specified for the number of descriptors in the subsets, a strategy often used in the literature [52].

After descriptor selection was completed, all ML models were trained again with the reduced number of descriptors and 10-fold CV. The model with the best performance in predicting the target properties was identified. This model was finalized by fine tuning of its parameters using 67% of the total dataset for training and the remaining 33% for testing. This final model was then used to predict  $\Delta G_{\text{exf}}$  for more than 100 solvents that did not exist in our dataset. The ML modeling was conducted using the open-source code Scikit-learn [53] package in the Python 3.7 environment.

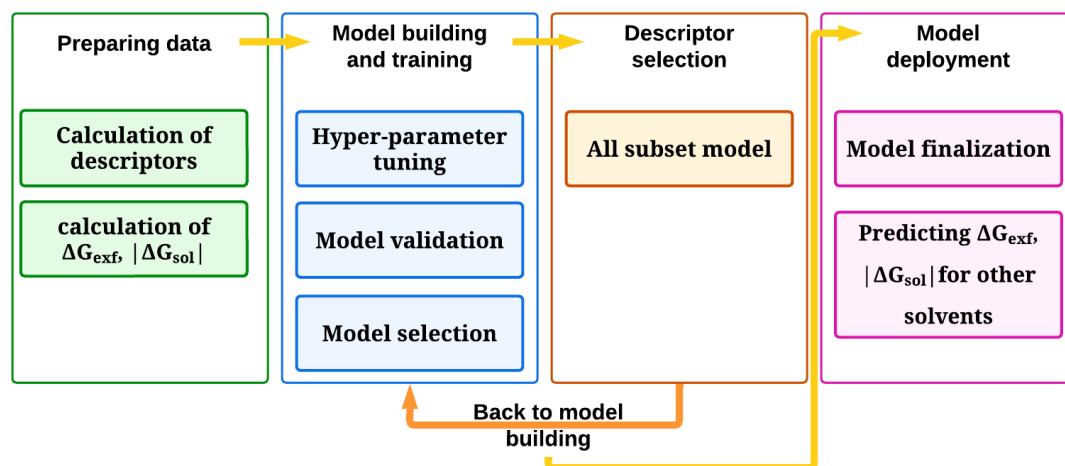


Figure 4.2. ML framework employed in this work. The flowchart outlines the systematic approach taken for dataset preparation, model training and validation, descriptor selection to reduce overfitting, and model finalization.

#### 4.2.4. MD details

MD simulations performed in this work are summarized in Table 4.1.

##### 4.2.4.1. Calculation of target properties

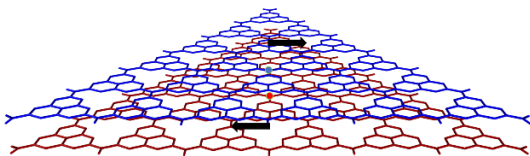
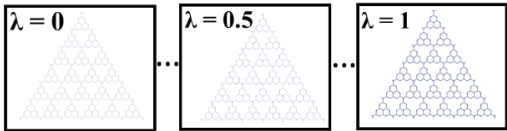
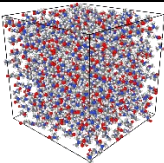
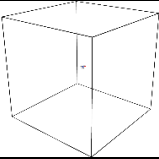
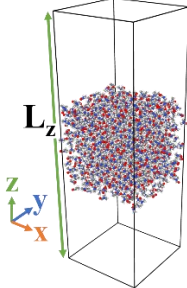
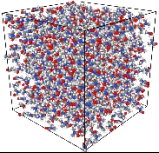
$\Delta G_{\text{exf}}$  was evaluated by calculating the PMF as two initially stacked nanosheets were separated along a defined reaction coordinate. Based on our previous work [21], the most probably

path for LPE of g-C<sub>3</sub>N<sub>4</sub> is via “shearing”, i.e., as one sheet slides over the other while maintaining a parallel configuration. Therefore, the reaction coordinate was chosen to be the distance  $D$  between the COMs of the two nanosheets and varied (from 0.3 nm to 4.0 nm) by moving one nanosheet laterally. A presentation of the reaction coordinate is shown in the first row of Table 4.1 with black arrows. Umbrella sampling [54] simulations were performed and the WHAM was used to calculate the PMF [55]. Specifically, for each solvent a total of 50 US windows were used and for each window, a harmonic biasing potential (with force constant 5000 kJ/mol·nm<sup>2</sup>) was applied to restrain the reaction coordinate. The PMF was set to zero at the smallest value of the reaction coordinate. As an example, the histograms obtained from US simulations in chloroform and the associated PMF curve are shown in Figure 4.3a and b, respectively. The PMF curve exhibits a global minimum at small  $D$  and a plateau when  $D$  is sufficiently large. The plateauing of the PMF curve indicates diminishing interaction between the two sheets when their lateral separation is larger than 4 nm. Within this separation, the interaction between the sheets is attractive, suggesting that spontaneous aggregation will occur when the two sheets are close. The difference between the minimum and plateau values of the PMF is defined as  $\Delta G_{\text{exf}}$  (see Figure 4.3b). Figure 4.3b shows that  $\Delta G_{\text{exf}}$  is lowered by a solvent, compared with the situation where two sheets are separated laterally in vacuum. Nevertheless, the attractive interaction shown by the PMF still requires external forces (e.g., via sonication or high-shear mixing) to disperse the sheets during LPE. When the concentration of the sheets is below the critical aggregation concentration (CAC) [56], LPE can produce a stable dispersion due to the large distance and therefore vanishing attraction between the fully dispersed sheets.

BAR approach [57] was used to calculate  $\Delta G_{\text{sol}}$  of a single nanosheet. Twenty-one states were defined, with the first state corresponding to no interaction between the g-C<sub>3</sub>N<sub>4</sub> nanosheet and the solvent. The vdW and electrostatic interactions were then gradually turned on to reach the final stage where the sheet and the solvent were fully interacting.  $\Delta G_{\text{sol}}$  was calculated based on the data from the twenty-one states. Figure 4.3c shows the change in  $\Delta G_{\text{sol}}$  for the process of turning on vdW and electrostatic interactions. It is worth noting that although both  $\Delta G_{\text{sol}}$  and  $\Delta G_{\text{exf}}$  serve as metrics for evaluating the dispersibility of the nanosheets in a solvent, they are not interchangeable. Specifically,  $\Delta G_{\text{sol}}$  represents an intrinsic property characterizing the interaction of a single sheet with the solvent. In contrast,  $\Delta G_{\text{exf}}$  in our study is derived from a PMF calculation and is contingent upon the chosen reaction coordinate. As evidenced by our previous work [21],

the value of  $\Delta G_{\text{exf}}$  can vary depending on the specific reaction coordinate used. To be precise,  $\Delta G_{\text{exf}}$  discussed in this study should be understood as the free energy required for exfoliating two sheets via shear.

Table 4.1. Summary of MD simulations performed in this study.

Simulation	Schematic	Setup
$\Delta G_{\text{exf}}$ calculation		<ul style="list-style-type: none"> <li>Two g-C<sub>3</sub>N<sub>4</sub> sheets in solvent</li> <li>US along reaction coordinate shown</li> <li>NPT ensemble, initial box size 9*10*9 nm<sup>3</sup></li> <li>50 US windows, 10 ns production run for each window; simulation time for 49 solvents = 24.5 <math>\mu</math>s</li> </ul>
$\Delta G_{\text{sol}}$ calculation		<ul style="list-style-type: none"> <li>One g-C<sub>3</sub>N<sub>4</sub> sheets in solvent</li> <li>BAR method</li> <li>NPT ensemble, initial box size 7*7*4 nm<sup>3</sup></li> <li>21 simulated windows, 1 ns production run for each window; simulation time for 49 solvents = 1 <math>\mu</math>s</li> </ul>
Liquid phase for solvent		<ul style="list-style-type: none"> <li>600 to 4000 solvent molecules</li> <li>NPT ensemble, initial box size 5*5*5 nm<sup>3</sup></li> <li>10 ns production run; simulation time for 49 solvents = 0.5 <math>\mu</math>s</li> </ul>
Gas phase for solvent		<ul style="list-style-type: none"> <li>One solvent molecule</li> <li>NVT ensemble, box size: 5*5*5 nm<sup>3</sup></li> <li>100 ns production run; simulation time for 49 solvents = 4.9 <math>\mu</math>s</li> </ul>
Surface tension calculation for solvent		<ul style="list-style-type: none"> <li>600 to 4000 solvent molecules</li> <li>NVT ensemble, box size 5*5*15 nm<sup>3</sup></li> <li>10 ns production run; simulation time for 49 solvents = 0.5 <math>\mu</math>s</li> </ul>
DoS		<ul style="list-style-type: none"> <li>600 to 4000 solvent molecules</li> <li>NVT ensemble, box size 5*5*5 nm<sup>3</sup></li> <li>100 ps production run; simulation time for 49 solvents = 0.005 <math>\mu</math>s</li> </ul>

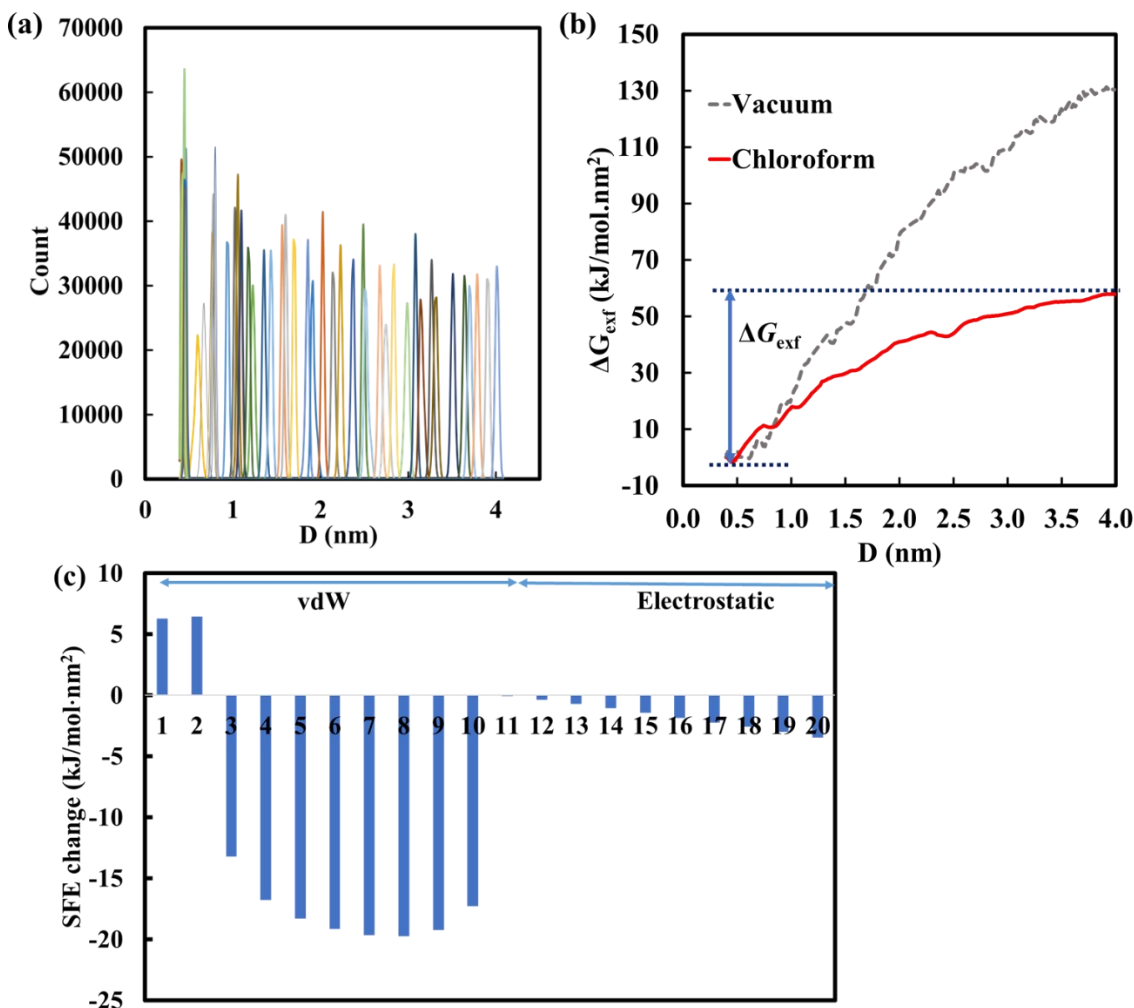


Figure 4.3. Illustration of computational methodology and free energy definitions. (a) Histograms from US simulations for two  $g\text{-C}_3\text{N}_4$  sheets in chloroform. (b) PMF curve and definition of  $\Delta G_{\text{exf}}$  for two  $g\text{-C}_3\text{N}_4$  sheets in chloroform. Corresponding PMF curve in vacuum is shown as a comparison to demonstrate lowering of  $\Delta G_{\text{exf}}$  by chloroform. (c) The change in solvation free energy (SFE) as the vdW and electrostatic interactions are turned on, via 21 steps, between a  $g\text{-C}_3\text{N}_4$  sheet and chloroform. The sum of the changes is the total solvation free energy for this solvent and its absolute value is denoted by  $\Delta G_{\text{sol}}$  in this work.

#### 4.2.4.2. Calculation of type 1 descriptors

Four additional simulations were performed for each solvent (Table 4.1) to evaluate type 1 descriptors related to its bulk properties. Such calculations follow the work of Coleman et al.,



which has been implemented in GROMACS [58]. For completeness, a brief description of the computations is provided below. For each solvent, a 10.0 ns NPT simulation (liquid phase simulation in Table 4.1) was performed to calculate the density ( $\rho$ )

$$\rho = \frac{M}{\langle V \rangle} \quad (4.3)$$

where  $M$  is the mass of the system and  $\langle V \rangle$  is the average volume of the simulation box. Static dielectric constant ( $\epsilon$ ) was computed based on the fluctuation in the magnitude of dipole moment ( $M_D$ ) obtained from the liquid phase simulation [59]:

$$\epsilon = 1 + \frac{4\pi}{3} \left( \frac{\langle M_D^2 \rangle - \langle M_D \rangle^2}{\langle V \rangle k_B T} \right) \quad (4.4)$$

where  $k_B$  is the Boltzmann constant and  $T$  is the temperature. Thermal expansion coefficient ( $\alpha_P$ ) and isothermal compressibility ( $\kappa_T$ ) were also computed from fluctuations in the liquid phase simulation, according to [60]

$$\langle \delta V \delta H \rangle = k_B T^2 \langle V \rangle \alpha_P \quad (4.5)$$

$$\langle \delta V^2 \rangle = k_B T \langle V \rangle \kappa_T \quad (4.6)$$

where  $\delta V$ ,  $\delta H$  and  $\delta V^2$  are respectively the standard deviation in volume, standard deviation in enthalpy and variance in volume. The enthalpy of vaporization ( $\Delta H_{vap}$ ) was evaluated from

$$\Delta H_{vap} = (\langle U_g \rangle + k_B T) - \langle U_l \rangle \quad (4.7)$$

where  $U_l$  is the potential energy of one mole of the solvent in the liquid phase evaluated from the liquid phase simulation in Table 4.1, and  $U_g$  is the potential energy of one mole of the solvent in the gas phase evaluated from the gas phase simulation in Table 4.1. The gas phase is modeled as an ideal gas with no intermolecular interaction, and the molar volume of the liquid phase is assumed negligible compared with that of the gas phase. To determine the surface tension ( $\gamma$ ), a simulation box with two liquid–vacuum interfaces was generated (Table 4.1) and  $\gamma$  was calculated from

$$\gamma(t) = \frac{L_z}{2} \left\langle P_z(t) - \frac{P_x(t) + P_y(t)}{2} \right\rangle \quad (4.8)$$

where  $P_i$  is the pressure component in direction  $i$  and  $L_z$  is the length of the box in  $z$  direction (refer to Table 4.1 for the xyz coordinates). For the calculation of constant pressure heat capacity ( $C_p$ ), the two-phase thermodynamics method introduced by the Goddard group [61] was applied. This method treats a liquid as a substance between a solid and an ideal gas, which enables the calculation of thermodynamic properties based on the density of states [41]

$$C_p = k_B \int_0^\infty [DoS_{gas}(v)W_{gas}(v) + DoS_{solid}(v)W_{solid}(v)]dv + VT \frac{\alpha_p^2}{\kappa_T} \quad (4.9)$$

Here  $DoS_{gas}$  and  $DoS_{solid}$  are respectively the density of states in a gas and a solid state, and  $W_{gas}(v)$  and  $W_{solid}(v)$  are the corresponding weighting factors [62]. Both  $DoS_{gas}$  and  $DoS_{solid}$  can be obtained from the  $DoS$  simulation in Table 4.1, which calculates  $DoS$  of the liquid as a function of frequency  $v$  from the Fourier transform of the mass-weighted velocity autocorrelation function.

#### 4.2.4.3. Simulation parameters

MD simulations were carried out using the GROMACS 2021.2 package [58]. Before each production run, the steepest descent algorithm was used for energy minimization, followed by a short (200 ps) equilibration in the NPT ensemble with a timestep of 2 fs. Berendsen thermostat and barostat were used to control the temperature and pressure at 300 K and 1 bar with a coupling time constant of 0.2 ps and 5.0 ps respectively. In the production run for all simulations in Table 4.1 (except  $DoS$ ), a timestep of 2 fs was applied and bond lengths involving a hydrogen atom were constrained using the LINCS algorithm. A cutoff of 1.2 nm was used for vdW (Lennard-Jones) and short-range electrostatic interactions, while the PME algorithm was employed for computing long-range electrostatic interactions. Periodic boundary condition was employed in all directions. For production runs in NVT ensemble (gas phase simulation and surface tension calculation in Table 4.1), the temperature was maintained at 300 K via the velocity-rescale algorithm with a time constant of 0.1 ps. For production runs in NPT ensemble ( $\Delta G_{\text{exf}}$ ,  $\Delta G_{\text{sol}}$  and liquid phase simulation in Table 4.1), the same thermostat was adopted while C-rescale scheme with a coupling constant

of 1.0 ps was used to achieve an isotropic pressure of 1 bar. In the production run for the *DoS* simulations, stricter energy conservation parameters were used, including a neighbor list buffer of 0.3 nm, combined with a switched Lennard-Jones and short-range electrostatics term (for 1.0–1.1 nm). See reference [63] for more explanations on the simulation parameters used.

## 4.3. Results

### 4.3.1. Quality of dataset

It is of vital importance that our dataset of solvents (with water removed as the outlier) is representative of common organic solvents studied in the literature. To examine this, the probability distribution of  $\rho$ ,  $\epsilon$  and  $\gamma$  for the organic solvents in our dataset are compared to an organic solvent dataset cured by Coleman et al. [41] which includes 146 common organic solvents (Figure 4.4). For both datasets, these properties are calculated from MD simulations, which allows for a fair and direct comparison. The good agreement in Figure 4.4 shows that our database is a sufficient sample of the population.

Table 4.2 summarizes the statistics of all descriptors for the solvents in our dataset, including their mean, standard deviation (std), minimum (min), the first quartile ( $Q_1$ ), the second quartile ( $Q_2$ ), the third quartile ( $Q_3$ ), and maximum (max). More information about the dataset and the histograms of all descriptors and target properties can be found in Appendix B, section B1.

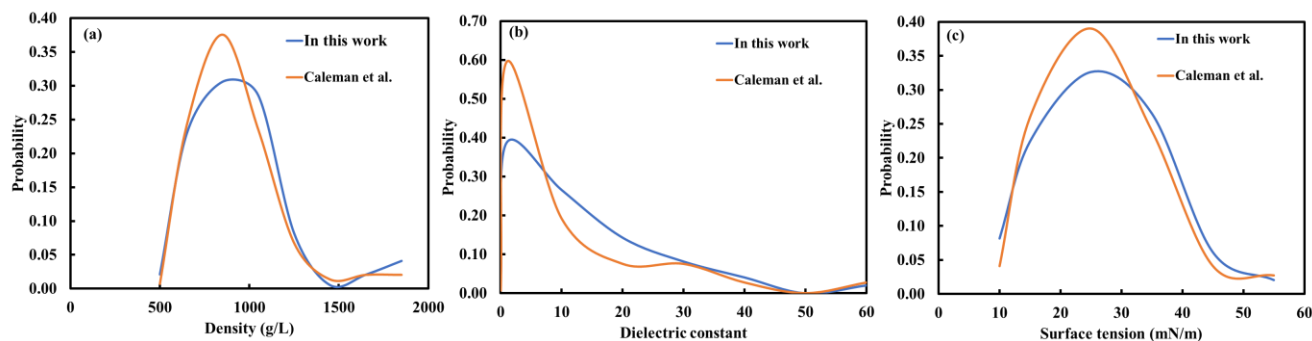


Figure 4.4. Comparison of solvent properties in this work and those in Coleman et al. [41] (containing a larger dataset). The probability distributions are compared for (a) density, (b) dielectric constant and (c) surface tension of the organic solvents in both works. The good agreement confirms that the organic solvents in our dataset constitute a sufficient sample.

Table 4.2. Statistics of descriptors and target properties used for ML modeling.

	Quantity	mean	std	min	25%	50%	75%	max
Macroscopic descriptors	$\Delta H_{vap}$ (kJ/mol)	46.43	14.83	23.36	34.02	46.68	55.17	89.37
	$C_p$ (J/mol·K)	157.59	59.09	70.07	118.89	149.54	176.73	333.23
	$\varepsilon$	21.21	16.37	1.01	6.70	15.20	35.22	59.19
	$\gamma$ (mN/m)	30.32	10.76	12.03	21.46	31.47	37.39	57.20
	$\kappa_T$ (1/GPa)	0.87	0.51	0.22	0.48	0.63	1.16	2.22
	$\alpha_P$ (10 <sup>-3</sup> /K)	1.25	0.41	0.59	0.98	1.15	1.50	2.12
	$\rho$ (g/L)	1058.92	326.03	584.90	863.03	1024.80	1108.07	2475.30
Molecular descriptors	$M_W$ (g/mol)	100.26	43.13	32.04	60.84	92.12	127.55	202.09
	DHBC	0.40	0.71	0	0	0	1	3
	AHBC	1.13	0.91	0	1	1	1	4
	HAC	6.23	2.63	2	4	6	8	13
	Arm	0.38	0.49	0	0	0	1	1
	$C_{mean}$	0.18	0.10	0.04	0.11	0.15	0.2025	0.48
	$C_{max}$	0.49	0.33	0.04	0.22	0.47	0.68	2
Target properties	$\Delta G_{exf}$ (kJ/mol·nm <sup>2</sup> )	57.82	8.72	44.37	50.68	55.72	62.63	78.68
	$\Delta G_{sol}$ (kJ/mol·nm <sup>2</sup> )	164.98	19.58	114.57	150.80	165.54	176.27	208.28

### 4.3.2. Model comparison

Six different ML methods are applied on the dataset to identify the one that predicts  $\Delta G_{exf}$  and  $\Delta G_{sol}$  with the highest accuracy. All the descriptors are included in the first round of modeling. Figure 4.5a shows the predicted (from ML models) vs. measured (from MD) values of  $\Delta G_{exf}$  after 10-fold CV.  $R^2$  score and RMSE are also displayed, where ETR produces the highest  $R^2 = 63.14$  and the lowest RMSE = 5.51 kJ/mol·nm<sup>2</sup>, followed by ABR, RFR, GBR, BBR, and LR. The two linear models (BRR and LR) have the lowest performance suggesting complex non-linear dependence of  $\Delta G_{exf}$  and  $\Delta G_{sol}$  on the descriptors. Figure 4.5b exhibits that the best prediction of  $\Delta G_{sol}$  is also achieved by ETR with  $R^2 = 80.79$  and RMSE = 8.51 kJ/mol·nm<sup>2</sup> and the two linear models are again inferior in comparison to the nonlinear models.

A closer examination of Figure 4.5 is conducted to understanding the origin of discrepancy between predicted and measured values. In Figure 4.5a, the predicted  $\Delta G_{\text{exf}}$  is closer to the measured value when  $\Delta G_{\text{exf}} < 67 \text{ kJ/mol}\cdot\text{nm}^2$  (except GBR which shows better agreement for larger  $\Delta G_{\text{exf}}$ ). In addition, all the nonlinear models overestimate  $\Delta G_{\text{exf}}$  for two solvents (dibromomethane and dichlorofluoromethane) highlighted by orange circles in Figure 4.5a. A similar behavior is observed for  $\Delta G_{\text{sol}}$ , where all models produce poor prediction for the two solvents with the lowest and highest  $\Delta G_{\text{sol}}$  (propane,1,2,3-triol and triethyl phosphate, respectively, highlighted by orange circles in Figure 4.5b). The inaccuracy associated with these seemingly outliers (dibromomethane and dichlorofluoromethane for  $\Delta G_{\text{exf}}$ , propane,1,2,3-triol and triethyl phosphate for  $\Delta G_{\text{sol}}$ ) can be attributed to the lack of solvents around these solvents. For example, according to Table 4.2, more than 75% of the solvents have  $\Delta G_{\text{exf}} < 67 \text{ kJ/mol}\cdot\text{nm}^2$ , which results in lower accuracy of predicting  $\Delta G_{\text{exf}}$  when its value is above  $67 \text{ kJ/mol}\cdot\text{nm}^2$ .

Reduced accuracy in the prediction of  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  due to lack of solvents can be a sign of overfitting. Moreover, the number of descriptors in this study (15 for  $\Delta G_{\text{exf}}$  and 14 for  $\Delta G_{\text{sol}}$ ) is relatively high compared to the number of solvents (48 used for ML modeling), which can lead to overfitting [64]. Potential overfitting is further investigated with ETR, the best performing model identified from Figure 4.5. Bootstrap method is implemented to break down the contributions of bias and variance to the predicting error [65],

$$\text{bias}^2 = \frac{1}{n_{\text{test}}} \sum_{l=1}^{n_{\text{test}}} (\bar{f}(X_l) - y_l)^2 \quad (4.10)$$

$$\text{variance} = \frac{1}{n_{\text{test}}} \sum_{l=1}^{n_{\text{test}}} \frac{1}{B} \sum_{b=1}^B (f(X_l; D^b) - \bar{f}(X_l))^2 \quad (4.11)$$

where  $n_{\text{test}}$  is the number of testing data,  $X_l$  contains descriptors for the  $l^{\text{th}}$  testing data, and  $y_l$  is the measured target property for the  $l^{\text{th}}$  testing data.  $B$  is the number of training sets (denoted by  $D^b$ ,  $b = 1, 2, \dots, B$ ) sampled from the original training data,  $f(X_l; D^b)$  is the target property predicted for  $X_l$  using the model trained from  $D^b$ .  $\bar{f}(X_l)$  is the average of  $f(X_l; D^b)$  over different training sets  $D^b$ . For  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  the squared biases are estimated to be  $4.68 \text{ kJ}^2/\text{mol}^2\cdot\text{nm}^4$  and  $19.56 \text{ kJ}^2/\text{mol}^2\cdot\text{nm}^4$  respectively, and the corresponding variances are  $21.33 \text{ kJ}^2/\text{mol}^2\cdot\text{nm}^4$  and  $56.19 \text{ kJ}^2/\text{mol}^2\cdot\text{nm}^4$ . This suggests that for both  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  the variance is 3 to 5 times larger

than the squared bias. A large variance overshadowing squared bias signals the statistical issue of overfitting. Although one way to address this problem is adding more solvents, this approach requires unrealistically high time and resource (each solvent requires over 540 ns of simulation). Alternatively, reducing the number of descriptors can also mitigate the impact of overfitting, especially for small datasets [66]. In the next section, descriptor selection is carried out using the ETR model.

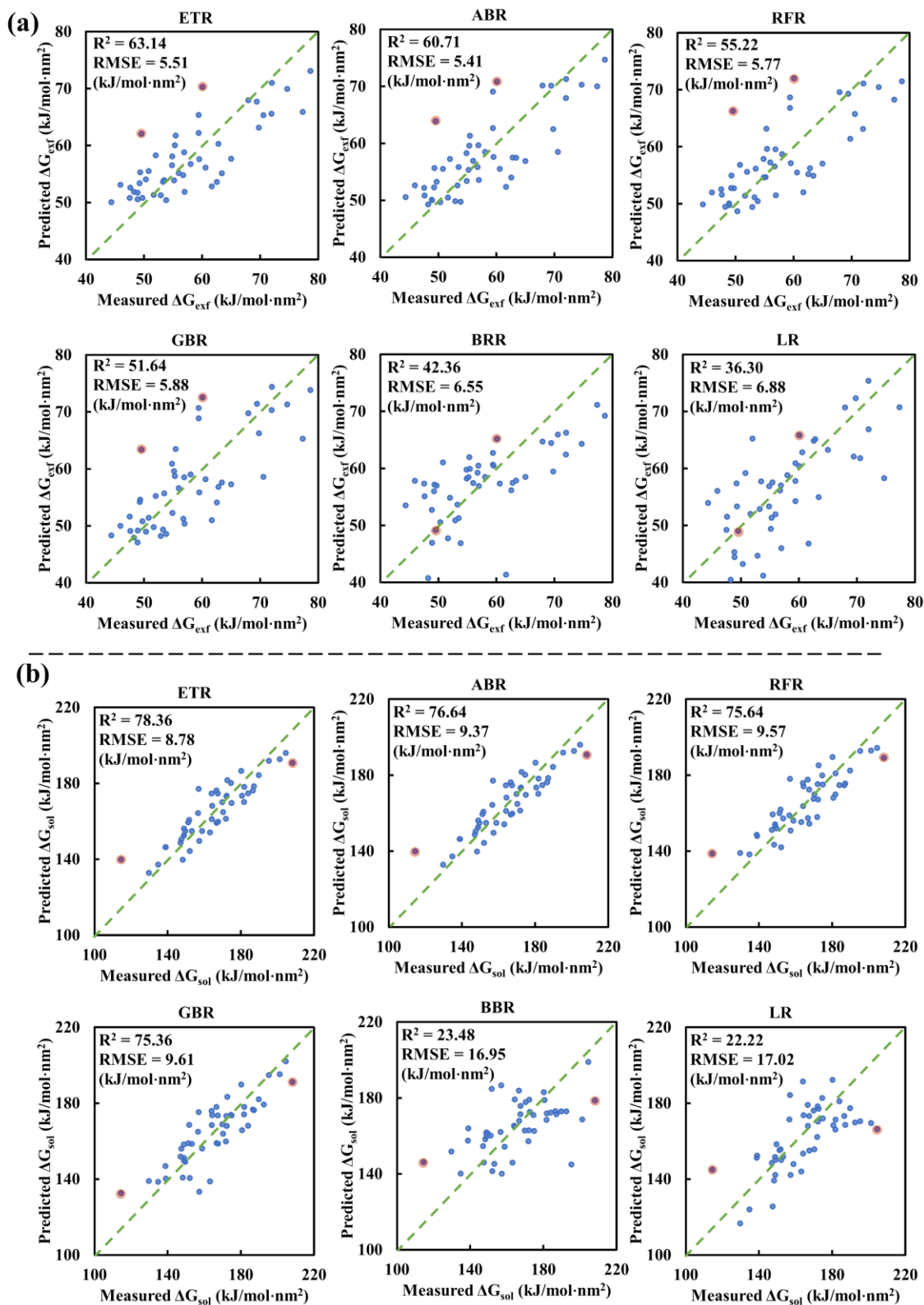


Figure 4.5. Correlation between predicted and measured values of free energies for the six ML models. All descriptors are included in the ML modeling. The scatter plots for (a)  $\Delta G_{\text{exf}}$  and (b)

$\Delta G_{\text{sol}}$  provide a visual assessment of model accuracy, which is quantified by  $R^2$  and RMSE values indicated in each subplot. ETR stands out as the most accurate model for both  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ .

### 4.3.3. Descriptor selection

After implementing the ASM method for ETR, the optimal subsets of descriptors are identified to be:  $\Delta H_{\text{vap}}$ ,  $\varepsilon$ ,  $\gamma$ ,  $\rho$  and  $M_W$  for  $\Delta G_{\text{exf}}$ , and  $\varepsilon$ ,  $\gamma$ ,  $\alpha_p$ ,  $M_W$ ,  $C_{\text{mean}}$  and  $C_{\text{max}}$  for  $\Delta G_{\text{sol}}$ . All the ML models are trained again with these descriptors and their optimal hyper-parameters are listed in Appendix B, section B3. Figure 4.6a shows the improvement in  $R^2$  and RMSE for ETR after descriptor selection. A 21% gain in  $R^2$  and 24% reduction in RMSE are observed in the prediction of  $\Delta G_{\text{exf}}$ . While the improvement is less significant for  $\Delta G_{\text{sol}}$ ,  $R^2$  still increases by 3% and RMSE decreases by 4%. The predicted vs. measured values of  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  after descriptor selection are shown in Appendix B, section B3 for all ML models. Comparison with Figure 4.5 shows improved performance of all ML models with the optimal subset of descriptors. Reducing the number of descriptors to a maximum of six was strategically applied to strike a balance among computational efficiency, model complexity, and prediction accuracy. This not only mitigates overfitting [67] but also facilitates the prediction of  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  for solvents that are out of our database, as fewer descriptors require less amount of calculations.

Some discussion is warranted for why the descriptors in the optimal subsets might play crucial roles in determining  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ . Our previous work [21] has shown that  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  are correlated with the mobility of solvent molecules around the g-C<sub>3</sub>N<sub>4</sub> nanosheet, with lower mobility leading to larger  $\Delta G_{\text{sol}}$  and smaller  $\Delta G_{\text{exf}}$ . Reduced solvent mobility can be caused by high values of  $\rho$  and  $M_W$ . Larger  $\Delta H_{\text{vap}}$  and  $\gamma$  both imply stronger solvent cohesion, which could make the insertion of a nanosheet into the solvent energetically more difficult. On the other hand, large  $\Delta H_{\text{vap}}$  and  $\gamma$  can also work synergistically with large  $\rho$  and  $M_W$  to reduce solvent mobility around the nanosheet. The effect of  $\varepsilon$  is multi-faceted. Larger  $\varepsilon$  corresponds to higher polarity of the solvent molecules, which can lead to stronger solvent cohesion that hinders solvation and exfoliation. Meanwhile, the more polar solvent molecules can exhibit increased affinity towards the partially charged atoms on the g-C<sub>3</sub>N<sub>4</sub> nanosheets, which favors solvation and exfoliation. Higher  $\varepsilon$  also provides stronger screening of electrostatic interactions between different nanosheets. Solvents with lower  $\alpha_p$  values offer a more stable molecular arrangement



around the nanosheets across temperature variations, potentially producing a more favorable  $\Delta G_{\text{sol}}$ . The two descriptors related to PACs ( $C_{\text{mean}}$  and  $C_{\text{max}}$ ) are found to be influential for  $\Delta G_{\text{sol}}$ , consistent with the work of Jämbeck et al. [43] Such impact is rooted in the electrostatic interactions between partially charged atoms on the nanosheets and those on the solvent molecules.

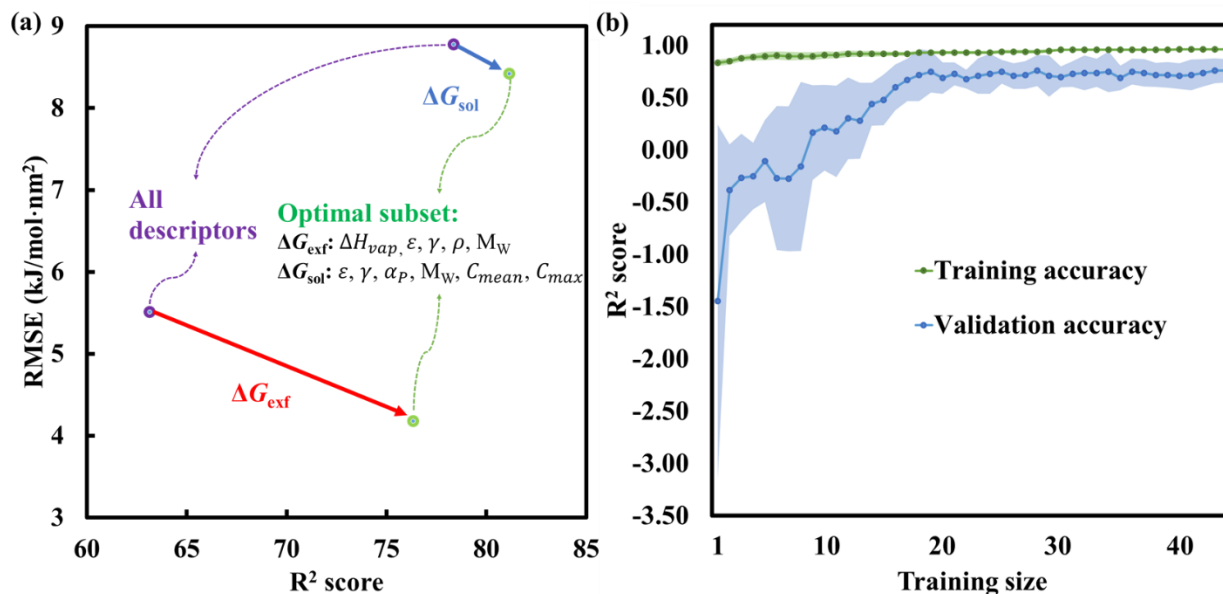


Figure 4.6. Refinement of the ETR model. (a) RMSE and  $R^2$  score with all descriptors and the optimal subset (listed in figure) identified from ASM. The arrows indicate the improvement of prediction accuracy after descriptor selection. (b) Learning curves (with  $\Delta G_{\text{exf}}$  being the target property) using 10-fold CV and  $R^2$  as the accuracy metric. The green and blue shadows show the standard deviation for training and validation accuracy, respectively. Both curves plateau after the training size reaches 20.

#### 4.3.4. Model finalization

It is well-known that the accuracy of an ML model depends on the size of the training set [66]. To study this effect, Figure 4.6b shows the change in  $R^2$  score for the ETR prediction of  $\Delta G_{\text{exf}}$  by varying the proportion of the training set (also known as the learning curve). The optimal subset of descriptors is used, along with 10-fold CV for validation. The size of the training set is varied from 1 to 43, with the remaining data serving as the testing set. The training accuracy curve is generated from the training set, while the validation accuracy curve is from the testing set. Standard

deviation is reported for each curve, to account for the different choices of training/testing set during 10-fold CV.

As illustrated in Figure 4.6b, the training accuracy is insensitive to the number of training data points and quickly reaches 0.90 when as few as five data points were used. A slight increasing trend continues as the training size increases, ultimately reaching 0.96 when more than 30 data points are utilized for training. On the other hand, validation accuracy is found to be highly dependent on the number of training data points, with even negative  $R^2$  score observed when less than nine data points are used for training. As the training size increases, validation accuracy also increases, ultimately reaching a plateau when the training set contains 20 or more data points. This indicates that the accuracy of the model is maintained as long as the training set includes more than 20 data points. However, it is recommended that the training set contains more data points than the testing set when the dataset is small [68]. For generating the final model, 67% (32 data points) of the dataset is used for training and 33% (16 data points) used for testing. The hyper-parameters used for this model remain the same as the ones listed in Appendix B, section B3.

A few more data are provided here to support the robustness of the finalized model. In Figure 4.6a, RMSE for  $\Delta G_{\text{exf}}$  obtained from the 10-fold cross-validation after descriptor selection is 4.18 kJ/mol.nm<sup>2</sup> (data on the bottom right). Using the finalized model with a 67/33 split between the training and testing data, we performed RMSE calculations for five random data splitting. The RMSE values are 4.35, 4.17, 4.24, 4.04 and 4.58 kJ/mol.nm<sup>2</sup> respectively, which are close to the RMSE from Figure 4.6a. We further employed the Bootstrap method [69] for 1000 iterations and the RMSE obtained is 4.23 kJ/mol.nm<sup>2</sup>, again comparable to the result from 10-fold cross-validation. The consistency among these data suggests that the finalized ETR model is well calibrated, with good applicability to be generalized for new, unseen data.

## 4.4. Discussion

The finalized ETR model can now be used to predict  $\Delta G_{\text{exf}}$  for solvents outside of our dataset. For this purpose, the dataset cured by Coleman et al [41] is used, which contains the experimental values for the optimal subset of descriptors ( $\Delta H_{\text{vap}}$ ,  $\epsilon$ ,  $\gamma$ ,  $\rho$  and  $M_W$ ) for 146 organic solvents. Additionally, we included seven organic solvents commonly used in the LPE process, bringing the total number of organic solvents in the dataset to 153. A complete list of the solvents

and their predicted  $\Delta G_{\text{exf}}$  values are given in Appendix B section B4, while the histogram of the predicted  $\Delta G_{\text{exf}}$  is shown in Figure 4.7. This figure demonstrates that most of the solvents have a moderate performance in exfoliating the g-C<sub>3</sub>N<sub>4</sub> sheets, with  $\Delta G_{\text{exf}}$  falling in the range of 50-60 kJ/mol·nm<sup>2</sup>. However, four solvents have a  $\Delta G_{\text{exf}}$  less than 50 kJ/mol·nm<sup>2</sup> and therefore are potentially effective solvents for LPE. They are, in the order of increasing  $\Delta G_{\text{exf}}$ : benzonitrile (PhCN), phenylmethanol (or benzyl alcohol, abbreviated Bn), NMP, and methanesulfonic acid (MSA). Some of these solvents have been previously explored for the LPE of 2D materials, including: PhCN for graphene [70], NMP for graphene [71], MoS<sub>2</sub> [72] and g-C<sub>3</sub>N<sub>4</sub> [73], as well as MSA for graphene [74] and boron nitride nanosheets [75]. However, Bn has never been attempted for the LPE of layered materials and could be a good candidate to consider.

To test the model predictions, four solvents—NMP, MSA, Bn, and Acetonitrile (ACN)—were chosen for experimental validation. They were selected based on the diversity in their chemical structure and properties, commercial availability, as well as range of  $\Delta G_{\text{exf}}$  prediction (low for NMP, MSA, Bn, while high for ACN). Particularly, Bn is a novel candidate identified by the model, while NMP and MSA are established benchmarks. The experimental details are given in Appendix B section B5, and the images demonstrating the dispersibility are shown in Figure 4.8. NMP, Bn and MSA all demonstrate good dispersibility of g-C<sub>3</sub>N<sub>4</sub>, producing a uniform dispersion for concentration as high as 5 mg/mL, which remains undisturbed after an hour. These observations are consistent with the developed model, which predicts small values of  $\Delta G_{\text{exf}}$  for NMP (49.43 kJ/mol·nm<sup>2</sup>), MSA (49.89 kJ/mol·nm<sup>2</sup>) and Bn (49.11 kJ/mol·nm<sup>2</sup>). Conversely, the dispersibility of g-C<sub>3</sub>N<sub>4</sub> in ACN is remarkably poor, also in line with our prediction ( $\Delta G_{\text{exf}} = 67.54$  kJ/mol·nm<sup>2</sup>), and the substance sediments at the bottom. Therefore, our model is shown to successfully predict the dispersibility of g-C<sub>3</sub>N<sub>4</sub> in these solvents.

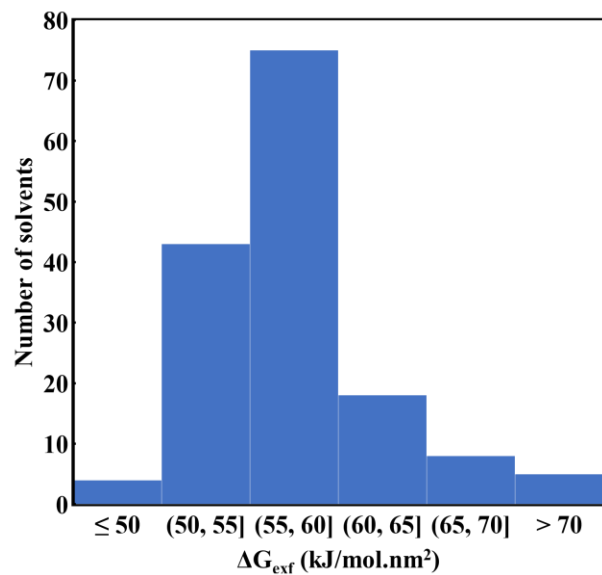


Figure 4.7. Histogram of predicted  $\Delta G_{\text{exf}}$  for an expanded set of organic solvents. The prediction is made using the ETR model with the optimal subset of descriptors, for the 153 organic solvents listed in Appendix B section B4. The distribution shows that most  $\Delta G_{\text{exf}}$  values are in the range of 50-60  $\text{kJ/mol}\cdot\text{nm}^2$ .

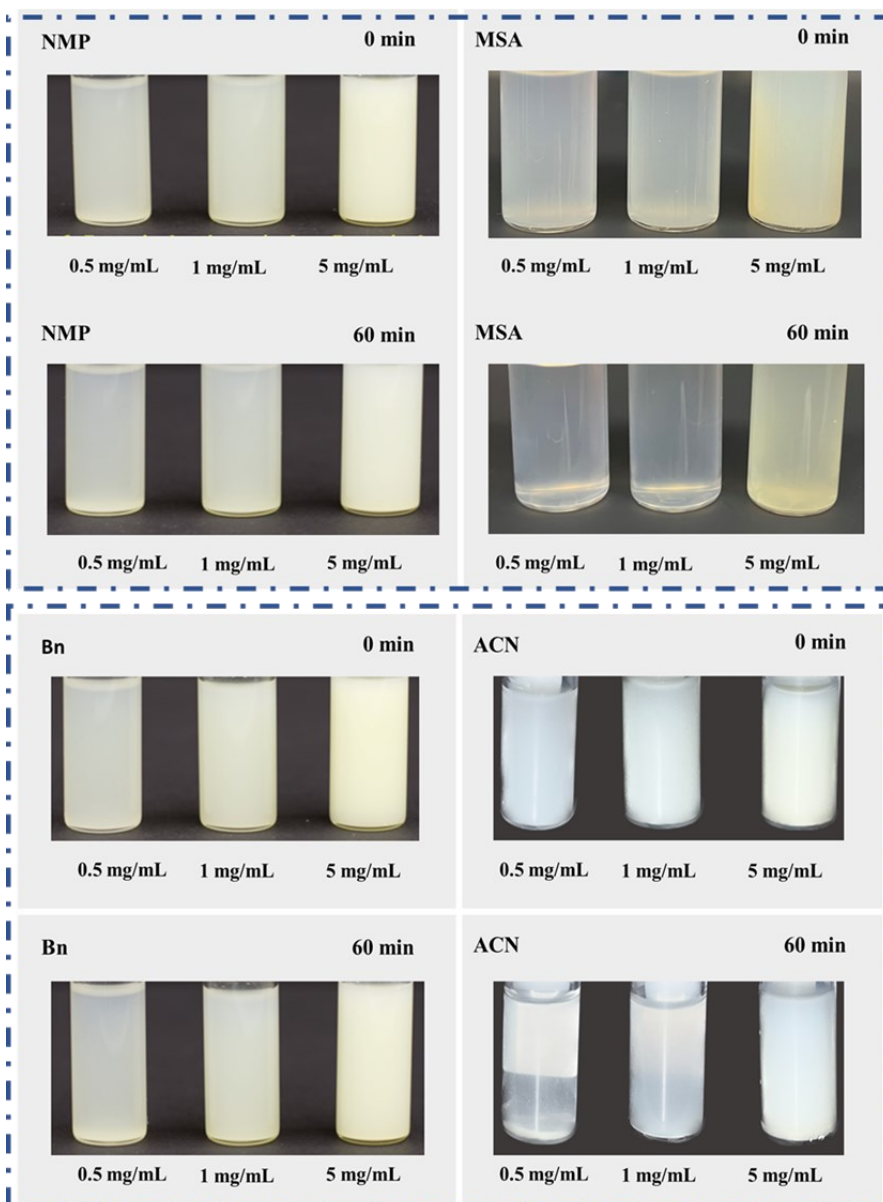


Figure 4.8. Macroscopic images demonstrating the dispersibility of bulk carbon nitride in various solvents. For each solvent, two sets of images are shown which correspond to different time (0 min: immediately after sonication; 60 min: 1 hour after sonication with no disturbance). Mass concentration of g-C<sub>3</sub>N<sub>4</sub> for each test is indicated below the respective image. ACN shows considerably poorer dispersibility compared with other solvents.

Although the LR model showed poorer performance than ETR in Figure 4.5, by introducing the optimal subset of descriptors, the accuracy of LR is also improved. The enhancement in its RMSE and R<sup>2</sup> score due to descriptor selection is illustrated in Appendix B

section B3. One advantage of the LR model is that it provides closed-form predictions for  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$ , and therefore can be a useful tool for initial screening of solvents where relatively low accuracy is needed. The closed-form LR predictions of  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  in terms of the optimal subset of descriptors are

$$\Delta G_{\text{exf}} = -0.35 \Delta H_{\text{vap}} + 0.12 \varepsilon - 0.03 \gamma - 0.01 \rho - 0.02 M_W + 80.56 \quad (4.12)$$

$$\Delta G_{\text{sol}} = 0.04 \varepsilon + 0.16 \gamma - 17.18 \alpha_P + 0.03 M_W + 28.08 C_{\text{mean}} + 27.29 C_{\text{max}} + 159.84 \quad (4.13)$$

Looking at Eqs. (4.12) and (4.13), it appears that  $\Delta G_{\text{exf}}$  has a negative correlation with  $\Delta H_{\text{vap}}$  and  $\gamma$ , correlates positively to  $\varepsilon$  and has negligible dependence on  $\rho$  and  $M_W$ . For  $\Delta G_{\text{sol}}$ , a highly negative correlation is found with  $\alpha_P$  while there is a strong positive correlation with  $C_{\text{mean}}$  and  $C_{\text{max}}$ .  $\varepsilon$ ,  $\gamma$  and  $M_W$  have much smaller impact on  $\Delta G_{\text{sol}}$ . In an earlier work, Coleman et al. [18] proposed that the free energy of exfoliation for graphene only depends on the surface energy difference between graphene and the solvent. On the contrary, Eqs. (4.12) and (4.13) show that  $\gamma$  is neither the only influencing parameter for  $\Delta G_{\text{exf}}$  nor the most impactful one.

One question arises from the LR predictions in Eqs. (4.12) and (4.13). If some of the descriptors (such as  $M_W$ ) has very weak correlation with the target properties, why are they selected during ASM? To answer this, Figure 4.9 shows the pair plots of the target properties measured from MD vs. three descriptors that have the highest coefficients in Eqs. (4.12) and (4.13). For  $\Delta G_{\text{sol}}$ , the pair plot with  $C_{\text{mean}}$  is not shown because it is similar to that with  $C_{\text{max}}$  (Appendix B section B6). Figure 4.9 sheds light on some of the correlations found in LR. Generally, LR fits a line between the target property and a descriptor, and if there is a significant linear relationship (as is the case for  $\Delta H_{\text{vap}}$  in Figure 4.9a and  $C_{\text{max}}$  in Figure 4.9d), a large coefficient will appear in the closed-form. On the contrary, if the relationship is not clear (Figure 4.9c, f) or it is a higher order correlation (Figure 4.9b, e), then LR alone is insufficient to describe the dependence of the target property on the descriptor. Using polynomial regression might help but implementing polynomial regression in our dataset only exacerbates overfitting. In fact, using even a second order polynomial led to negative  $R^2$  (data not presented). As shown in Appendix B section B6,  $M_W$  does not exhibit a clear linear relationship with  $\Delta G_{\text{exf}}$  or  $\Delta G_{\text{sol}}$ . This explains why despite its small coefficients in Eqs. (4.12) and (4.13),  $M_W$  is retained during descriptor selection

using ASM and ETR (a nonlinear ML model). While Eqs. (4.12) and (4.13) are useful for initial solvent screening, care should be exercised not to over-interpret the influence of the descriptors.

Another interesting observation from Figure 4.9 is that solvents with better performance (lower  $\Delta G_{\text{exf}}$  and higher  $\Delta G_{\text{sol}}$ ) tend to be associated with a certain range for each descriptor. Specifically, five solvents that have the lowest  $\Delta G_{\text{exf}}$  and highest  $\Delta G_{\text{sol}}$  are highlighted with red color and their name are listed in Figure 4.9a and d. Clearly, these data points form a cluster in each plot and the clustering is more evident for  $\Delta G_{\text{exf}}$ . This motivated us to reexamine  $\Delta G_{\text{exf}}$  predicted for all 153 organic solvents (Appendix section B4 and Figure 4.7) and the range of descriptor values for achieving small  $\Delta G_{\text{exf}}$ . The range of descriptors that correspond to predicted  $\Delta G_{\text{exf}} < 51 \text{ kJ/mol}\cdot\text{nm}^2$  are shown in Figure 4.10a, which serves as a guideline for initial screening of solvents for LPE of g-C<sub>3</sub>N<sub>4</sub> nanosheets.

The effect of descriptor on the target properties is further examined through the Shapley additive explanations (SHAP) values [76]. SHAP values allocate the contribution from each descriptor towards the model prediction, thereby allowing us to understand the impact of individual descriptors. The relative importance of each descriptor in predicting  $\Delta G_{\text{exf}}$  from the ETR model is shown in Figure 4.10b and the corresponding results for  $\Delta G_{\text{sol}}$  are given in Appendix B section B6. On the horizontal axis, each point represents an individual instance or sample from the dataset. Each instance represents a specific set of descriptor values from the dataset used to generate predictions and assess descriptor contributions. The predicted  $\Delta G_{\text{exf}}$  associated with these instances are shown at the top, below which the descriptors are ordered by their importance from high to low, as determined by the mean absolute SHAP values. A higher SHAP value signifies a stronger effect of the descriptor on the prediction, while a lower value implies a weaker effect. The results reveal that the  $\Delta H_{\text{vap}}$  has the highest influence on the prediction of  $\Delta G_{\text{exf}}$ , followed by  $M_w$ ,  $\varepsilon$ ,  $\gamma$ , and  $\rho$ . Together with Figure 4.10a, these findings can be used to guide solvent selection for LPE of g-C<sub>3</sub>N<sub>4</sub>, especially to identify solvents that exhibit favorable characteristics in terms of heat of vaporization and molecular weight.

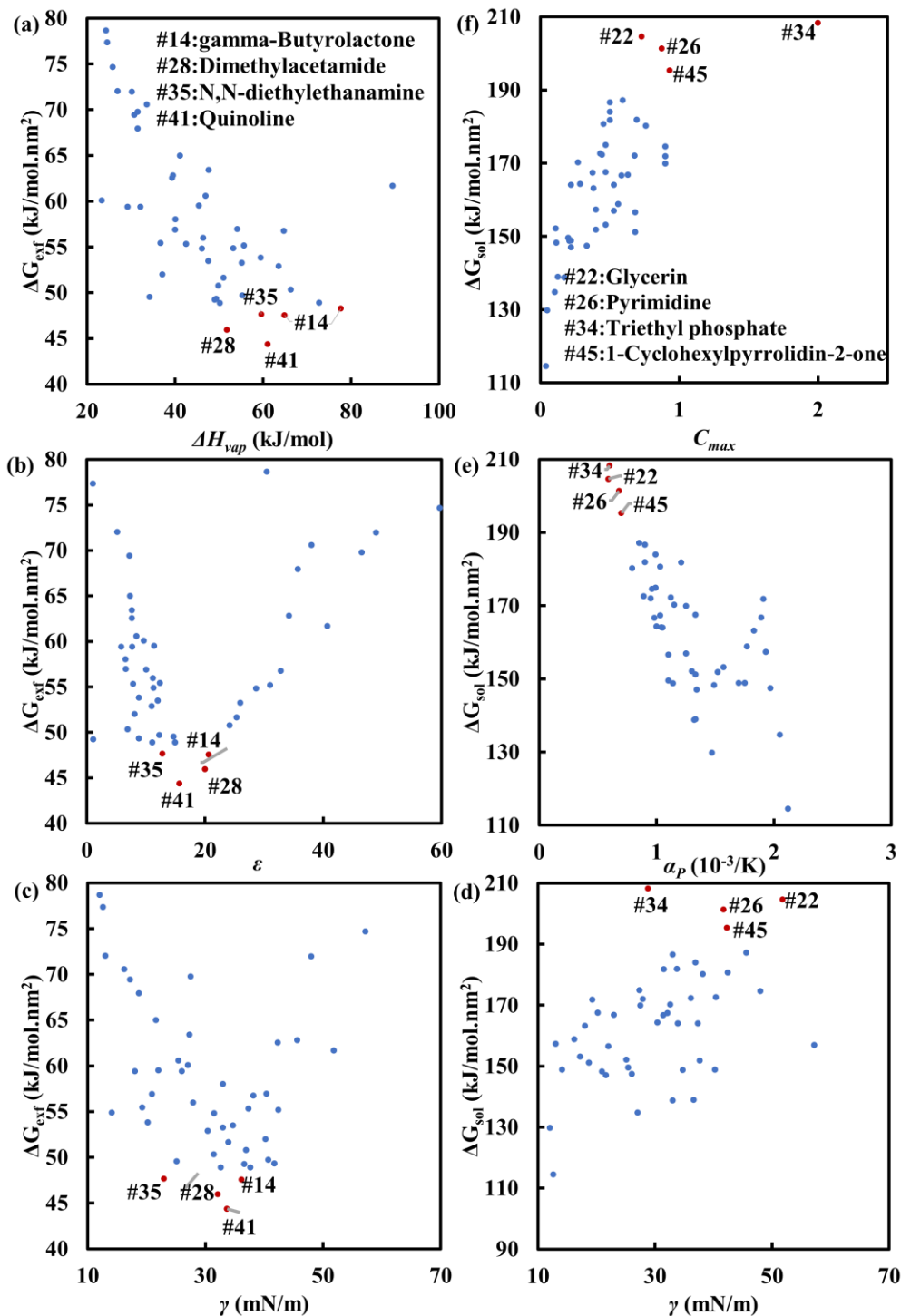


Figure 4.9. Correlation between measured  $\Delta G_{\text{exf}}$  and key descriptors. Pair plots of  $\Delta G_{\text{exf}}$  vs. (a)  $\Delta H_{\text{vap}}$ , (b)  $\epsilon$  and (c)  $\gamma$ . Pair plots of  $\Delta G_{\text{sol}}$  vs. (d)  $C_{\text{max}}$ , (e)  $\alpha_P$  and (f)  $\gamma$ .  $\Delta G_{\text{sol}}$  and  $\Delta G_{\text{exf}}$  are measured (from MD) values for the 48 organic solvents in our dataset. Five solvents with the lowest  $\Delta G_{\text{exf}}$  and highest  $\Delta G_{\text{sol}}$  are highlighted with red color; their numbering is consistent with



Appendix B section B1, and their name are given in (a) and (d). The correlation is linear in (a) and (d), nonlinear in (b) and (e), and unclear in (c) and (f).

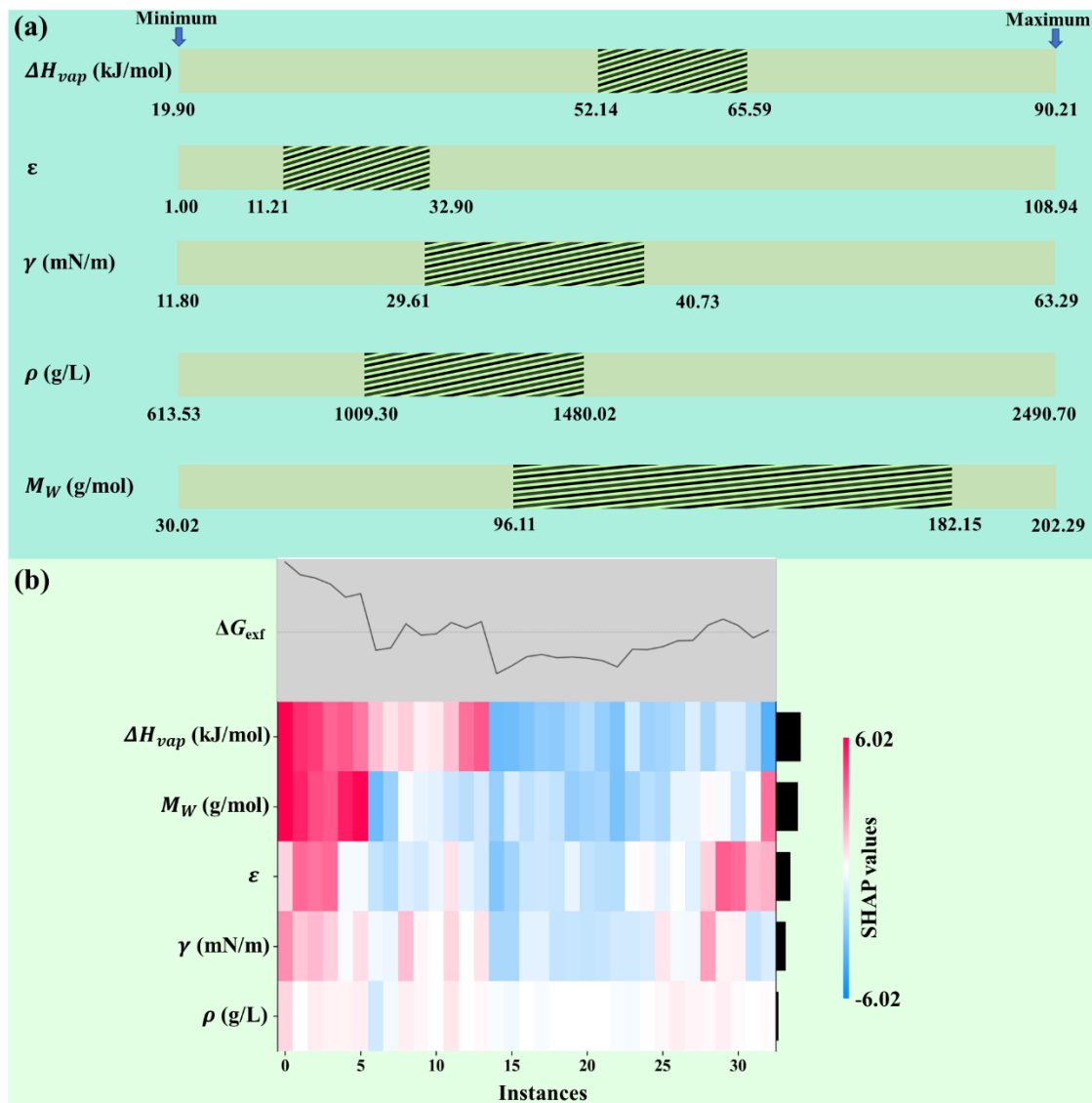


Figure 4.10. Guidance on solvent selection for LPE of g-C<sub>3</sub>N<sub>4</sub> gained from ML modeling. (a) Proposed ranges of descriptors (shown with pattern bars), based on the solvents with predicted  $\Delta G_{exf} < 51$  kJ/mol·nm<sup>2</sup>. The maximum and minimum of each descriptor is also shown for the experimental values of 153 organic solvents listed in Appendix B section B4. (b) Heatmap of SHAP values for each descriptor in the ETR model. The color intensity represents the magnitude of the SHAP value, with red indicating a positive impact and blue indicating a negative impact on

the predicted  $\Delta G_{\text{exf}}$ . The descriptors are ordered by their impact on  $\Delta G_{\text{exf}}$ , with decreasing significance from top to bottom.

## 4.5. Limitations and future considerations

While MD simulations offer valuable insights into material behaviors at the molecular level, it is crucial to acknowledge their inherent limitations. Of particular note is the quality of interatomic potentials used in the simulations, which can affect the accuracy of the simulation results. To mitigate this limitation, extensive validations were conducted during the parameterization of the interatomic potentials used in our work. Key properties of g-C<sub>3</sub>N<sub>4</sub>, such as bond length/angle distributions, surface energy and hydration free energy, were rigorously compared with results from DFT calculations. For the solvents, validations were carried out by finding good agreement when comparing their calculated density, dielectric constant and surface tension against experimental data. Details of these validations can be found in our previous work [21] and in Coleman et al [41].

Another inherent limitation of MD simulation is the demand of high computational resources when conducting large-scale simulations. As such, simplifications (e.g., size reduction compared to real-world systems) are almost unavoidable in order to make the running time of the simulations realistic. In this work, to calculate  $\Delta G_{\text{exf}}$  we considered the PMF between two g-C<sub>3</sub>N<sub>4</sub> nanosheets, while in reality LPE likely functions on a multi-layer structure. The choice of reducing the number of layers to two is based on the computational demand (the current simulations for two layers already took months on a national high-performance computing platform), as well as the consideration that the closest layer is expected to dominate  $\Delta G_{\text{exf}}$ , a notion also supported by other studies [12,15]. This approach serves the primary purpose of this work, which is to establish a model that correlates the free energy with the solvent properties and use the same model across a range of solvents to guide their selection for LPE. While the validity of the model is proven by experimental dispersibility tests, we acknowledge that the free energy to exfoliate a unit area of g-C<sub>3</sub>N<sub>4</sub> nanosheet from a multi-layer structure can be different from  $\Delta G_{\text{exf}}$  reported in this work, and is subjected to future investigation.

Acknowledgment of limitations can also be extended to the ML modeling. While descriptor reduction mitigates the risk of overfitting and reduces model complexity, this process could potentially affect applicability of the model to new solvents. Data augmentation methods such as Bootstrap and synthetic minority over-sampling technique (SMOTE) can increase the number of samples and help achieve better predictive performance. In this work, we applied the Bootstrap method for the ETR model with the optimal subset of descriptors, and the results indicate that our finalized model is robust and well calibrated. Nevertheless, as new solvents are added to our database, it is prudent to closely monitor the model performance to ensure it remains accurate or else adjust the descriptors in the model. Another limitation is that while the ML models here are effective in predicting free energies based on solvent descriptors, they do not account for other factors that may be crucial in practical situations, such as sonication and centrifugation settings, solvent stability, cost, availability, and environmental impact. The models are built on data generated under specific conditions and might not fully capture variability arising from different experimental setups. In addition, the applications of g-C<sub>3</sub>N<sub>4</sub> are broad and include areas such as photovoltaic cells, catalysts, and water treatment systems. Solvents identified as being effective for LPE might have limitations in subsequent applications. For instance, a solvent that is excellent for exfoliating g-C<sub>3</sub>N<sub>4</sub> may not be ideal for depositing a thin and uniform film of the material for photovoltaic applications. These factors should be taken into consideration for the successful application of the ML models beyond the scope of this study.

## 4.6. Conclusion

Molecular dynamics simulations were combined with machine learning (ML) algorithms to predict the free energy of exfoliation and solvation free energy for the liquid phase exfoliation (LPE) of g-C<sub>3</sub>N<sub>4</sub> nanosheets. A dataset involving 48 organic solvents was built using extensive MD simulations that evaluated the two free energies as well as bulk properties of the solvents. Six different ML methods were attempted, considering the two free energies as the target properties and 15 descriptors for the free energy of exfoliation (14 for the solvation free energy). Extra-trees regressor demonstrated the best ability to predict both target properties. For some descriptors (e.g., dielectric constant, surface tension), a certain range was identified which would lead to better LPE performance. The established model allowed us to predict the free energies of exfoliation and

solvation for g-C<sub>3</sub>N<sub>4</sub> in different solvents based on available experimental data, thus reducing the need to perform additional lengthy simulations. In fact, by applying the ML model to over 100 additional solvents, we identified benzyl alcohol (Bn) as a promising solvent, which had not been previously used for the LPE of 2D materials. Experimental dispersibility tests on four solvents (NMP, MSA, Bn, and ACN) provided evidence for the efficacy of the ML model and substantiated the applicability of our model in guiding solvent selection for LPE. The approach undertaken here can be adapted to predict the performance of solvents for other layered materials.

#### 4.7. Acknowledgement

Digital Research Alliance of Canada is gratefully acknowledged for providing the computing resources and technical support. TT acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Grant number: RGPIN-2018-04281) and the Canada Research Chair program (CRC; Grant number: TIER1 2021-00023). KS thanks NSERC for financial support (Grant number: RGPIN-2020-04620). The authors appreciate the assistance from John Garcia while performing the dispersibility tests.

#### 4.8. References

- [1] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [2] G.Q. Zhao, J. Zou, J. Hu, X. Long, F.P. Jiao, A critical review on graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>)-based composites for environmental remediation, *Sep Purif Technol* 279 (2021) 119769. <https://doi.org/10.1016/J.SEPPUR.2021.119769>.
- [3] Y. Liu, D. Xie, M. Song, L. Jiang, G. Fu, L. Liu, J. Li, Water desalination across multilayer graphitic carbon nitride membrane: Insights from non-equilibrium molecular dynamics simulations, *Carbon N Y* 140 (2018) 131–138. <https://doi.org/10.1016/j.carbon.2018.08.043>.

- [4] A. Alaghmandfard, K. Ghandi, A Comprehensive Review of Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)–Metal Oxide-Based Nanocomposites: Potential for Photocatalysis and Sensing, *Nanomaterials* 2022, Vol. 12, Page 294 12 (2022) 294. <https://doi.org/10.3390/NANO12020294>.
- [5] Y. Dong, Q. Wang, H. Wu, Y. Chen, C.H. Lu, Y. Chi, H.H. Yang, Graphitic Carbon Nitride Materials: Sensing, Imaging and Therapy, *Small* 12 (2016) 5376–5393. <https://doi.org/10.1002/SMLL.201602056>.
- [6] L. Bai, H. Huang, S. Yu, D. Zhang, H. Huang, Y. Zhang, Role of transition metal oxides in g-C<sub>3</sub>N<sub>4</sub>-based heterojunctions for photocatalysis and supercapacitors, *Journal of Energy Chemistry* 64 (2022) 214–235. <https://doi.org/10.1016/J.JECHEM.2021.04.057>.
- [7] K.M. Alam, N. Chaulagain, E. Shahini, M. Masud Rana, J. Garcia, N. Kumar, A.E. Kobryn, S. Gusarov, T. Tang, K. Shankar, Low bandgap carbon nitride nanoparticles incorporated in titania nanotube arrays by in situ electrophoretic anodization for photocatalytic CO<sub>2</sub> reduction, *Chemical Engineering Journal* 456 (2023) 141067. <https://doi.org/10.1016/J.CEJ.2022.141067>.
- [8] K.M. Alam, C.E. Jensen, P. Kumar, R.W. Hooper, G.M. Bernard, A. Patidar, A.P. Manuel, N. Amer, A. Palmgren, D.N. Purschke, N. Chaulagain, J. Garcia, P.S. Kirwin, L.C.T. Shoute, K. Cui, S. Gusarov, A.E. Kobryn, V.K. Michaelis, F.A. Hegmann, K. Shankar, Photocatalytic Mechanism Control and Study of Carrier Dynamics in CdS@C<sub>3</sub>N<sub>5</sub>Core-Shell Nanowires, *ACS Appl Mater Interfaces* 13 (2021) 47418–47439. <https://doi.org/10.1021/ACSAMI.1C08550>/ASSET/IMAGES/LARGE/AM1C08550\_0011.JPEG .
- [9] S. Angizi, M. Ali Akbar, M. Darestani-Farahani, al -, S. Dhanraj Nehate, S. Sundaresh, R. Peale, P. Kumar, S. Mulmi, D. Laishram, K.M. Alam, U.K. Thakur, V. Thangadurai, K. Shankar, Water-splitting photoelectrodes consisting of heterojunctions of carbon nitride with a p-type low bandgap double perovskite oxide, *Nanotechnology* 32 (2021) 485407. <https://doi.org/10.1088/1361-6528/ABEDEC>.
- [10] B. Abreu, B. Almeida, P. Ferreira, R. M. F. Fernandes, D.M. Fernandes, E.F. Marques, A critical assessment of the role of ionic surfactants in the exfoliation and stabilization of 2D nanosheets: The case of the transition metal dichalcogenides MoS<sub>2</sub>, WS<sub>2</sub> and MoSe<sub>2</sub>, *J Colloid Interface Sci* 626 (2022) 167–177. <https://doi.org/10.1016/J.JCIS.2022.06.097>.

- [11] I. Benabdallah, A. Kara, M. Benaissa, Exfoliation and re-aggregation mechanisms of black phosphorus: A molecular dynamics study, *Appl Surf Sci* 507 (2020).  
<https://doi.org/10.1016/j.apsusc.2019.144826>.
- [12] R. Biswas, Molecular dynamics studies on the exfoliation of graphene in room temperature ionic liquids, *J Mol Liq* 337 (2021) 116592. <https://doi.org/10.1016/j.molliq.2021.116592>.
- [13] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.
- [14] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268.  
<https://doi.org/10.1021/acsnano.5b02683>.
- [15] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648.  
<https://doi.org/10.1021/ja1064284>.
- [16] G. Zhou, P. Rajak, S. Susarla, P.M. Ajayan, R.K. Kalia, A. Nakano, P. Vashishta, Molecular Simulation of MoS<sub>2</sub> Exfoliation, *Sci Rep* 8 (2018) 1–9. <https://doi.org/10.1038/s41598-018-35008-z>.
- [17] Y. Hernandez, M. Lotya, D. Rickard, S.D. Bergin, J.N. Coleman, Measurement of multicomponent solubility parameters for graphene facilitates solvent discovery, *Langmuir* 26 (2010) 3208–3213. <https://doi.org/10.1021/la903188a>.
- [18] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. v. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuvsen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/SCIENCE.1194975>.

- [19] C. Wu, S. Xue, Z. Qin, M. Nazari, G. Yang, S. Yue, T. Tong, H. Ghasemi, F.C.R. Hernandez, S. Xue, D. Zhang, H. Wang, Z.M. Wang, S. Pu, J. Bao, Making g-C<sub>3</sub>N<sub>4</sub> ultra-thin nanosheets active for photocatalytic overall water splitting, *Appl Catal B* 282 (2021) 119557. <https://doi.org/10.1016/j.apcatb.2020.119557>.
- [20] S.P. Pattnaik, A. Behera, S. Martha, R. Acharya, K. Parida, Facile synthesis of exfoliated graphitic carbon nitride for photocatalytic degradation of ciprofloxacin under solar irradiation, *J Mater Sci* 54 (2019) 5726–5742. <https://doi.org/10.1007/s10853-018-03266-x>.
- [21] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910. <https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [22] J.S. Delaney, ESOL: Estimating aqueous solubility directly from molecular structure, *J Chem Inf Comput Sci* 44 (2004) 1000–1005. [https://doi.org/10.1021/CI034243X/SUPPL\\_FILE/CI034243XSI20040112\\_053635.TXT](https://doi.org/10.1021/CI034243X/SUPPL_FILE/CI034243XSI20040112_053635.TXT).
- [23] C. Reichardt, T. Welton, *Solvents and Solvent Effects in Organic Chemistry: Fourth Edition*, Solvents and Solvent Effects in Organic Chemistry: Fourth Edition (2010). <https://doi.org/10.1002/9783527632220>.
- [24] A.P. Sgouros, C.J. Revelas, A.T. Lakkas, D.N. Theodorou, Solvation Free Energy of Dilute Grafted (Nano)Particles in Polymer Melts via the Self-Consistent Field Theory, *J Phys Chem B* (2022). <https://doi.org/10.1021/ACS.JPCB.2C05306>.
- [25] L. Pereyaslavets, G. Kamath, O. Butin, A. Illarionov, M. Olevanov, I. Kurnikov, S. Sakipov, I. Leontyev, E. Voronina, T. Gannon, G. Nawrocki, M. Darkhovskiy, I. Ivahnenko, A. Kostikov, J. Scaranto, M.G. Kurnikova, S. Banik, H. Chan, M.G. Sternberg, S.K.R.S. Sankaranarayanan, B. Crawford, J. Potoff, M. Levitt, R.D. Kornberg, B. Fain, Accurate determination of solvation free energies of neutral organic compounds from first principles, *Nature Communications* 2022 13:1 13 (2022) 1–7. <https://doi.org/10.1038/s41467-022-28041-0>.
- [26] F.H. Vermeire, W.H. Green, Transfer learning for solvation free energies: From quantum chemistry to experiments, *Chemical Engineering Journal* 418 (2021) 129307. <https://doi.org/10.1016/J.CEJ.2021.129307>.

- [27] Z. Wan, Q. de Wang, D. Liu, J. Liang, Accelerating the optimization of enzyme-catalyzed synthesis conditions via machine learning and reactivity descriptors, *Org Biomol Chem* 19 (2021) 6267–6273. <https://doi.org/10.1039/D1OB01066B>.
- [28] S. Perumal, R. Atchudan, T.N.J. Immanuel Edison, J.J. Shim, Y.R. Lee, Exfoliation and Noncovalent Functionalization of Graphene Surface with Poly-N-Vinyl-2-Pyrrolidone by In Situ Polymerization, *Molecules* 2021, Vol. 26, Page 1534 26 (2021) 1534. <https://doi.org/10.3390/MOLECULES26061534>.
- [29] Y. Saito, K. Shin, K. Terayama, S. Desai, M. Onga, Y. Nakagawa, Y.M. Itahashi, Y. Iwasa, M. Yamada, K. Tsuda, Deep-learning-based quality filtering of mechanically exfoliated 2D crystals, *Npj Computational Materials* 2019 5:1 5 (2019) 1–6. <https://doi.org/10.1038/s41524-019-0262-4>.
- [30] E.M.D. Siriwardane, R.P. Joshi, N. Kumar, D. Çakır, Revealing the Formation Energy-Exfoliation Energy-Structure Correlation of MAB Phases Using Machine Learning and DFT, *ACS Appl Mater Interfaces* 12 (2020) 29424–29431. [https://doi.org/10.1021/ACSAMI.0C03536/ASSET/IMAGES/LARGE/AM0C03536\\_0003.JPEG](https://doi.org/10.1021/ACSAMI.0C03536/ASSET/IMAGES/LARGE/AM0C03536_0003.JPEG).
- [31] Z. Wan, Q. de Wang, Machine Learning Prediction of the Exfoliation Energies of Two-Dimension Materials via Data-Driven Approach, *Journal of Physical Chemistry Letters* 12 (2021) 11470–11475. [https://doi.org/10.1021/ACS.JPCLETT.1C03335/ASSET/IMAGES/LARGE/JZ1C03335\\_0006.JPEG](https://doi.org/10.1021/ACS.JPCLETT.1C03335/ASSET/IMAGES/LARGE/JZ1C03335_0006.JPEG).
- [32] W.K. Darkwah, Y. Ao, Mini Review on the Structure and Properties (Photocatalysis), and Preparation Techniques of Graphitic Carbon Nitride Nano-Based Particle, and Its Applications, *Nanoscale Research Letters* 2018 13:1 13 (2018) 1–15. <https://doi.org/10.1186/S11671-018-2702-3>.
- [33] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon N Y* 141 (2019) 580–607. <https://doi.org/10.1016/J.CARBON.2018.09.082>.



- [34] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [35] X. Chen, R. Hu, DFT-based study of single transition metal atom doped g-C<sub>3</sub>N<sub>4</sub> as alternative oxygen reduction reaction catalysts, *Int J Hydrogen Energy* 44 (2019) 15409–15416. <https://doi.org/10.1016/J.IJHYDENE.2019.04.057>.
- [36] Z. Bonakchi, A. Nakhaei Pour, S. Soheili, Molecular simulation of methane on various g-C<sub>3</sub>N<sub>4</sub> isomers: collision, adsorption, desorption, and diffusion studies, *Journal of the Iranian Chemical Society* 19 (2022) 3649–3657. <https://doi.org/10.1007/S13738-022-02562-3/TABLES/2>.
- [37] W.L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, *Proc Natl Acad Sci U S A* (2005). <https://doi.org/10.1073/pnas.0408037102>.
- [38] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).
- [39] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [40] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/science.1194975>.
- [41] C. Caleman, P.J. Van Maaren, M. Hong, J.S. Hub, L.T. Costa, D. Van Der Spoel, Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension,

isothermal compressibility, volumetric expansion coefficient, and dielectric constant, *J Chem Theory Comput* 8 (2012) 61–74. <https://doi.org/10.1021/ct200731v>.

- [42] V. Consonni, R. Todeschini, Molecular descriptors, *Challenges and Advances in Computational Chemistry and Physics* 8 (2010) 29–102. [https://doi.org/10.1007/978-1-4020-9783-6\\_3](https://doi.org/10.1007/978-1-4020-9783-6_3).
- [43] J.P.M. Jcambeck, F. Mocci, A.P. Lyubartsev, A. Laaksonen, Partial atomic charges and their impact on the free energy of solvation, *J Comput Chem* 34 (2013) 187–197. <https://doi.org/10.1002/JCC.23117>.
- [44] J. Tukey, *Exploratory data analysis*, 1977. [http://theta.edu.pl/wp-content/uploads/2012/10/exploratorydataanalysis\\_tukey.pdf](http://theta.edu.pl/wp-content/uploads/2012/10/exploratorydataanalysis_tukey.pdf) (accessed September 2, 2022).
- [45] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, *Machine Learning* 2006 63:1 63 (2006) 3–42. <https://doi.org/10.1007/S10994-006-6226-1>.
- [46] J.H. Friedman, Stochastic gradient boosting, *Comput Stat Data Anal* 38 (2002) 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2).
- [47] R.E. Schapire, Explaining adaboost, *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik* (2013) 37–52. [https://doi.org/10.1007/978-3-642-41136-6\\_5/COVER](https://doi.org/10.1007/978-3-642-41136-6_5/COVER).
- [48] F.A. da Silva, A.P. Viana, C.C.G. Correa, E.A. Santos, J.A.V.S. de Oliveira, J.D.G. Andrade, R.M. Ribeiro, L.S. Glória, Bayesian ridge regression shows the best fit for SSR markers in *Psidium guajava* among Bayesian models, *Scientific Reports* 2021 11:1 11 (2021) 1–11. <https://doi.org/10.1038/s41598-021-93120-z>.
- [49] C.A. Ramezan, T.A. Warner, A.E. Maxwell, Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification, *Remote Sensing* 2019, Vol. 11, Page 185 11 (2019) 185. <https://doi.org/10.3390/RS11020185>.
- [50] J.M. Sutter, J.H. Kalivas, Comparison of Forward Selection, Backward Elimination, and Generalized Simulated Annealing for Variable Selection, *Microchemical Journal* 47 (1993) 60–66. <https://doi.org/10.1006/MCHJ.1993.1012>.

- [51] M.A. Hall, L.A. Smith, Practical feature subset selection for machine learning, *Computer Science* Volume 20 No 1 (1998) 181–191.  
<https://researchcommons.waikato.ac.nz/handle/10289/1512> (accessed August 13, 2022).
- [52] K.P. Shroff, H.H. Maheta, A comparative study of various feature selection techniques in high-dimensional data set to improve classification accuracy, 2015 International Conference on Computer Communication and Informatics, ICCCI 2015 (2015).  
<https://doi.org/10.1109/ICCCI.2015.7218098>.
- [53] F. Pedregosa FABIANPEDREGOSA, V. Michel, O. Grisel OLIVIERGRISEL, M. Blondel, P. Prettenhofer, R. Weiss, J. Vanderplas, D. Cournapeau, F. Pedregosa, G. Varoquaux, A. Gramfort, B. Thirion, O. Grisel, V. Dubourg, A. Passos, M. Brucher, M. Perrot andÉdouardand, andÉdouard Duchesnay, Fré. Duchesnay EDOUARDDUCHESNAY, Scikit-learn: Machine learning in Python, *Jmlr.Org* 12 (2011) 2825–2830.  
<https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://> (accessed October 12, 2023).
- [54] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199.  
[https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [55] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [56] N. Ye, Z. Wang, S. Wang, H. Fang, D. Wang, Aqueous aggregation and stability of graphene nanoplatelets, graphene oxide, and reduced graphene oxide in simulated natural environmental conditions: complex roles of surface and solution chemistry, *Environmental Science and Pollution Research* 25 (2018) 10956–10965. <https://doi.org/10.1007/S11356-018-1326-6/FIGURES/7>.
- [57] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).

- [58] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [59] M. Neumann, O. Steinhauser, G. Stuart Pawley, Consistent calculation of the static and frequency-dependent dielectric constant in computer simulations, *Https://Doi.Org/10.1080/00268978400101081 52* (2006) 97–113. <https://doi.org/10.1080/00268978400101081>.
- [60] M.P. Allen, D.J. Tildesley, *Computer simulation of liquids*, 2017. <https://books.google.com/books?hl=en&lr=&id=WFExDwAAQBAJ&oi=fnd&pg=PP1&ots=VGTF41KD4V&sig=M1bd3YNU5iqGsmQa-Bt47iBd6Tg> (accessed August 21, 2022).
- [61] S.T. Lin, P.K. Maiti, W.A. Goddard, Two-phase thermodynamic model for efficient and accurate absolute entropy of water from molecular dynamics simulations, *Journal of Physical Chemistry B* 114 (2010) 8191–8198. [https://doi.org/10.1021/JP103120Q/ASSET/IMAGES/LARGE/JP-2010-03120Q\\_0004.JPEG](https://doi.org/10.1021/JP103120Q/ASSET/IMAGES/LARGE/JP-2010-03120Q_0004.JPEG).
- [62] S.T. Lin, M. Blanco, W.A. Goddard, The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of Lennard-Jones fluids, *J Chem Phys* 119 (2003) 11792. <https://doi.org/10.1063/1.1624057>.
- [63] D. van der Spoel, P.J. van Maaren, The origin of layer structure artifacts in simulations of liquid water, *J Chem Theory Comput* 2 (2006) 1–11. <https://doi.org/10.1021/CT0502256>.
- [64] C.-Y. Chang, M.-T. Hsu, E.X. Esposito, Y.J. Tseng, Oversampling to Overcome Overfitting: Exploring the Relationship between Data Set Composition, Molecular Descriptors, and Predictive Modeling Methods, (2013). <https://doi.org/10.1021/ci4000536>.
- [65] B. Efron, R. Tibshirani, *An introduction to the bootstrap*, 1994. [https://books.google.com/books?hl=en&lr=&id=gLlpIUxRntoC&oi=fnd&pg=PR14&ots=AaAuZ9QdF7&sig=x-o2EoZMRM8w7J\\_-BTLo8GSY3TI](https://books.google.com/books?hl=en&lr=&id=gLlpIUxRntoC&oi=fnd&pg=PR14&ots=AaAuZ9QdF7&sig=x-o2EoZMRM8w7J_-BTLo8GSY3TI) (accessed September 19, 2022).
- [66] Y. Zhang, C. Ling, A strategy to apply machine learning to small datasets in materials science, *Npj Computational Materials* 2018 4:1 4 (2018) 1–8. <https://doi.org/10.1038/s41524-018-0081-z>.
- [67] Y. Liu, T. Zhao, W. Ju, S. Shi, Materials discovery and design using machine learning, *Journal of Materiomics* 3 (2017) 159–177. <https://doi.org/10.1016/J.JMAT.2017.08.002>.

- [68] Y. Nademi, T. Tang, H. Uludağ, Modeling Uptake of Polyethylenimine/Short Interfering RNA Nanoparticles in Breast Cancer Cells Using Machine Learning, *Adv Nanobiomed Res* 1 (2021) 2000106. <https://doi.org/10.1002/ANBR.202000106>.
- [69] I. Tsamardinos, E. Greasidou, G. Borboudakis, Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation, *Mach Learn* 107 (2018) 1895–1922. <https://doi.org/10.1007/S10994-018-5714-4/FIGURES/7>.
- [70] R. Singh, C.C. Tripathi, Enhancing Liquid-Phase Exfoliation of Graphene with Addition of Anthracene in Organic Solvents, *Arab J Sci Eng* 42 (2017) 2417–2424. <https://doi.org/10.1007/S13369-016-2379-9/METRICS>.
- [71] B. Liang, K. Liu, P. Liu, L. Qian, G. Zhao, W. Pan, C. Chen, Organic salt-assisted liquid-phase shear exfoliation of expanded graphite into graphene nanosheets, *Journal of Materiomics* 7 (2021) 1181–1189. <https://doi.org/10.1016/J.JMAT.2021.03.007>.
- [72] D. Sahoo, B. Kumar, J. Sinha, S. Ghosh, S.S. Roy, B. Kaviraj, Cost effective liquid phase exfoliation of MoS<sub>2</sub> nanosheets and photocatalytic activity for wastewater treatment enforced by visible light, *Scientific Reports* 2020 10:1 10 (2020) 1–12. <https://doi.org/10.1038/s41598-020-67683-2>.
- [73] Y.J. Yuan, Z. Shen, S. Wu, Y. Su, L. Pei, Z. Ji, M. Ding, W. Bai, Y. Chen, Z.T. Yu, Z. Zou, Liquid exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets to construct 2D-2D MoS<sub>2</sub>/g-C<sub>3</sub>N<sub>4</sub> photocatalyst for enhanced photocatalytic H<sub>2</sub> production activity, *Appl Catal B* 246 (2019) 120–128. <https://doi.org/10.1016/J.APCATB.2019.01.043>.
- [74] W. Du, X. Jiang, L. Zhu, From graphite to graphene: direct liquid-phase exfoliation of graphite to produce single- and few-layered pristine graphene, *J Mater Chem A Mater* 1 (2013) 10592–10606. <https://doi.org/10.1039/C3TA12212C>.
- [75] Y. Wang, Z. Shi, J. Yin, Boron nitride nanosheets: large-scale exfoliation in methanesulfonic acid and their composites with polybenzimidazole, *J Mater Chem* 21 (2011) 11371–11377. <https://doi.org/10.1039/C1JM10342C>.
- [76] S.M. Lundberg, P.G. Allen, S.-I. Lee, A Unified Approach to Interpreting Model Predictions, (n.d.). <https://doi.org/10.5555/3295222.3295230>.

# Chapter 5: Hierarchy in Binary Liquid Phase Exfoliation of Graphitic Carbon Nitride: Dissecting the Dominance of One Solvent

## 5.1. Introduction

LPE has been identified as a critical method for producing 2D nanosheets consisting of one or a few layers (e.g. graphene or g-C<sub>3</sub>N<sub>4</sub>), a process essential for harnessing the distinct photocatalytic, electronic, and structural properties of these 2D materials [1–4]. The selection of solvents, a cornerstone of this process, has traditionally focused on single-component systems [5,6]. Yet, the exploration of binary solvent mixtures opens new avenues by offering the potential to enhance exfoliation efficiency via customization of solvent compositions. For example, Lin et al. [7] reported that ethanol and water mixed at 1:3 volume ratio maximized nanosheet concentration and markedly enhanced the visible-light photocatalytic capabilities of g-C<sub>3</sub>N<sub>4</sub> nanosheets. Shen et al. [8] optimized the exfoliation efficiency by aligning the surface tension of binary solvents with those of 2D materials. Their findings identified a number of optimal volume ratios for IPA and water binary mixture, such as 1:1 for graphene, WS<sub>2</sub>, h-BN and MoSe<sub>2</sub>; 1:4 for Bi<sub>2</sub>Se<sub>3</sub> and SnS<sub>2</sub>; and 7:3 for MoS<sub>2</sub>. Advancing these methodologies, Zou et al. [9] introduced a three-step tactic for LPE of g-C<sub>3</sub>N<sub>4</sub> nanosheets, involving the initial preparation of bulk g-C<sub>3</sub>N<sub>4</sub>, followed by solvent exfoliation using various solvents, and finally optimizing the exfoliation conditions based on MD simulations. They identified Water:IPA mixed at 1:1.75 volume ratio as being the most effective. This strategy notably increased the nanosheet yield to an unprecedented concentration of 5.03 mg/mL. Together, these studies have paved the way for the development of efficient and scalable methods to synthesize 2D nanomaterials for a wide range of uses.

While significant progress has been made in employing mixed solvents for LPE, it remains unclear why some mixtures outshine others. Existing studies often hinge on the idea that matching the surface tension of solvents with the 2D material optimizes exfoliation efficiency. However,

insights from our prior work [10,11] revealed that the reality of LPE is more intricate, with multiple factors like the size, polarity and structure of the solvent molecules impacting exfoliation success even for a single solvent. The situation becomes more interesting in a solvent mixture where different solvents are mutually interacting. This motivated us to take an approach that evaluates solvent performance through a free energy lens, using a quantity called the free energy of exfoliation ( $\Delta G_{\text{exf}}$ ) calculated from MD simulations. In our exploration of how a binary solvent mixture performs in LPE of g-C<sub>3</sub>N<sub>4</sub> compared to its singular counterparts, we discovered an interesting phenomenon: while some binary solvents have  $\Delta G_{\text{exf}}$  values between those of their singular counterparts, there exist situations where  $\Delta G_{\text{exf}}$  is significantly skewed towards one of the constituting solvents. As a result, one of the solvents in the mixture can exhibit dominance in LPE, by suppressing either the “bad” solvent with high  $\Delta G_{\text{exf}}$  or the “good” solvent with low  $\Delta G_{\text{exf}}$ . Detailed analysis of the MD simulation data allowed us to dissect the dominant role of one solvent, and shed light on why certain mixtures achieve superior performance.

Besides physical insights, this work presents a systematic and quantitative approach that can be more broadly applied to gauge the effectiveness of solvent mixtures. This approach includes an initial solvent screening process using an established ML model, followed by more precise free energy calculations using umbrella sampling. Finally, based on the potential dominance of one solvent in a binary mixture, we demonstrate that cheaper and cleaner LPE of 2D nanomaterials can be achieved by mixing certain organic solvents with water (a known bad solvent for g-C<sub>3</sub>N<sub>4</sub>), at no significant cost of the LPE performance.

## 5.2. Methods

### 5.2.1. Structure of simulated g-C<sub>3</sub>N<sub>4</sub> and solvents

For this study, g-C<sub>3</sub>N<sub>4</sub> is constructed from heptazine units, as opposed to triazine-based counterparts [12,13]. This choice is guided by the prevalence of heptazine in g-C<sub>3</sub>N<sub>4</sub> compositions, as corroborated by the literature [14]. Consistent with our previous work [10], the molecular model employs an equilateral triangular nanosheet, incorporating 15 heptazine units with each side extending 4.0 nm, as illustrated in Figure 5.1. The parameters requisite for describing the intra- and inter- molecular interactions are derived using the LigParGen and PolyParGen servers [15,16]. These parameters align with the OPLS-AA force field [17] and incorporate the Charge Model 5

[18] with a scaling factor of 1.20 for partial atomic charges. This approach and the accuracy of the parameters it generates have been previously validated [10,11].

A total of 171 solvents are considered for initial screening (see details in Section 5.2.2), as listed in Appendix C, Section C1. Two binary solvents that exhibit significant dominance by one component are then subjected to further investigation. These two binary solvents are: NMP mixed with Cyclohexane (CH), and methanol (MET) mixed with Dichloromethane (DCM), both at a 1:1 volume ratio. The chemical structure and molecular mass of the four constituting solvents are shown in Figure 5.1.

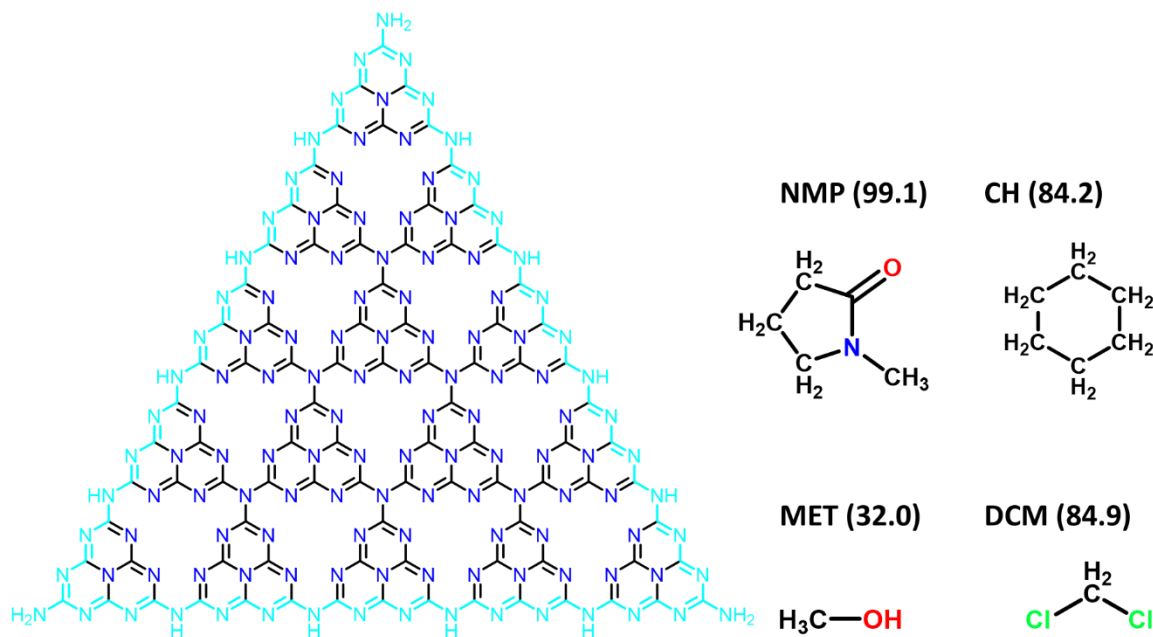


Figure 5.1. Molecular structure of the 15-unit g-C<sub>3</sub>N<sub>4</sub> alongside four solvents used in US simulations. The molecular mass of each solvent is provided in parentheses. For the nanosheet, atoms highlighted in cyan are considered as periphery, while the remaining atoms are considered as interior.

### 5.2.2. Initial screening of binary mixtures

In the previous chapter, a ML model has been established based on Extra Trees Regressor, which is capable of accurately predicting the  $\Delta G_{\text{exf}}$  for a single solvent based on physicochemical properties significant in LPE—namely, heat of vaporization ( $\Delta H_{\text{vap}}$ ), dielectric constant ( $\epsilon$ ), surface tension ( $\gamma$ ), mass density ( $\rho$ ), and molar weight ( $M_W$ ). Physically,  $\Delta G_{\text{exf}}$  is the free energy required to separate a unit area of two stacked g-C<sub>3</sub>N<sub>4</sub> nanosheets from their equilibrium distance



to a state where they are no longer interacting (see quantitative definition of  $\Delta G_{\text{exf}}$  in Figure 5.4a). A lower  $\Delta G_{\text{exf}}$  signifies not only reduced energy expenditure during LPE but also enhanced stability of the separated nanosheets post-exfoliation, underscoring its critical role in measuring solvent efficacy.

Built on this ML model, initial screening of binary solvents is performed. These binary solvents are created based on the 171 solvents specified in Appendix C Section C1 and prepared at a 1:1 volume ratio. Properties ( $\Delta H_{\text{vap}}$ ,  $\varepsilon$ ,  $\gamma$ ,  $\rho$ ,  $MW$ ) of the binary solvents are calculated from the properties of the individual components according to following equations. For  $\Delta H_{\text{vap}}$ ,  $\varepsilon$ ,  $\gamma$ ,

$$P_{\text{mix}} = x_1 P_1 + x_2 P_2 \quad (5.1)$$

where  $P_{\text{mix}}$  represents the property of the binary mixture,  $x_1$  and  $x_2$  are the mole fractions, and  $P_1$  and  $P_2$  are the respective properties ( $\Delta H_{\text{vap}}$ ,  $\varepsilon$ ,  $\gamma$ ) of the individual solvents. At 1:1 volume ratio, the mass density  $\rho$  of the mixture is given by

$$\rho_{\text{mix}} = \frac{\rho_1 + \rho_2}{2} \quad (5.2)$$

where  $\rho_1$  and  $\rho_2$  denote the mass densities of the individual solvents.  $MW$  of a mixture with 1:1 volume ratio can be expressed as

$$MW_{\text{mix}} = \frac{\rho_1 + \rho_2}{\left(\frac{\rho_1}{MW_1}\right) + \left(\frac{\rho_2}{MW_2}\right)} \quad (5.3)$$

where  $MW_1$  and  $MW_2$  are the molar weights of the individual solvents. Based on these properties,  $\Delta G_{\text{exf}}$  is predicted for every possible binary combination, corresponding to a total of 14,535 binary solvents. Additionally,  $\Delta G_{\text{exf}}$  is predicted for the 171 pure solvents.

By examining the predicted  $\Delta G_{\text{exf}}$  values, it is determined whether the combination of two solvents results in  $\Delta G_{\text{exf}}$  close to the arithmetic mean of its components, or dominated by one of the components. Selected binary solvents are further investigated by MD simulations, to verify the ML predictions as well as gain deeper insights into the molecular interactions driving the observed behaviors. The schematic representation of the methodological steps is depicted in Figure 5.2, illustrating the workflow from property calculation,  $\Delta G_{\text{exf}}$  prediction using ETR, to the selection and detailed analysis of noteworthy binary solvent systems.

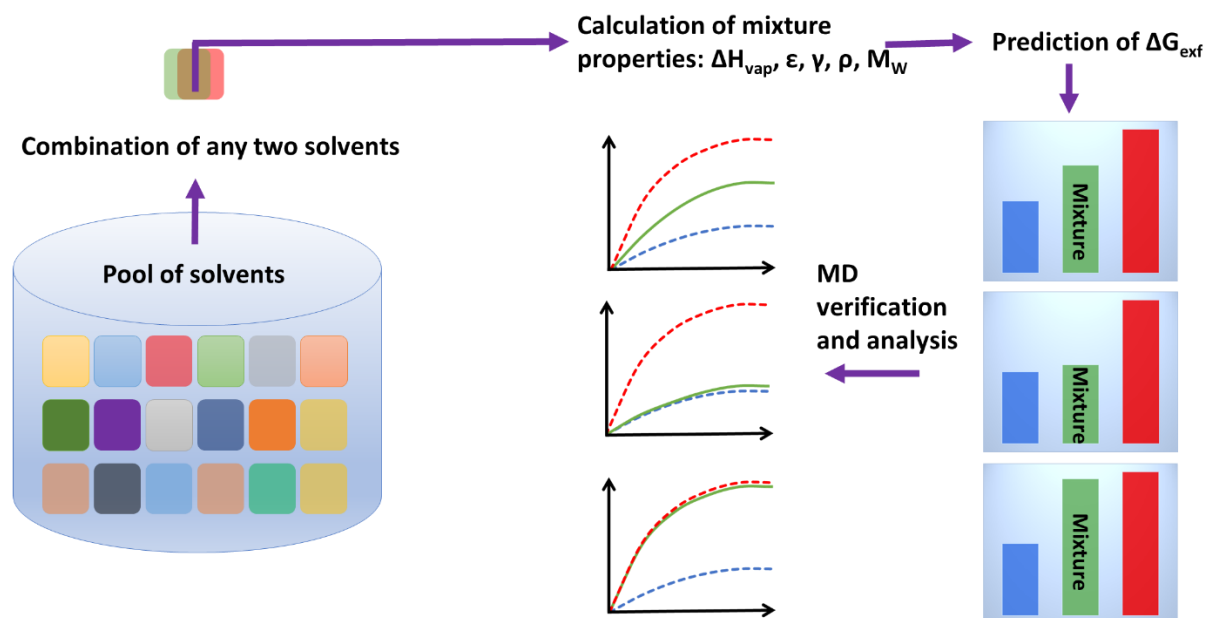


Figure 5.2. Flowchart illustrating steps taken in this work to examine LPE of g-C<sub>3</sub>N<sub>4</sub> by binary solvents.

### 5.2.3. Simulation details

$\Delta G_{\text{exf}}$  is evaluated via the computation of the PMF, as two initially adjacent nanosheets are separated along a predetermined reaction coordinate. Drawing from insights gained in our prior investigations [10,11], the most feasible pathway for the LPE of g-C<sub>3</sub>N<sub>4</sub> is identified as “shearing” where one nanosheet glides over another while maintaining a parallel orientation. Consequently, the reaction coordinate is designated as the distance  $D$  between the COMs of two adjacent nanosheets, and varied from 0.3 nm to 3.5 nm by laterally shifting one nanosheet (see inset of Figure 5.4a for a schematic depiction). The US [19] technique is utilized to perform a series of restrained simulations, from which the PMF is derived through the WHAM [20]. For each solvent, a total of 40 US windows are employed, within each window a harmonic biasing potential (force constant 5000 kJ/mol·nm<sup>2</sup>) is enforced on the reaction coordinate. The baseline of the PMF is set at the minimum value of the reaction coordinate.

Along with  $\Delta G_{\text{exf}}$ , the  $\Delta G_{\text{sol}}$  of an individual nanosheet is evaluated using the BAR method [21]. Beginning with a state indicative of zero interaction between the g-C<sub>3</sub>N<sub>4</sub> nanosheet and the solvent, the vdW and electrostatic interactions are incrementally activated through nineteen

intermediate states, culminating in a fully interactive state between the nanosheet and the solvent.  $\Delta G_{\text{sol}}$  is then calculated based on the data gathered across these twenty-one states. Furthermore, to investigate the dynamics of solvent molecules near the nanosheets, additional simulations are performed where a single or two stacked nanosheets are equilibrated in a binary mixture, as well as in their singular counterparts.

All simulations are conducted utilizing the GROMACS software [22] (version 2022.3). For each US simulation in which the reaction coordinate is restrained, the system first undergoes an energy minimization employing the steepest descent method. Following this, the system is equilibrated in an NPT ensemble for a duration of 200 ps, with temperature and pressure maintained at 300 K and 1 bar, respectively, using the Berendsen thermostat (time constant = 0.2 ps) and barostat (time constant = 5 ps). Subsequently, a production run is executed for 10 ns, during which the bond lengths that include a hydrogen atom are restricted using the LINCS algorithm. Temperature is maintained at 300 K using the Nosé-Hoover algorithm with a time constant of 0.1 ps, while an isotropic pressure of 1 bar is regulated using the C-rescale scheme with a time constant of 1.0 ps. Periodic boundary conditions are applied in all three directions. The computation of long-range electrostatic interactions is facilitated through the particle-mesh Ewald technique. For short-range nonbonded interactions, a cutoff radius of 1.2 nm is used. The equations of motion are integrated with a 2-fs time step in both equilibration and production stages. The simulation for  $\Delta G_{\text{sol}}$  calculation follows the same procedure and settings, except that there is only one nanosheet in the system and it is unrestrained, and that the production simulation for each intermediate state is run for 1 ns. The same parameters, including thermostat and barostat, are also adopted for the simulations where a single or two stacked nanosheets are equilibrated. Each of these simulations is run for 20 ns and data from the last 10 ns are used for statistical analysis.

## 5.3. Results and discussion

### 5.3.1. ML prediction of $\Delta G_{\text{exf}}$ in binary mixtures

Figure 5.3a shows the normalized histograms for  $\Delta G_{\text{exf}}$  of the binary solvent mixtures and the arithmetic average  $\Delta G_{\text{exf}}$  of the two components in each binary solvent. All data are predicted by the ML model. The histogram for the binary solvent mixtures appears to be wider compared to the histogram for the average of the individual solvents. In addition, more binary systems exhibit

$\Delta G_{\text{exf}}$  values lower than 50 kJ/mol·nm<sup>2</sup>, suggesting LPE performance better than what would be predicted by taking the arithmetic average of the two constituting solvents. To further investigate this behavior quantitatively, we categorized the binary solvent mixtures into three groups: above-par, par-reaching and sub-par performance. A binary solvent is considered above-par if its  $\Delta G_{\text{exf}}$  is in the bottom 25% of the range between the lower and higher  $\Delta G_{\text{exf}}$  values of the two solvents it contains. A sub-par performance is identified when  $\Delta G_{\text{exf}}$  of the binary solvent is in the top 25% of this range. Binary solvents that do not fall into either category are classified as par-reaching.

The results of this categorization show that 43.1% of the binary solvent mixtures exhibit above-par performance, 36.3% exhibit par-reaching performance, and 20.6% exhibit sub-par performance. Figure 5.3b presents a scatter plot of  $\Delta G_{\text{exf}}$  of binary solvents against the arithmetic average of its individual components. The red dashed line represents the line of equality, where  $\Delta G_{\text{exf}}$  of the binary mixture would equal the average of its components. Clearly, the scattered data show deviation from the red line, and if a straight line were to be fit to the data, its slope would be larger than 1. In other words, for small average  $\Delta G_{\text{exf}}$  the binary mixtures have a higher tendency to exhibit above-par performance, and for large average  $\Delta G_{\text{exf}}$  the binary mixtures have a higher tendency to show sub-par performance.

From the entire dataset, we selected three systems for further investigation: one with par-reaching performance (NMP:MET), one with above-par performance (NMP:CH), and the last with sub-par performance (MET:DCM). These three systems are highlighted in Figure 5.3b with yellow, green, and red colors, respectively. MD simulations and PMF calculations are performed to validate the ML predications as well as to study the molecular interactions driving their behaviors. The above constituting solvents are also chosen based on their wide availability in LPE of g-C<sub>3</sub>N<sub>4</sub> and other 2D materials, as well as the consideration that the two solvents in a binary mixture should be fully miscible.

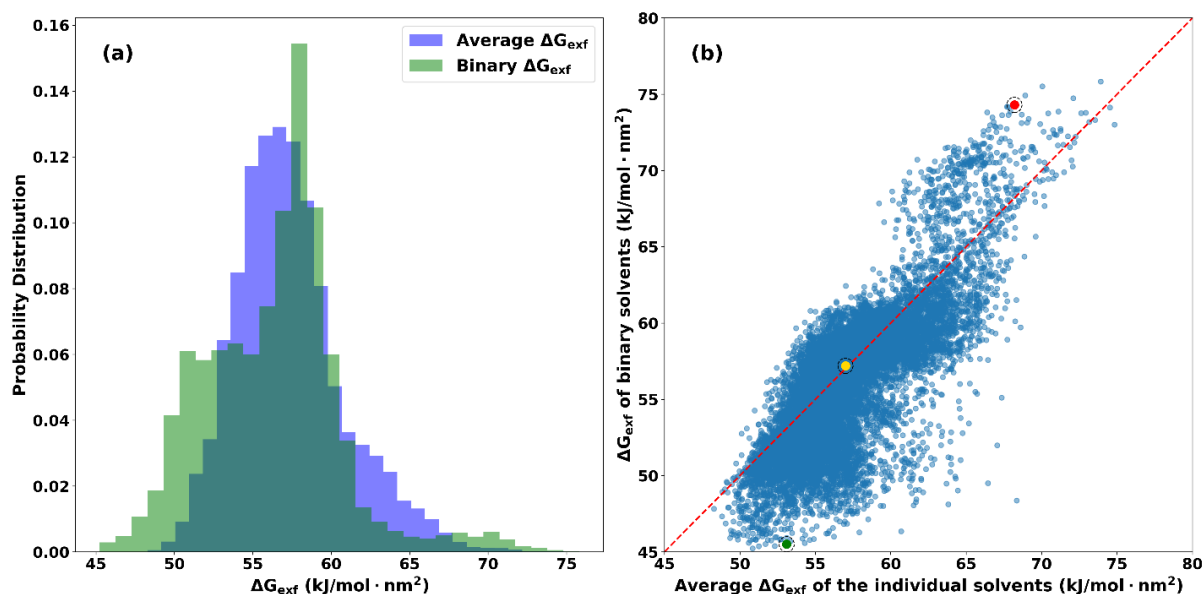


Figure 5.3. (a) Normalized histograms and (b) scatter plot of  $\Delta G_{\text{exf}}$  for the binary solvent mixtures as well as the arithmetic average of the two solvents in each binary mixture.

### 5.3.2. Validation of solvent dominance with MD

Figure 5.4 shows the PMF profiles for the three selected binary solvent mixtures and their singular counterparts, with part (a) specifically highlighting  $\Delta G_{\text{exf}}$  as the difference between the minimum PMF and the value where the PMF curve plateaus. In Figure 5.4a,  $\Delta G_{\text{exf}}$  of NMP:MET mixture is approximately the average of  $\Delta G_{\text{exf}}$  of NMP and MET, consistent with the ML prediction in Figure 5.3b and the traditional expectation that a solvent mixture behaves merely as a composite of its individual components. Conversely, the NMP:CH system in Figure 5.4b demonstrates an alignment of their  $\Delta G_{\text{exf}}$  with the more effective component (NMP) which has lower  $\Delta G_{\text{exf}}$ . On the other end of the spectrum,  $\Delta G_{\text{exf}}$  of the MET:DCM system in Figure 5.4c closely mirrors that of the less efficient component (MET) with higher  $\Delta G_{\text{exf}}$ . These behaviors confirm the findings from ML modeling that exfoliation efficacy of binary mixture can go beyond mere arithmetic averages of constituent solvents and be dominated by one solvent.

A natural question arises from the above observations: what leads to the dominance of one solvent in a 1:1 binary mixture? Building on insights from our prior research (Chapter 3 and 4) [10,11], we first examine  $\Delta G_{\text{sol}}$  of the solvents in Figure 5.4b and c, since higher  $\Delta G_{\text{sol}}$  typically

indicate improved performance in LPE of g-C<sub>3</sub>N<sub>4</sub> [10]. Figure 5.5 shows the step-by-step change in solvation free energy as the vdW and electrostatic interactions are turned on in the BAR method. In Figure 5.5a,  $\Delta G_{\text{sol}}$  for NMP, CH and their mixture are respectively 213.8, 138.6 and 203.4 kJ/mol·nm<sup>2</sup>. The similar values of  $\Delta G_{\text{sol}}$  for NMP and NMP:CH mixture are consistent with the observation in Figure 5.4b where the PMF of NMP:CH mixture follows closely that of NMP. CH has much smaller  $\Delta G_{\text{sol}}$ , indicating weaker nanosheet-solvent interaction. The distinction between CH, NMP and their mixture becomes particularly pronounced upon activating the electrostatic interaction: CH's non-polar nature results in negligible changes, whereas NMP and the mixture exhibit substantial shifts owing to NMP's polarity. Unlike NMP and CH which have contrasting polarity, both MET and DCM are polar solvents, and the difference between their  $\Delta G_{\text{sol}}$  is notably smaller (Figure 5.5b). However,  $\Delta G_{\text{sol}}$  of the MET:DCM mixture (131.5 kJ/mol·nm<sup>2</sup>) still aligns better with that of MET (130.5 kJ/mol·nm<sup>2</sup>) than DCM (136.3 kJ/mol·nm<sup>2</sup>), consistent with the PMF agreement observed in Figure 5.4c. DCM has slightly larger  $\Delta G_{\text{sol}}$  suggestive of stronger nanosheet-solvent interaction. This is reflected in the greater magnitude of vdW interaction compared to MET (the first 10 steps in Figure 5.5b), despite smaller magnitude of electrostatic interaction (the last 10 steps in Figure 5.5b).

Figure 5.4 and Figure 5.5 demonstrate that the dominance of one solvent in  $\Delta G_{\text{exf}}$  of the mixture is positively correlated with the agreement between  $\Delta G_{\text{sol}}$  of the mixture and that of the dominant solvent. However, these data alone do not explain why the binary mixture would be skewed towards the good solvent (NMP, lower  $\Delta G_{\text{exf}}$  and higher  $\Delta G_{\text{sol}}$ ) for the NMP:CH system while toward the bad solvent (MET, higher  $\Delta G_{\text{exf}}$  and lower  $\Delta G_{\text{sol}}$ ) for the MET:DCM system. Nor does it explain why DCM, which exhibits only a slightly stronger nanosheet-solvent interaction seen in  $\Delta G_{\text{sol}}$ , possess a considerably lower  $\Delta G_{\text{exf}}$  than MET and MET:DCM mixture. Detailed atomistic analyses are conducted next to unravel the mechanisms driving the diverse spectrum of exfoliation behaviors observed among the solvent mixtures.

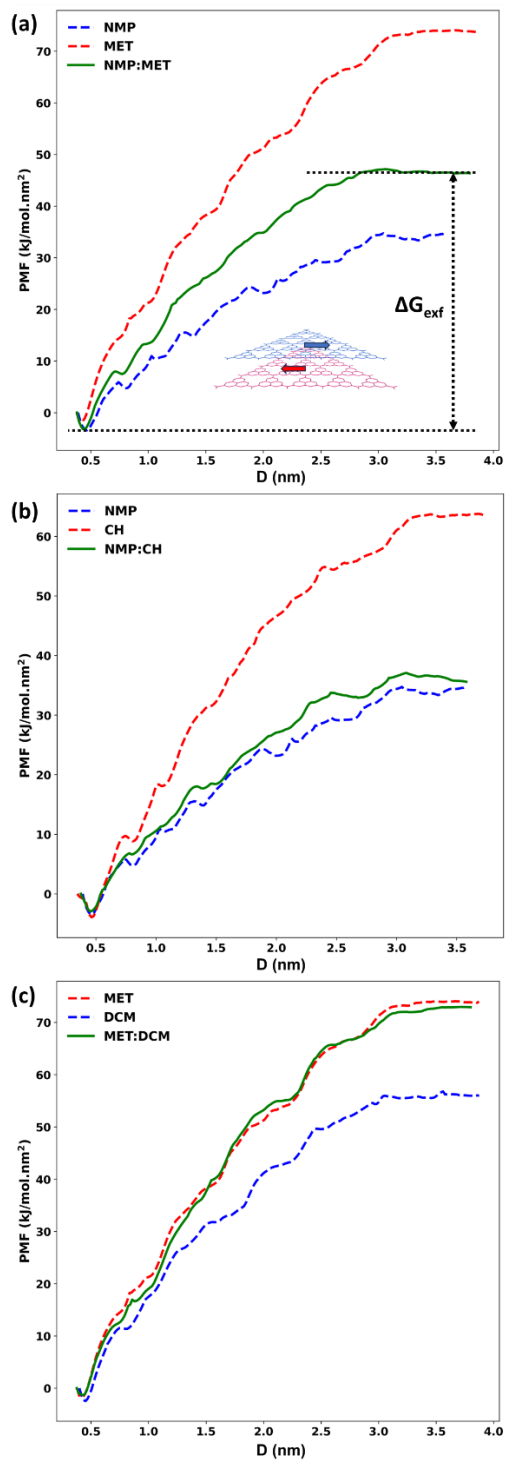


Figure 5.4. PMF profiles for different solvents and their mixture in LPE of g-C<sub>3</sub>N<sub>4</sub>: (a) NMP, MET, and their mixture, (b) NMP, CH, and their mixture, and (c) MET, DCM, and their mixture. The definitions of reaction coordinate and  $\Delta G_{\text{ext}}$  are shown in (a).

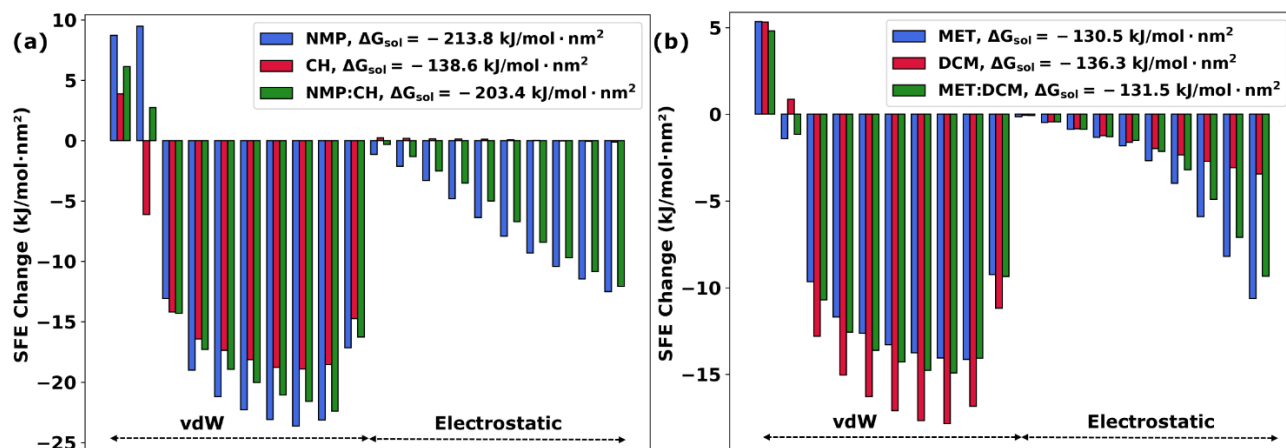


Figure 5.5. Comparison of change in solvation free energy (SFE) as vdW and electrostatic interactions are turned on in twenty steps; (a) NMP, CH, and their mixture; (b) MET, DCM, and their mixture.

### 5.3.3. Dominance of NMP in NMP:CH mixture

Inspired by a previous finding that the mobility of solvent molecules around the nanosheets plays a critical role in LPE [10], equilibrium simulations for two stacked nanosheets are used to quantify the solvent mobility ( $M_{\text{ave}}$ ). Here,  $M_{\text{ave}}$  is defined as the average fraction of solvent molecules that leave the first solvation layer for no more than 0.5 ns; larger  $M_{\text{ave}}$  thus corresponds to higher solvent mobility. The definitions of first, second and third solvation layers are provided in Appendix C Section C2, with the first solvation layer particularly important to the LPE process due to the significant solvent-sheet interactions in this region. The first 5 columns in Figure 5.6 present  $M_{\text{ave}}$  for pure NMP, pure CH, NMP in the binary mixture, CH in the binary mixture, and both NMP and CH in the mixture. Pure NMP stands out with the lowest  $M_{\text{ave}}$  at 35%, suggesting a stable adsorption layer. This low mobility implies high affinity between NMP molecules and the nanosheets, which according to our prior studies correlates with enhanced LPE efficiency [10]. Similar observation was also reported by Gupta et al. [23] in their combined experimental and computational investigation. In stark contrast, pure CH exhibits an  $M_{\text{ave}}$  of 66%, indicating a more transient interaction with the nanosheets and a reduced propensity for aiding the exfoliation process. When combined into the NMP:CH mixture, the mobility of NMP increases slightly to 38% while simultaneously the mobility of CH decreases to 55%. The overall  $M_{\text{ave}}$  is 41% for the



mixture, a number that is more closely aligned with NMP rather than being intermediate between NMP and CH, which would have been the expected outcome if both solvents contributed equally to the dynamics of the mixture in the vicinity of the nanosheets. This confirms the dominance of NMP in the behavior of the binary solvent near the nanosheets, consistent with the observations made in Figure 5.4.

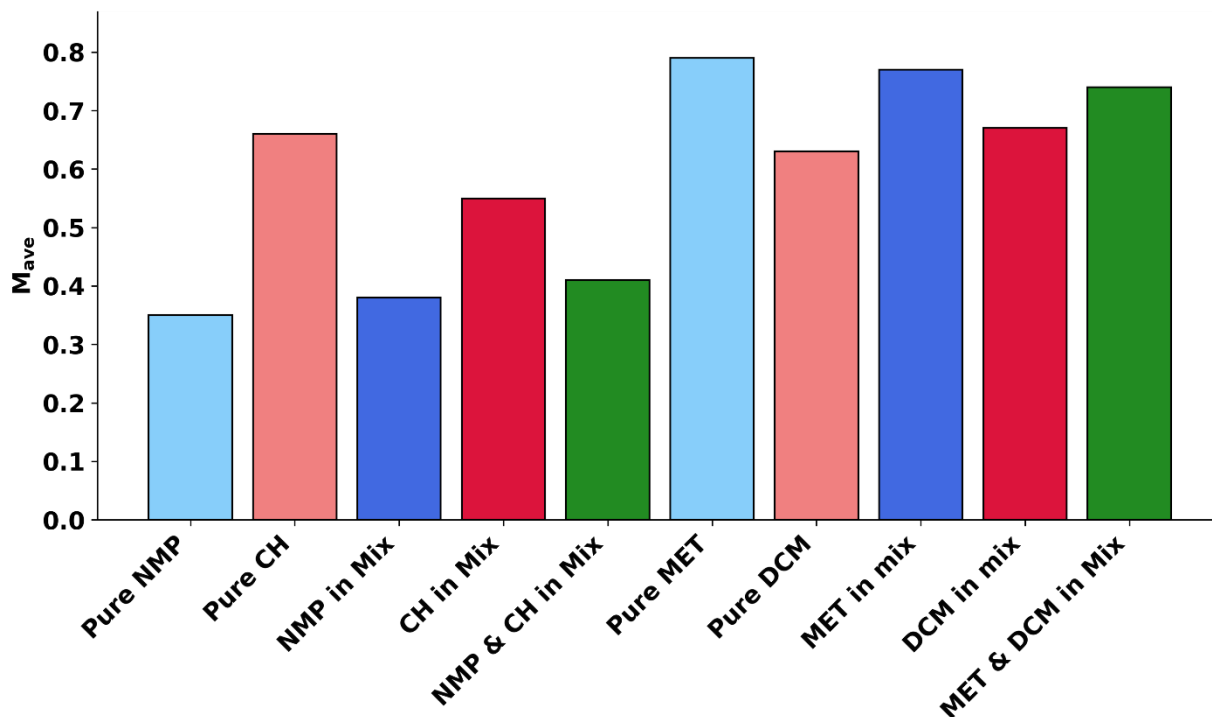


Figure 5.6.  $M_{ave}$  of NMP, CH, MET and DCM in the binary mixtures and their singular counterparts. “NMP & CH in Mix” considers both NMP and CH molecules in the first solvation layer; the same applies to “MET & DCM in Mix”.

Using the  $M_{ave}$  data in Figure 5.6, we can compute the composition of the two types of solvents within the first solvation layer, according to:

$$M_{ave,NMP} \times n_{NMP} + M_{ave,CH} \times n_{CH} = M_{ave,mix} \times (n_{NMP} + n_{CH}) \quad (5.4)$$

Or in another form:

$$\frac{n_{NMP}}{n_{CH}} = \frac{M_{ave,CH} - M_{ave,mix}}{M_{ave,mix} - M_{ave,NMP}} \quad (5.5)$$

where  $n_i$  stands for the average number of molecules of solvent  $i$  within the first solvation layer. Substituting in  $M_{ave,NMP} = 38\%$ ,  $M_{ave,CH} = 55\%$  and  $M_{ave,mix} = 41\%$ , the ratio  $n_{NMP}:n_{CH}$  is found to be 4.66. This implies that over 80% of the solvent molecules in the first solvation layer are NMP, despite the initial ratio of  $n_{NMP}:n_{CH} = 1.1:1$ . This theoretical prediction is substantiated by the direct count from the simulation data, as shown in Figure 5.7a. The graphical representation confirms the prevalence of NMP over CH in the immediate vicinity of the nanosheets. As we move to the second solvation layer, this NMP prevalence halves, and inside the third solvation layer, CH rises to a slightly larger percentage than NMP. This layered arrangement of solvent molecules is further elucidated by the density profiles in Figure 5.7b. Here, the regions of high NMP concentration correspond to lower CH presence than in the bulk solution, illustrating an active displacement of CH by NMP.

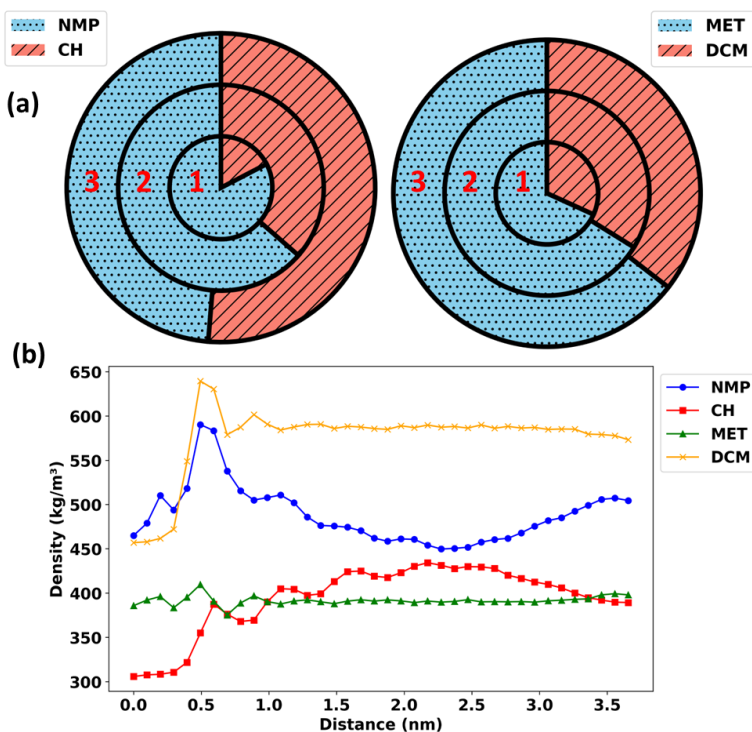


Figure 5.7. (a) Schematic illustration of the average percentage of NMP, CH, MET and DCM molecules in the first three solvation layers of two stacked nanosheets. (b) The corresponding density profiles of NMP, CH, MET, and DCM in their corresponding mixtures. Here, "Distance" is measured from the midpoint between the two nanosheets along the direction perpendicular to the nanosheet plane. All solvent molecules in the systems are considered in the calculation. Each density profile is symmetric about Distance = 0 and only half of the profile is shown.

In addition to being displaced from the vicinity of the nanosheets, CH molecules also undergo orientation changes that affect their affinity with the nanosheets. The molecular orientation is quantified by calculating the angle between the normal of the aromatic plane in CH (and NMP) and that of the nanosheet (see inset of Figure 5.8 for illustration). An angle of 90 degrees indicates a solvent molecule perpendicular to the nanosheets, while angles of 0 or 180 degrees depict parallel alignment. Figure 5.8 shows that the orientation of NMP molecules is remarkably consistent within the first solvation layer, regardless of whether they are in the singular or binary solvent. This suggests a strong, potentially directional, affinity towards the g-C<sub>3</sub>N<sub>4</sub> surfaces that is robust against solvent compositional changes. On the other hand, CH shows considerable reorientation in the mixture, moving to a more perpendicular stance relative to the nanosheets. Caused by the presence of NMP, this reorientation reduces the contact area between each CH molecule and the nanosheet, further weakening the influence of CH in the LPE of g-C<sub>3</sub>N<sub>4</sub> by the binary mixture.

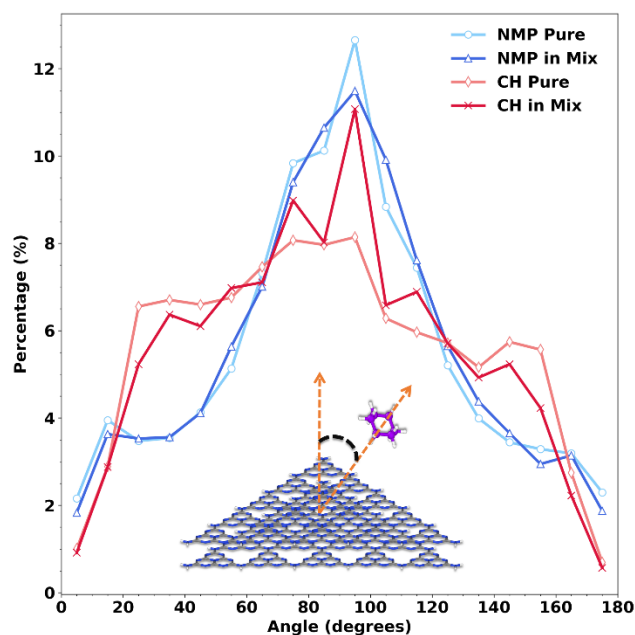


Figure 5.8. Probability distribution for the orientation of solvent molecules in the first solvation layer relative to the nanosheets.

So far, it has been established that NMP's dominance in the mixture stems from its greater adherence to the g-C<sub>3</sub>N<sub>4</sub> nanosheets and lower mobility in the first solvation layer, which

influences the mixture to behave more akin to NMP rather than CH. The next question is what drives the NMP molecules to preferentially attract to the g-C<sub>3</sub>N<sub>4</sub> nanosheets, and what specific interactions are at play? To address this, we turn to the radial distribution function (RDF), which elucidates the subtleties of intermolecular interactions. As Figure 5.9a illustrates, from the singular NMP system to the mixture the RDF profile of the aromatic ring in NMP around the nitrogen atoms in the g-C<sub>3</sub>N<sub>4</sub> nanosheets is notably elevated (upward blue arrow). Correspondingly, presence of CH around the nanosheets is suppressed in the mixture (downward red arrow). For both NMP and CH, the first RDF peaks of their aromatic rings are at approximately the same distance (~0.38 nm) from the g-C<sub>3</sub>N<sub>4</sub> nitrogen. However, NMP uniquely approaches the nanosheets from its oxygen end, evidenced by the first peak in the RDF of NMP oxygen being located as close as 0.27 nm to the g-C<sub>3</sub>N<sub>4</sub> nitrogen (green curve in Figure 5.9a). This peak corresponds to the formation of H-bonds between NMP oxygen and the hydrogens of the nanosheets on the periphery (see the definition of periphery and interior regions of a nanosheet in Figure 5.1). The longevity of these H-bond is substantial (see quantification of H-bond lifetime in Appendix C, Section C3), with NMP in the mixture showing a H-bond lifetime of 33.8 ps—a slight decrease from 43.4 ps in pure NMP but still significant enough to foster a persistent interaction conducive to LPE.

In addition to the first and sharpest peak at 0.27 nm, the RDF of NMP oxygen in Figure 5.9a also shows a pronounced peak at around 0.35 nm. This peak corresponds to another type of interaction NMP exploits with the nanosheets via its oxygen atoms that are situated near the aromatic structures of the nanosheets. The so-called “lone-pair  $\pi$  interaction”, akin to those found in water-protein studies, involves the electron-rich lone pairs of oxygen interacting with the electron-poor  $\pi$  systems of aromatic rings [24–26]. Both types of non-covalent interactions promote the attraction of NMP to g-C<sub>3</sub>N<sub>4</sub> and its stability around the nanosheets, while CH’s lack of ability to establish H-bond or lone-pair  $\pi$  interaction, due to its non-polar nature, leads to its depletion from the nanosheets.

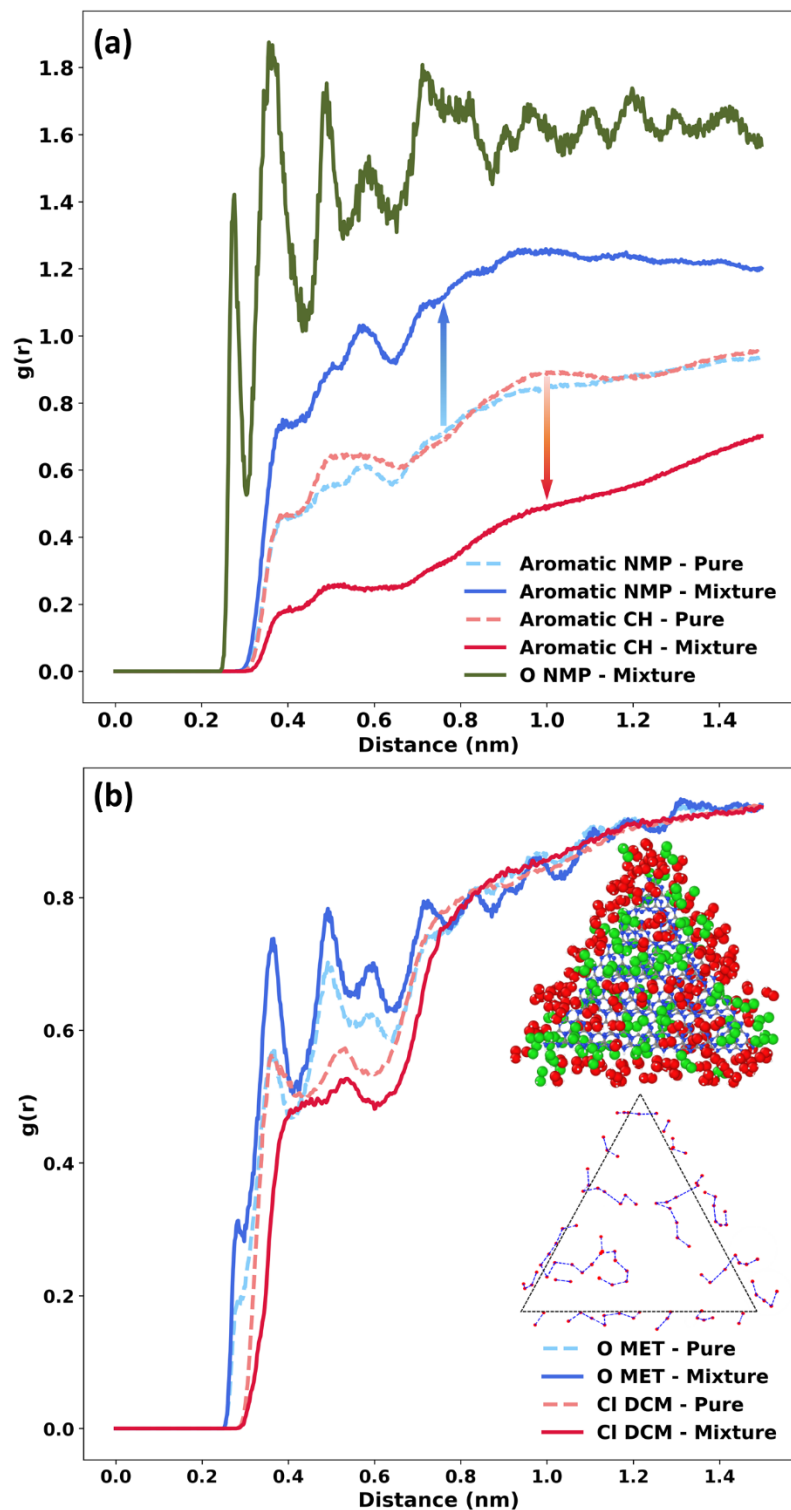


Figure 5.9. (a) RDF plots showing the distributions of aromatic rings in NMP and CH around nitrogen in the two stacked  $g\text{-C}_3\text{N}_4$  nanosheets, in both singular and binary mixture systems. The distribution of oxygen in NMP around nitrogen in  $g\text{-C}_3\text{N}_4$  nanosheets is also shown for the binary

mixture system; the corresponding result for pure NMP is very similar and hence not shown. (b) RDF plots showing the distributions of oxygen in MET and chlorine in DCM around nitrogen atoms of the two stacked g-C<sub>3</sub>N<sub>4</sub> nanosheets, in both singular and binary systems. The top inset shows a snapshot of the first solvation layer at  $t = 20$  ns, with MET and DCM molecules in red and green, respectively. The H-bond network between MET molecules is shown with blue lines in the bottom inset.

#### 5.3.4. Dominance of MET in MET:DCM mixture

Both  $\Delta G_{\text{exf}}$  and  $\Delta G_{\text{sol}}$  data, in Figure 5.4 and Figure 5.5 respectively, show that the behavior of the MET:DCM mixture is closer to its MET component. Examination of the solvent mobility data (last 5 columns in Figure 5.6) reveals that in their pure form, MET has higher mobility compared to DCM, a trend that persists even when MET is mixed with DCM. The resulting mixture exhibits mobility similar to its MET component, mirroring the alignment seen between the NMP:CH mixture and its NMP component. Extending this reasoning to the number of solvent molecules within the first solvation layer, we anticipate a similar pattern of dominance. Based on Eq. (5.5) and  $M_{\text{ave}}$  data in Figure 5.6, the ratio  $n_{\text{MET}}:n_{\text{DCM}}$  in the first solvation layer is approximately 2.3:1, despite the initial ratio of 1.8:1. Direct counts of solvent molecules confirm this prediction (Figure 5.7a). Examining Figure 5.7b, the dominance of MET is also evident in the density profiles, where near the nanosheets the density of MET exceeds its bulk value, while DCM has a significantly lower density compared to the bulk.

Despite the similarity to the NMP:CH mixture, where the behavior of the binary solvent is dominated by the component with stronger presence around the nanosheets, a critical difference exists: here, the solvent with higher mobility dominates the mixture. The dominance of MET in the MET:DCM mixture therefore results in poorer LPE performance, converging to that of the “bad” solvent (MET), opposite to the case of the NMP:CH mixture.

A natural question arises: why does MET dominate the first solvation layer (in fact, the first three solvation layers according to Figure 5.7a) even though it has higher mobility? To answer this, we must delve into the solvent-sheet interactions. Figure 5.9b shows the RDF curves of oxygen in MET and chlorine in DCM around nitrogen atoms of the nanosheets. For pure MET, the RDF of its oxygen shows a small peak at approximately 0.28 nm, indicating H-bonding with the nanosheets’ periphery. DCM, on the other hand, cannot form H-bonds due to the absence of highly electronegative atoms such as nitrogen, oxygen, or fluorine. Consequently, there is no peak

at this distance in the RDF for pure DCM. Subsequent peaks around 0.36 nm for both pure solvents suggest interactions of oxygen atoms in MET and chlorine atoms in DCM with the  $\pi$  systems of the nanosheets. Cl- $\pi$  interaction, induced by chlorine's high polarizability, generates strong dispersion forces, with interaction energies ranging from -2.01 to -3.60 kcal/mol [27–29]. In comparison, O- $\pi$  interaction between MET and the nanosheets is generally weaker, with energies between -0.11 and -2.12 kcal/mol [29–31]. The pronounced strength of interaction between DCM and the interior of the nanosheet offsets DCM's lack of H-bonding capability and explains why DCM has a considerably lower  $\Delta G_{\text{exf}}$  than MET and the MET:DCM mixture, even though its  $\Delta G_{\text{sol}}$  (governed by interactions in both interior and periphery of the nanosheet) is of similar magnitude.

When MET and DCM are mixed, the RDF curves experience notable changes. As shown in Figure 5.9b, the RDF of the dominant solvent (MET here) is elevated, while that of the less important solvent (DCM here) is diminished at short distances. Specifically, the distinct peak at 0.36 nm in the singular DCM system, indicative of Cl- $\pi$  interactions, disappears in the mixture. Concurrently, the first peak in the RDF of pure MET (at 0.28 nm) becomes more pronounced in the mixture. These results suggest that DCM's Cl atoms are displaced to a larger distance from the nanosheets, while MET's H-bond formation with the nanosheets' periphery is enhanced. Consequently, we expect an accumulation of MET molecules in the periphery of the nanosheets. To verify this, solvent molecules in the first solvation layer are categorized into periphery and interior regions based on whether their COMs fall within 0.50 nm (for MET) or 0.53 nm (for DCM) from the respective regions of the nanosheets. These distances correspond to the first prominent peaks observed in the RDF curves, as detailed in Appendix, section C2. The results reveal that 67% of the MET molecules in the first solvation layer are situated in the periphery, compared to only 47% for DCM. Consequently, MET constitutes 77% of the total molecules in the periphery region. The top inset of Figure 5.9b provides a snapshot of the first solvation layer, clearly illustrating the dominance of MET, especially in the periphery.

In addition, a MET molecule can form H-bonds with other MET molecules. As shown in the bottom inset of Figure 5.9b, which highlights the H-bond network of MET with itself (rather than with the nanosheets), MET's H-bond network begins at the periphery and extends into the interior. This H-bond network contributes to the dominant occupation of the solvation layers by the MET molecules, not only in the periphery but also in the interior. MET molecules remain highly mobile due to their small MW, thus the H-bond network is dynamic and there are frequent

exchanges between MET molecules in the first solvation layer and outside. However, the H-bond network is stable with balanced H-bond breaking and reforming, ensuring the dominant presence of MET in the vicinity of the nanosheets. This is why Figure 5.7a shows prevalent occupation of all three solvation layers by MET, in contrast to NMP which, being higher immobile, becomes less dominant in the second or third solvation layers.

### 5.3.5. Application: eco-friendly LPE with water-organic solvent mixtures

Exploring cost-effective and environmentally-friendly methods for the LPE of g-C<sub>3</sub>N<sub>4</sub>, a natural option is to consider water as a solvent. However, water is known to be a poor performer for g-C<sub>3</sub>N<sub>4</sub> LPE, with  $\Delta G_{\text{exf}}$  reaching 119.5 kJ/mol·nm<sup>2</sup> [10], the highest among all solvents that we have simulated. With the discovery of dominance of one solvent in a binary mixture, it is of interest to see if water can be added to an organic solvent with only minor deterioration of the LPE performance. To achieve this, we conducted a series of MD simulation, each 20 ns long and contained a single nanosheet equilibrated in 25 common organic solvents (listed in Appendix C, Section C4)— either in pure form or mixed with water in a volumetric ratio of 1:3, where water constitutes 25% and the organic solvent 75%. These organic solvents are commonly used in the LPE of layered nanomaterials and are miscible with water. From the insights gained earlier, we assess the efficacy of the binary mixture based on a ratio ( $\beta$ ) that quantifies the dominance of the organic solvent molecular near the nanosheet.

$$\beta = \frac{\left(\frac{n_{OS}}{n_W}\right)_{\text{First layer}}}{\left(\frac{n_{OS}}{n_W}\right)_{\text{Bulk}}} \quad (5.6)$$

where  $n_{OS}$  and  $n_W$  are the number of organic solvent and water molecules, respectively, and the subscripts indicate whether the measurement is done in the first solvation layer or for the bulk. A larger  $\beta$  indicates greater dominance of the organic solvent in the first solvation layer.

Our findings indicate that none of the systems demonstrates water as the dominant solvent, suggesting that mixing organic solvents with water consistently exhibits above-par performance. Notably, three organic solvents—IPA, MSA, and EtOH—emerge as candidates potentially dominating their mixtures with water. PMF results (Figure 5.10a) further show the extent of their dominance. In particular, when mixed with water IPA is found to significantly reduce the  $\Delta G_{\text{exf}}$



value towards itself, with  $\Delta G_{\text{exf}}$  of the mixture only 16.3% higher than that of pure IPA. RDF of IPA oxygen around nitrogen atoms of the nanosheets (Figure 5.10b) shows that compared to pure IPA, the lone-pair  $\pi$  interactions in IPA:water mixture not only occur at shorter distances (moving from 0.38 nm to 0.35 nm) but also have greater intensities. Snapshot of the first solvation layer (inset of Figure 5.10b) also reveals that IPA molecules dominate both the interior and periphery of the nanosheet, displacing water molecules out of the first solvation layer.

Our results align well with experimental studies on various 2D layered nanomaterials, which commonly report the highest nanosheet concentrations using IPA-water mixtures [8,9,32] or ethanol-water mixtures [7,33–35]. Additionally, our results recommend MSA as another solvent to be considered for mixing with water. Adopting water-inclusive solvent systems offers a strategy for scaling up nanomaterial production, reducing reliance on costly solvents, and aligning with environmental sustainability goals. Consequently, this approach promises broader adoption of g- $\text{C}_3\text{N}_4$  nanosheets in diverse applications, from energy storage to environmental remediation.

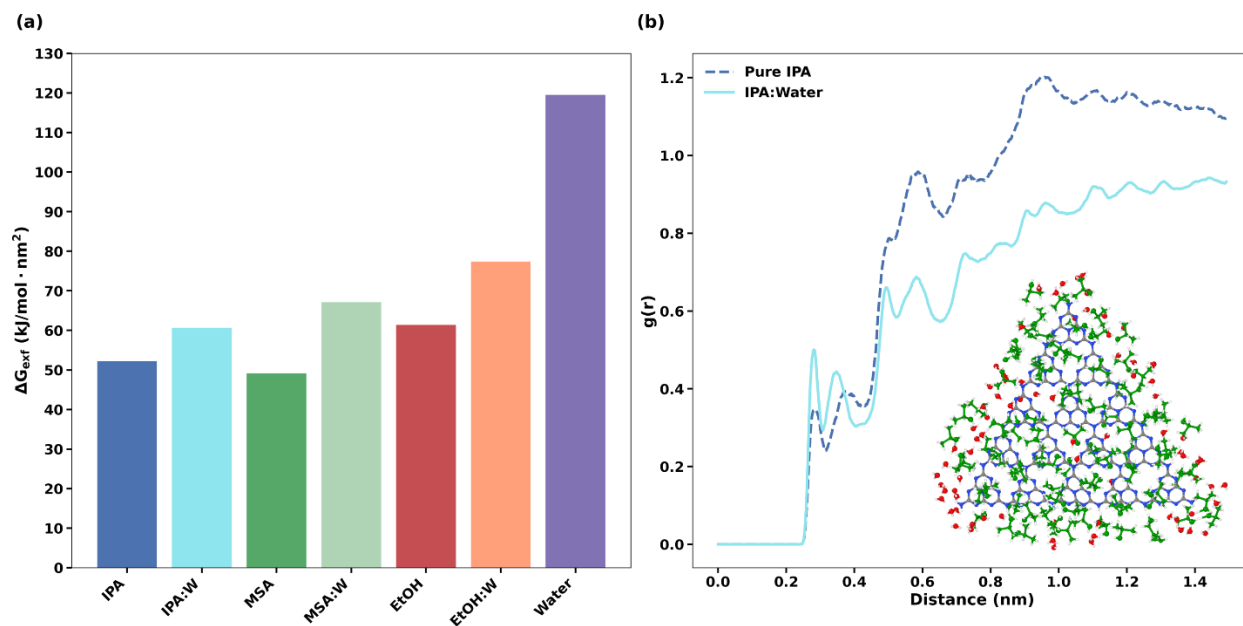


Figure 5.10. (a)  $\Delta G_{\text{exf}}$  for three organic solvents and their mixture with water.  $\Delta G_{\text{exf}}$  for pure water is also shown for comparison. (b) RDF of IPA oxygen around nitrogen atoms of the nanosheet, in pure IPA and in IPA:water mixture. Inset show a snapshot of the first solvation layer at  $t = 20$  ns where the IPA and water molecules are shown with green and red color, respectively.

## 5.4. Discussion

ML modeling and MD simulations reveal that when two solvents are combined to create a binary mixture, their  $\Delta G_{\text{exf}}$  for g-C<sub>3</sub>N<sub>4</sub> nanosheets can significantly deviate from the arithmetic average of  $\Delta G_{\text{exf}}$  of the individual components. The derivation is stronger when the arithmetic average is very large or very small. In these situations, one solvent is often found to have prevalent presence in the vicinity of the nanosheets and consequently dominate the LPE performance of the binary mixture. This can skew  $\Delta G_{\text{exf}}$  of the binary mixture towards the “good” solvent (if the dominant solvent has low  $\Delta G_{\text{exf}}$ , for example NMP in the NMP:CH mixture) or the “bad” solvent (if the dominant solvent has high  $\Delta G_{\text{exf}}$ , for example MET in the MET:DCM mixture).

The correlation between  $\Delta G_{\text{exf}}$  and solvent mobility, identified in our previous work [10], is reaffirmed in this study: higher solvent mobility in the first solvation layer leads to higher  $\Delta G_{\text{exf}}$  and poorer LPE performance. On the other hand, the relationship between solvent mobility and dominance near the nanosheets is more intricate. In the case of the NMP:CH mixture, NMP, being a more polar solvent, demonstrates stronger anchoring to the nanosheets through hydrogen bonding and lone-pair  $\pi$  interactions, both of which are absent in the non-polar CH. This strong anchoring, coupled with the higher MW of NMP, results in lower mobility for NMP, thereby establishing its dominance in the mixture. In contrast, for the MET:DCM mixture, both solvents are polar. MET shows anchoring to both periphery and interior of the nanosheets. However, this is counterbalanced by MET's lower MW. DCM, while only anchoring to the interior, has a higher MW, resulting in overall lower mobility. Despite this, the mobility difference between DCM and MET is not as pronounced as that between NMP and CH. The ability of MET to form a dynamic yet stable H-bond network, extending from the periphery to the interior, enables the prevalence of MET in the solvation layers around the nanosheets.

Our results underscore that molecular polarity, and similarly surface tension, alone does not dictate solvent-nanosheet interactions or solvent mobility. Halim et al. [33] provided insights into this aspect by demonstrating that larger molecule size of the cosolvent significantly enhances exfoliation yield through steric repulsion, which prevents the restacking of exfoliated layers. Their study focused on solvent systems composed of water mixed with different alcohols—MET, ethanol, IPA, and t-butyl alcohol. Their findings are in line with ours where solvents with higher MW usually having lower mobility and hence lower  $\Delta G_{\text{exf}}$ . Water, for instance, has high polarity

but also high mobility due to its small MW, and exhibits poor LPE performance. Results from this study therefore challenge existing, largely surface tension based, solvent selection criteria for LPE, and emphasize the importance of molecular details in the solvation layers around the nanosheets.

## 5.5. Conclusion

This study elucidates the intricate molecular mechanics governing the performance of binary solvent mixtures in the liquid-phase exfoliation (LPE) of graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) nanosheets. By integrating molecular dynamics simulations with machine learning predictions, we demonstrate that the performance of binary solvent systems can be significantly influenced by the dominance of one solvent, which can lead to above-par or sub-par LPE performance. This dominance is dictated by the ability of certain solvents to occupy the first solvation layer, their interaction strength with g-C<sub>3</sub>N<sub>4</sub> nanosheets, and the resulting molecular orientation and mobility. Insights from this study suggests a novel strategy for optimizing solvent mixtures, and particularly the potential for employing more environmentally friendly and cost-effective solvent systems, such as water-organic solvent mixtures, without compromising exfoliation performance.

## 5.6. Acknowledgement

Digital Research Alliance of Canada is acknowledged for providing the computing resources and technical support. TT acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Grant number: RGPIN-2018-04281) and the Canada Research Chair program (CRC; Grant number: TIER1 2021-00023). KS thanks NSERC for financial support (Grant number: RGPIN-2020-04620). ES acknowledges support from Alberta Innovates Graduate Student Scholarship and Myer Horowitz Graduate Students' Association Graduate Scholarship.

## 5.7. References

- [1] J. Fu, J. Yu, C. Jiang, B. Cheng, g-C<sub>3</sub>N<sub>4</sub>-Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.

- [2] X. Ma, J. Hu, H. He, S. Dong, C. Huang, X. Chen, New Understanding on Enhanced Photocatalytic Activity of g-C<sub>3</sub>N<sub>4</sub>/BiPO<sub>4</sub> Heterojunctions by Effective Interfacial Coupling, *ACS Appl Nano Mater* 1 (2018) 5507–5515. <https://doi.org/10.1021/acsanm.8b01012>.
- [3] S. Patnaik, D.P. Sahoo, K. Parida, An overview on Ag modified g-C<sub>3</sub>N<sub>4</sub> based nanostructured materials for energy and environmental applications, *Renewable and Sustainable Energy Reviews* 82 (2018) 1297–1312. <https://doi.org/10.1016/J.RSER.2017.09.026>.
- [4] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.
- [5] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268. [https://doi.org/10.1021/ACSNANO.5B02683/ASSET/IMAGES/MEDIUM/NN-2015-026836\\_0009.GIF](https://doi.org/10.1021/ACSNANO.5B02683/ASSET/IMAGES/MEDIUM/NN-2015-026836_0009.GIF).
- [6] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648. <https://doi.org/10.1021/ja1064284>.
- [7] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [8] J. Shen, Y. He, J. Wu, C. Gao, K. Keyshar, X. Zhang, Y. Yang, M. Ye, R. Vajtai, J. Lou, P.M. Ajayan, Liquid Phase Exfoliation of Two-Dimensional Materials by Directly Probing and Matching Surface Tension Components, *Nano Lett* 15 (2015) 5449–5454. [https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E\\_0005.JPEG](https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E_0005.JPEG).
- [9] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight from molecular

dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954.

<https://doi.org/10.1007/s10934-021-01047-7>.

- [10] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910.  
<https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [11] E. Shahini, N. Chaulagain, K. Shankar, T. Tang, Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning, *ACS Appl Mater Interfaces* 15 (2023) 53786–53801.  
[https://doi.org/10.1021/ACSAMI.3C09347/SUPPL\\_FILE/AM3C09347\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.3C09347/SUPPL_FILE/AM3C09347_SI_001.PDF).
- [12] W.K. Darkwah, Y. Ao, Mini Review on the Structure and Properties (Photocatalysis), and Preparation Techniques of Graphitic Carbon Nitride Nano-Based Particle, and Its Applications, *Nanoscale Research Letters* 2018 13:1 13 (2018) 1–15. <https://doi.org/10.1186/S11671-018-2702-3>.
- [13] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon N Y* 141 (2019) 580–607.  
<https://doi.org/10.1016/J.CARBON.2018.09.082>.
- [14] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [15] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017).  
<https://doi.org/10.1093/nar/gkx312>.
- [16] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a.  
<https://doi.org/10.2477/jccjie.2018-0034>.
- [17] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate

- quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).
- [18] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [19] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [20] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [21] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).
- [22] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [23] A. Gupta, V. Arunachalam, S. Vasudevan, Liquid-Phase Exfoliation of MoS<sub>2</sub> Nanosheets: The Critical Role of Trace Water, *Journal of Physical Chemistry Letters* 7 (2016) 4884–4890. [https://doi.org/10.1021/ACS.JPCLETT.6B02405/ASSET/IMAGES/LARGE/JZ-2016-02405F\\_0005.JPEG](https://doi.org/10.1021/ACS.JPCLETT.6B02405/ASSET/IMAGES/LARGE/JZ-2016-02405F_0005.JPEG).
- [24] G.A. DiLabio, E.R. Johnson, Lone pair- $\pi$  and  $\pi$ - $\pi$  interactions play an important role in proton-coupled electron transfer reactions, *J Am Chem Soc* 129 (2007) 6199–6203. [https://doi.org/10.1021/JA068090G/SUPPL\\_FILE/JA068090GSI20070312\\_120206.PDF](https://doi.org/10.1021/JA068090G/SUPPL_FILE/JA068090GSI20070312_120206.PDF).
- [25] T.J. Mooibroek, P. Gamez, J. Reedijk, Lone pair- $\pi$  interactions: a new supramolecular bond?, *CrystEngComm* 10 (2008) 1501–1515. <https://doi.org/10.1039/B812026A>.
- [26] A. Jain, V. Ramanathan, R. Sankararamakrishnan, Lone pair  $\cdots \pi$  interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds, *Protein Sci* 18 (2009) 595. <https://doi.org/10.1002/PRO.67>.

- [27] Y.N. Imai, Y. Inoue, I. Nakanishi, K. Kitaura, Cl- $\pi$  interactions in protein-ligand complexes, *Protein Science* 17 (2008) 1129–1137. <https://doi.org/10.1110/PS.033910.107>.
- [28] C. Amsaraj, R. Bharathikannan, P. Muthuraja, M. Rajkumar, Chlorine directed C–Cl ...  $\pi$  and C–H...Cl interactions in acridinium 3,5-dichlorosalicylate: Synthesis, X-ray diffraction and theoretical analysis, *J Mol Struct* 1220 (2020) 128759. <https://doi.org/10.1016/J.MOLSTRUC.2020.128759>.
- [29] K.E. Riley, K.A. Tran, Strength and Character of R–X... $\pi$  Interactions Involving Aromatic Amino Acid Sidechains in Protein-Ligand Complexes Derived from Crystal Structures in the Protein Data Bank, *Crystals* 2017, Vol. 7, Page 273 7 (2017) 273. <https://doi.org/10.3390/CRYST7090273>.
- [30] A. Jain, C.S. Purohit, S. Verma, R. Sankararamakrishnan, Close contacts between carbonyl oxygen atoms and aromatic centers in protein structures:  $\pi$ ... $\pi$  or lone-pair... $\pi$  interactions?, *Journal of Physical Chemistry B* 111 (2007) 8680–8683. [https://doi.org/10.1021/JP072742L/SUPPL\\_FILE/JP072742LSI20070604\\_010130.PDF](https://doi.org/10.1021/JP072742L/SUPPL_FILE/JP072742LSI20070604_010130.PDF).
- [31] T. Korenaga, H. Tanaka, T. Ema, T. Sakai, Intermolecular oxygen atom... $\pi$  interaction in the crystal packing of chiral amino alcohol bearing a pentafluorophenyl group, *J Fluor Chem* 122 (2003) 201–205. [https://doi.org/10.1016/S0022-1139\(03\)00089-7](https://doi.org/10.1016/S0022-1139(03)00089-7).
- [32] U. Halim, C.R. Zheng, Y. Chen, Z. Lin, S. Jiang, R. Cheng, Y. Huang, X. Duan, A rational design of cosolvent exfoliation of layered materials by directly probing liquid–solid interaction, *Nature Communications* 2013 4:1 4 (2013) 1–7. <https://doi.org/10.1038/ncomms3213>.
- [33] U. Halim, C.R. Zheng, Y. Chen, Z. Lin, S. Jiang, R. Cheng, Y. Huang, X. Duan, A rational design of cosolvent exfoliation of layered materials by directly probing liquid–solid interaction, *Nature Communications* 2013 4:1 4 (2013) 1–7. <https://doi.org/10.1038/ncomms3213>.
- [34] K.-G. Zhou, N.-N. Mao, H.-X. Wang, Y. Peng, H.-L. Zhang, K.-G. Zhou, N.-N. Mao, H.-X.H. Wang, -L Zhang, Y. Peng, A Mixed-Solvent Strategy for Efficient Exfoliation of Inorganic Graphene Analogues, *Angewandte Chemie International Edition* 50 (2011) 10839–10842. <https://doi.org/10.1002/ANIE.201105364>.

- [35] X. Gu, Y. Zhao, K. Sun, C.L.Z. Vieira, Z. Jia, C. Cui, Z. Wang, A. Walsh, S. Huang, Method of ultrasound-assisted liquid-phase exfoliation to prepare graphene, *Ultrason Sonochem* 58 (2019) 104630. <https://doi.org/10.1016/J.ULTSONCH.2019.104630>.



# Chapter 6: Molecular Dynamics Analysis on Liquid-Phase Exfoliation of Functionalized Graphitic Carbon Nitride

## 6.1. Introduction

g-C<sub>3</sub>N<sub>4</sub> has emerged as a versatile 2D material with remarkable potential in various applications such as photocatalysis [1], energy conversion [2], and environmental remediation [3]. Its intrinsic properties, including high chemical stability, metal-free composition, and visible-light-responsive photocatalytic activity, position g-C<sub>3</sub>N<sub>4</sub> as a promising candidate for advancing these technologies [4]. Pristine g-C<sub>3</sub>N<sub>4</sub> faces several inherent limitations, such as relatively low specific surface area compared to other 2D materials like graphene, rapid charge recombination, limited light absorption range, and poor dispersibility in both aqueous and organic mediums [5]. Also because of the poor dispersibility, the scalable production of high-quality g-C<sub>3</sub>N<sub>4</sub> nanosheets remains a significant challenge, impeding their widespread industrial application. To overcome these issues, chemical functionalization has emerged as a valuable strategy to modify and enhance the synthesis and properties of g-C<sub>3</sub>N<sub>4</sub>, thereby expanding its range of applications [6].

Functionalization of g-C<sub>3</sub>N<sub>4</sub> with groups such as sulfonic (SO<sub>3</sub>H) [7], carboxyl (COOH) [8], amine (NH<sub>2</sub>) [9], hydroxyl (OH) [10], and aldehyde (CHO) [11] has been shown to significantly enhance various properties of g-C<sub>3</sub>N<sub>4</sub>, including photocatalytic efficiency, charge separation, and chemical stability. These functional groups can introduce additional active sites, improve electronic properties, and increase the hydrophilicity of g-C<sub>3</sub>N<sub>4</sub>, thereby broadening its applicability in diverse fields such as environmental remediation, energy conversion, and sensor technology [6]. For example, Yousefi et al. [10] demonstrated that oxidized g-C<sub>3</sub>N<sub>4</sub> nanosheets are effective adsorbents for organic dyes and tetracycline in water remediation. These oxidized materials showed high adsorption capacities, 70-600 mg/g for organic dyes and 895 mg/g for tetracycline, surpassing non-oxidized counterparts. The enhanced adsorption was attributed to the introduction of oxygen-containing groups, improving interactions with adsorbates and making these materials promising for water purification applications.

Furthermore, functionalization has been shown to lead to better dispersion of g-C<sub>3</sub>N<sub>4</sub> in solvents, which is crucial for processes like LPE [4]. LPE involves dispersing bulk g-C<sub>3</sub>N<sub>4</sub> in a liquid medium followed by applying ultrasonic energy to overcome the van der Waals interactions between layers. The choice of solvent is crucial in this process, as it directly influences the efficiency of exfoliation and the stability of the resulting nanosheets in dispersion [12]. When functionalization is taken into account, finding an optimal solvent becomes even more complex due to the varied interactions between different functional groups and solvents.

The effect of functionalization on LPE for some 2D layered materials, such as graphene, has been extensively studied. For example, Haar et al. [13] explored the impact of functionalization on the LPE of graphite into single- and few-layer graphene using alkanes terminated with various functional groups, including COOH, OH, and NH<sub>2</sub>. Their study revealed that the type of functional group attached to the alkanes significantly influenced the exfoliation yield (YW), defined as the ratio of exfoliated graphene weight to the initial graphite weight. Among the functional groups tested, COOH was particularly effective, leading to a ~100% increase in YW when exfoliated in NMP. This enhanced performance was attributed to the strong dipole–dipole interactions facilitated by the COOH groups, which promoted better stabilization and dispersion of the graphene sheets in the solvent.

LPE of functionalized g-C<sub>3</sub>N<sub>4</sub> has been studied to a much less extent. Kumru et al. [7] investigated the enhancement of dispersibility of g-C<sub>3</sub>N<sub>4</sub> particles in both aqueous and organic media using a one-pot grafting approach. They grafted functionalized olefinic molecules with low polymerization tendencies onto the g-C<sub>3</sub>N<sub>4</sub> surface, which significantly improved its dispersibility while preserving its intrinsic properties. The study highlighted that functionalization with SO<sub>3</sub>H groups significantly enhanced the hydrophilicity of g-C<sub>3</sub>N<sub>4</sub>, resulting in rapid dispersion and increased solid content in water, achieving up to 10% solid content. The introduction of SO<sub>3</sub>H groups also led to more negative zeta potentials, indicating improved dispersion stability. Additionally, the research demonstrated that NH<sub>2</sub> functionalization imparted pH-dependent dispersibility to g-C<sub>3</sub>N<sub>4</sub> in aqueous media. The NH<sub>2</sub>-functionalized g-C<sub>3</sub>N<sub>4</sub> dispersed uniformly under acidic conditions (pH = 4) but precipitated when the pH was increased to a basic range (pH = 9), allowing for reversible dispersion through reacidification.

Chapter 3 [14] has shed light on the intricate mechanisms underlying the LPE of pristine g-C<sub>3</sub>N<sub>4</sub> nanosheets. We demonstrated that the free energy of exfoliation ( $\Delta G_{\text{exf}}$ ), defined as the

energy required to separate a unit area of two stacked nanosheets from their equilibrium distance, effectively quantifies the ease of exfoliation. Using MD-based PMF calculations, we revealed the correlations of  $\Delta G_{\text{exf}}$  with several solvent-nanosheet interaction metrics such as solvent mobility in the first solvation layer, H-bond lifetime, and the orientation of solvent molecules around the nanosheet [14]. Furthermore,  $\Delta G_{\text{exf}}$  is influenced by intensive properties of the solvents, including surface tension, dielectric constant, and density [15]. This comprehensive understanding enabled us to leverage MD data, supplemented with ML analysis, to predict effective solvents for LPE. Experimental validation through dispersibility tests confirmed the accuracy of our ML model. Building on these insights, we now turn our attention to functionalized g-C<sub>3</sub>N<sub>4</sub> nanosheets.

Specifically, we ask the following questions: can the same correlations and predictive capabilities be extended to functionalized variants? How can we use MD simulations to assist in solvent selection for functionalized g-C<sub>3</sub>N<sub>4</sub> nanosheets? These questions underscore the need for systematic investigations to elucidate the molecular-level interactions between functionalized g-C<sub>3</sub>N<sub>4</sub> and solvents. Such investigations can provide valuable guidelines for optimizing the exfoliation process and enhancing the performance of g-C<sub>3</sub>N<sub>4</sub>-based applications.

## 6.2. Methods

### 6.2.1. Structures of functionalized g-C<sub>3</sub>N<sub>4</sub>

The g-C<sub>3</sub>N<sub>4</sub> nanosheet is constructed using tri-s-triazine-based (heptazine) units instead of s-triazine-based counterparts, since experimental and theoretical studies have reported that the heptazine g-C<sub>3</sub>N<sub>4</sub> is more stable than the s-triazine-based g-C<sub>3</sub>N<sub>4</sub> [16,17]. The model features an equilateral triangular sheet composed of 15 heptazine units, each side measuring 4.0 nm, as depicted in Figure 6.1a. This configuration serves as the baseline g-C<sub>3</sub>N<sub>4</sub> model or pristine g-C<sub>3</sub>N<sub>4</sub>. Hydrogen atoms are added to the pristine g-C<sub>3</sub>N<sub>4</sub> to saturate any unbonded nitrogen atoms, ensuring stability of the structure during the simulation.

To investigate the effect of functional groups on LPE of g-C<sub>3</sub>N<sub>4</sub>, several common functional groups from the literature are selected: SO<sub>3</sub>H, COOH, NH<sub>2</sub>, OH, and CHO. The functionalization is carried out by uniformly placing 15 functional groups at the nanosheet periphery (Figure 6.1b-f). Functionalizing the periphery, as opposed to the inert basal planes, significantly enhances the material reactivity and accessibility for modifications, thereby

improving photocatalytic activities through better charge separation and extended light absorption capabilities [18,19]. For the liquid medium, DMF which is a common solvent for the LPE of  $g\text{-C}_3\text{N}_4$  [20] is used for all the systems. We recognize that the size of the nanosheets and the number of functional groups can affect the LPE performance and the overall properties of the nanosheets. For the purpose of this study, these variables are kept constant to focus solely on the effects of the functional groups.

To describe the intra- and intermolecular interactions, the LigParGen and PolyParGen servers [21,22] are used to generate the necessary parameters. These parameters, which conform to the OPLS-AA force field [23], incorporate the Charge Model 5 [24] with a partial atomic charge (PAC) scaling factor of 1.20. The effectiveness and accuracy of these generated parameters have been substantiated in previous research [14,15].

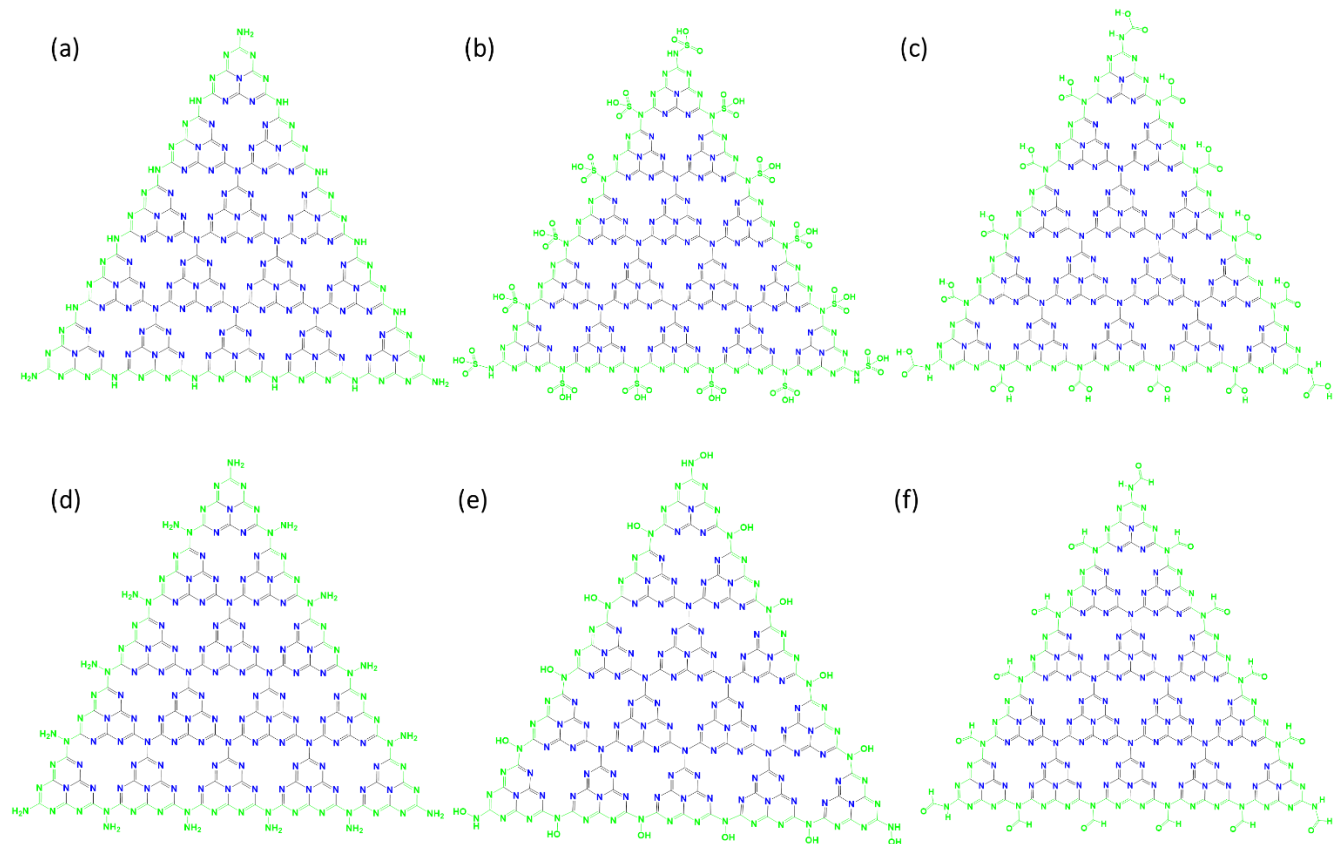


Figure 6.1. The molecular structure of (a) pristine  $g\text{-C}_3\text{N}_4$  and  $g\text{-C}_3\text{N}_4$  functionalized with (b) sulfonic, (c) carboxyl, (d) amine, (e) hydroxyl, and (f) aldehyde. The atoms shown with green color are defined to form the “periphery” of the nanosheet. The remaining atoms are defined to belong to the “interior”.

#### 6.2.1.1. Simulation details

To evaluate  $\Delta G_{\text{exf}}$ , PMF is computed by separating two initially adjacent nanosheets along a predefined reaction coordinate. Building on insights from previous studies [14,15], the most viable pathway for LPE of g-C<sub>3</sub>N<sub>4</sub> is identified as “shearing,” where one sheet glides over another while maintaining a parallel orientation. Consequently, the reaction coordinate is defined as the distance between the COMs of two adjacent nanosheets, varying from 0.35 nm to 3.8 nm by laterally shifting one nanosheet. US method [25] is employed to perform a series of restrained simulations, from which the PMF is derived using the WHAM [26]. For each functionalization, 40 US windows are used, within which a harmonic biasing potential (force constant of 5000 kJ/mol·nm<sup>2</sup>) is applied to the reaction coordinate. The PMF zero is set at the minimum value of the reaction coordinate.

All simulations are performed using GROMACS software [27] (version 2022.3). For each US simulation where the reaction coordinate is restrained, the system undergoes energy minimization using the steepest descent method. This is followed by equilibration in an NpT ensemble for 200 ps, with temperature and pressure maintained at 300 K and 1 bar using the Berendsen thermostat (time constant = 0.2 ps) and barostat (time constant = 5 ps). A production run is then carried out for 10 ns, during which bond lengths involving hydrogen atoms are constrained using the LINCS algorithm. Temperature is maintained at 300 K with the Nosé-Hoover thermostat (time constant = 0.1 ps), while isotropic pressure of 1 bar is controlled using the Parrinello-Rahman barostat (time constant = 1.0 ps). Periodic boundary conditions are applied in all three dimensions. Long-range electrostatic interactions are computed using the particle-mesh Ewald method, and short-range nonbonded interactions are assigned a cutoff radius of 1.2 nm. The equations of motion are integrated with a 2-fs time step during both equilibration and production stages.

To investigate the dynamics of solvent molecules near the nanosheets, additional simulations are performed with a single layer (1L) or two stacked layers (2L) of nanosheets equilibrated in DMF. The simulation parameters, including thermostat and barostat settings, are identical to those described above. Each simulation is run for 20 ns, with data from the last 10 ns used for statistical analysis.

## 6.3. Results

### 6.3.1. $\Delta G_{\text{exf}}$ comparison for different functionalizations

The PMF curves for pristine g-C<sub>3</sub>N<sub>4</sub> and functionalized g-C<sub>3</sub>N<sub>4</sub> with SO<sub>3</sub>H, COOH, NH<sub>2</sub>, OH, and CHO groups are shown in Figure 6.2a. For all systems except NH<sub>2</sub>, the PMF curves show a global minimum at approximately 0.5 nm, consistent with parallel-displaced  $\pi$ - $\pi$  stacking—a configuration where the  $\pi$ -electron clouds of adjacent layers overlap but are slightly offset, allowing the two nanosheets to stabilize at this separation. However, the NH<sub>2</sub> functionalized system does not exhibit a minimum and its curve is monotonic at short distances. This suggests that due to the presence of NH<sub>2</sub> functional groups, the two nanosheets tend to adopt a face-to-face  $\pi$ - $\pi$  stacking configuration. Additionally, the NH<sub>2</sub> groups, with their lone pairs of electrons, form H-bonds with hydrogen atoms from other NH<sub>2</sub> groups on the adjacent nanosheet, which averages to 3.6 H-bonds over a 10 ns simulation time. Both interactions provide a unique stabilization mechanism for stacked NH<sub>2</sub> functionalized g-C<sub>3</sub>N<sub>4</sub>.

All PMF curves reach a plateau around 3.50 nm, indicating negligible interaction between the nanosheets beyond this center-of-mass separation. The difference in PMF value between the global minimum and the plateau represents the free energy required to separate two stacked nanosheets from their equilibrium distance until they are no longer interacting, defined as  $\Delta G_{\text{exf}}$ . For instance,  $\Delta G_{\text{exf}}$  for the pristine system is shown in Figure 6.2a. For NH<sub>2</sub>, the PMF is simply calculated as the difference between the minimum PMF value and the plateau value. Among the functionalized g-C<sub>3</sub>N<sub>4</sub> samples, SO<sub>3</sub>H exhibits the lowest  $\Delta G_{\text{exf}}$ , suggesting the highest enhancement in exfoliation efficiency compared to pristine g-C<sub>3</sub>N<sub>4</sub>. COOH, NH<sub>2</sub>, and OH also show reduced  $\Delta G_{\text{exf}}$  values. Interestingly, CHO shows a higher  $\Delta G_{\text{exf}}$  compared to pristine g-C<sub>3</sub>N<sub>4</sub>, indicating that functionalization with CHO does not enhance, but rather hinders the exfoliation process. These findings underscore the significant impact of functional groups on  $\Delta G_{\text{exf}}$ . To understand the molecular mechanisms at play, we now examine the mobility of solvent molecules around the nanosheets, as this factor is closely tied to the efficiency of the exfoliation process [14].

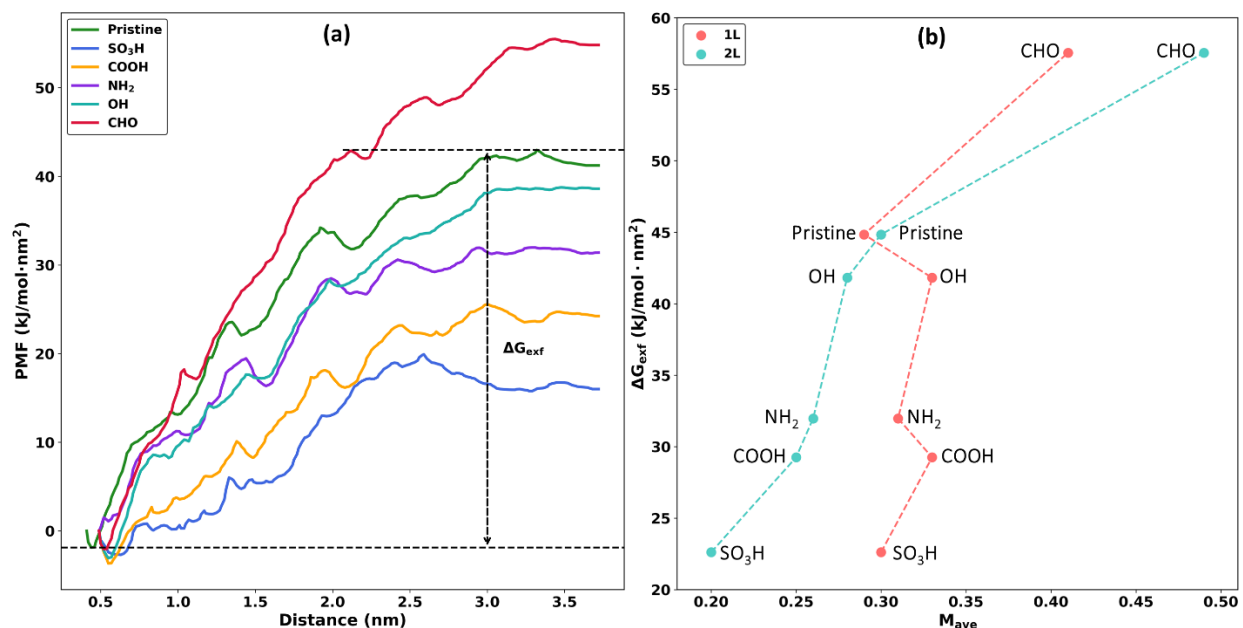


Figure 6.2. (a) The PMF curves of pristine g-C<sub>3</sub>N<sub>4</sub> and g-C<sub>3</sub>N<sub>4</sub> functionalized with different groups. (b) The relationship between  $\Delta G_{\text{exf}}$  and  $M_{\text{ave}}$  for 1L and 2L configurations of each type of nanosheet.

### 6.3.2. Solvent mobility around nanosheets

The solvent mobility,  $M_{\text{ave}}$ , is defined as the average fraction of solvent molecules leaving the first solvation layer—the region within 0.5 nm of the nanosheets—over successive 0.5 ns increments during the last 10 ns of a 20 ns simulation. This involves defining the first solvation layer at  $t = 10$  ns, calculating how many solvent molecules leave this layer after 0.5 ns, and repeating this process every 0.5 ns for 10 ns to obtain 20 mobility results, which are then averaged to determine the final  $M_{\text{ave}}$ . Lower mobility suggests stronger solvent-nanosheet interactions, which could reduce the adhesion between layers of the 2D nanomaterials, making exfoliation easier.

$M_{\text{ave}}$  is first calculated for 1L systems. Unlike our previous research on pristine systems, where there was a clear correlation between  $\Delta G_{\text{exf}}$  and  $M_{\text{ave}}$  for different solvents, this correlation no longer holds when dealing with functionalized nanosheets in the same solvent. Recognizing that  $\Delta G_{\text{exf}}$  is determined from two initially stacked nanosheets, we extended our calculations to include  $M_{\text{ave}}$  from 2L systems. The new calculations reveal a significant correction between  $M_{\text{ave}}$  and  $\Delta G_{\text{exf}}$  for the 2L systems (Figure 6.2b). To understand this change, we first divide each

nanosheet into periphery and interior parts (refer to Figure 6.1a for the definition of both parts). A periphery-to-interior ratio ( $\alpha$ ) can then be defined as the ratio of the perimeter along the periphery to the area of the interior (details for calculating perimeter and area are provided in Appendix D Section D1). As we transition from 1L to 2L systems,  $\alpha$  increases from 0.32 nm<sup>-1</sup> to 0.64 nm<sup>-1</sup>, and because the functional groups are only located in the periphery, it is expected that the impact of functionalization on solvent-nanosheet and solvent-solvent interactions is manifested to a greater extent in the 2L system. The roles of functional groups are therefore best addressed by examining how these interactions evolve as  $\alpha$  increases.

A notable trend emerges when increasing  $\alpha$ , i.e. when transitioning from 1L to 2L configurations (Figure 6.2b). For SO<sub>3</sub>H, COOH, NH<sub>2</sub>, and OH systems,  $M_{\text{ave}}$  decreases; for pristine g-C<sub>3</sub>N<sub>4</sub>,  $M_{\text{ave}}$  remains almost unchanged while for CHO,  $M_{\text{ave}}$  increases. This suggests that adding more functional groups can enhance solvent-nanosheet interactions in some systems, while it may be detrimental in others. To unravel why certain functional groups lead to more stable solvation layers (lower  $M_{\text{ave}}$ ) compared to others, we divided  $M_{\text{ave}}$  calculations into the periphery and interior regions. The periphery region includes solvent molecules within 0.5 nm of the periphery atoms, while the interior region includes solvent molecules within 0.5 nm of the nanosheet(s) but not in the periphery region. As shown in Figure 6.3,  $M_{\text{ave}}$  values are generally lower in the interior region compared to the periphery for both  $\alpha$  values. This can be attributed to stronger interactions between DMF and the interior part of the nanosheets through lone-pair- $\pi$  interactions [28]—where the lone pair of electrons on the oxygen in DMF interacts with the  $\pi$ -electron clouds of the aromatic structures of the nanosheets. These interactions are absent in the periphery, leading to higher  $M_{\text{ave}}$  values.

Although  $M_{\text{ave}}$  values are higher in the periphery for all systems, this does not imply that the periphery is less important than the interior in influencing solvent-nanosheet interactions. In fact, the periphery harbors significantly more solvent molecules than the interior (Figure 6.3). For instance, in the SO<sub>3</sub>H system, the average number of solvent molecules remaining in the periphery after 0.5 ns is 68.25, compared to 39.90 in the interior. This greater number of solvent molecules in the periphery suggests that the interactions occurring at the edges of the nanosheets can be crucial for the overall stability and dispersibility of the functionalized g-C<sub>3</sub>N<sub>4</sub>.

Comparing the 1L and 2L systems in Figure 6.3, it is observed that the SO<sub>3</sub>H and COOH systems show a decrease in  $M_{\text{ave}}$  for both periphery and interior regions when  $\alpha$  increases, with the



periphery experiencing a more pronounced decrease. This trend contrasts with pristine g-C<sub>3</sub>N<sub>4</sub>, which shows almost no change in  $M_{ave}$  for either region. NH<sub>2</sub> and OH systems, exhibit little to no change in the interior region but a considerable decrease in the periphery region. The CHO system, however, behaves uniquely: while it shows a decrease in  $M_{ave}$  in the interior region, the periphery experiences an increase. The increased mobility in the periphery for the CHO system outweighs the improvement in the interior, leading to an overall increase in solvent mobility for the first solvation layer when increasing  $\alpha$  (Figure 6.2b). This striking disparity reaffirms the substantial influence the periphery can have on the overall solvent mobility and, consequently,  $\Delta G_{exf}$ . Hence, we ask the questions: how do the functional groups interact with the solvent in the periphery region? Do they impact the interactions in the interior? Do they also affect solvent-solvent and nanosheet-nanosheet interactions?

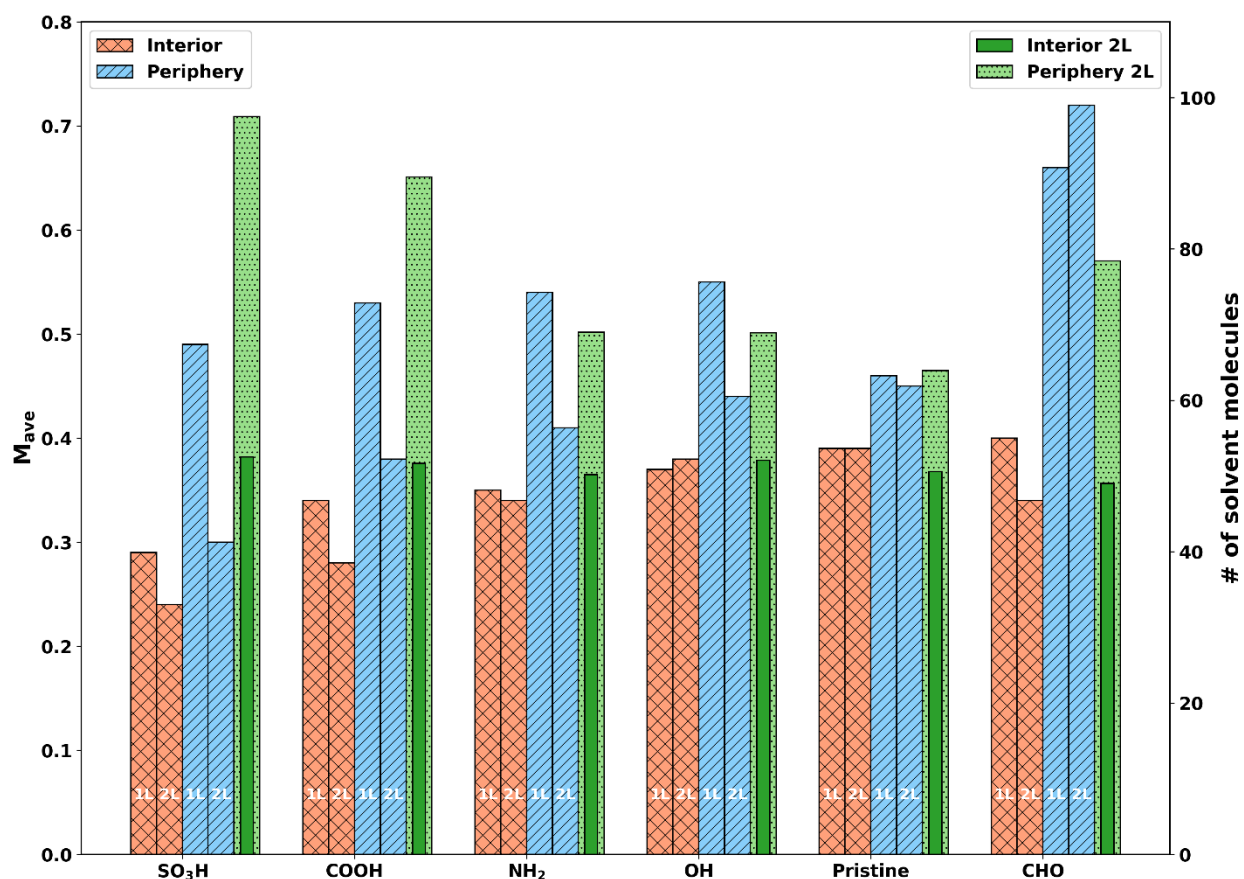


Figure 6.3.  $M_{ave}$  and number of solvent molecules in the periphery and interior regions of the first solvation layer, for both 1L and 2L configurations.

### 6.3.3. Functionalization impacts interactions in both periphery and interior regions

We first focus on the interactions occurring in the periphery region, where functional groups are present. Narrowing our analysis to three systems with distinct  $\Delta G_{\text{exf}}$  values,  $\text{SO}_3\text{H}$ , pristine, and CHO, we can generate a more in-depth understanding on these effects. Again, analysis will be conducted for both 1L and 2L systems, to monitor the changes upon increasing the presence of the functional groups ( $\alpha$  from 0.32 to 0.64  $\text{nm}^{-1}$ ).

Since the periphery is the only region where DMF molecules can form H-bonds with the nanosheets, a detailed analysis is conducted on the number of H-bonds and their lifetimes (for the calculation of H-bond lifetime, refer to Appendix D section D3), as these are crucial for solvent mobility in the periphery. Figure 6.4a plots the H-bond lifetime vs.  $M_{\text{ave}}$ . The number of H-bonds (inset of Figure 6.4a) in the pristine system is comparable to  $\text{SO}_3\text{H}$  for both 1L and 2L, but the H-bond lifetime is significantly shorter. This likely results from the fact that the nature of the H-bonds differs between the  $\text{SO}_3\text{H}$ -functionalized and pristine nanosheets. The  $\text{SO}_3\text{H}$  functional groups form  $\text{O}-\text{H}\cdots\text{O}$  H-bonds, which are stronger (21 kJ/mol) [29], while the pristine nanosheet forms  $\text{N}-\text{H}\cdots\text{O}$  H-bonds, which are weaker (8 kJ/mol) [29]. For the CHO system, although its H-bond lifetime is comparable to  $\text{SO}_3\text{H}$ , only a few H-bonds are formed with DMF (inset of Figure 6.4a), leads to higher  $M_{\text{ave}}$  values. As  $\alpha$  increases, the H-bond lifetime for the pristine and CHO-functionalized nanosheets remains constant, while the  $\text{SO}_3\text{H}$  system exhibits a significant increase in H-bond lifetime, underscoring the superior impact of  $\text{SO}_3\text{H}$  in stabilizing H-bonds.

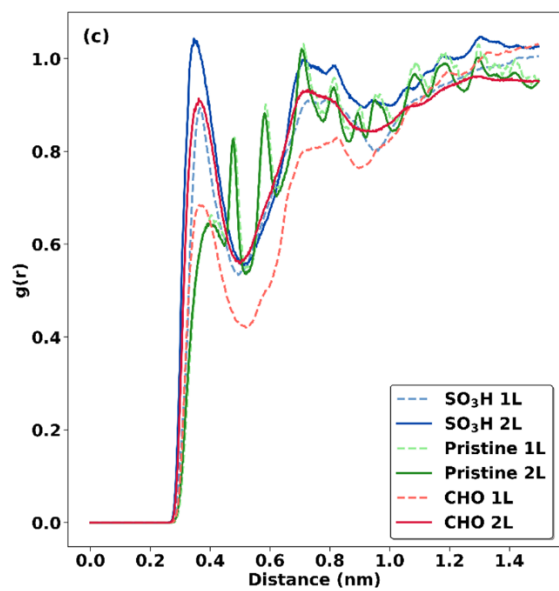
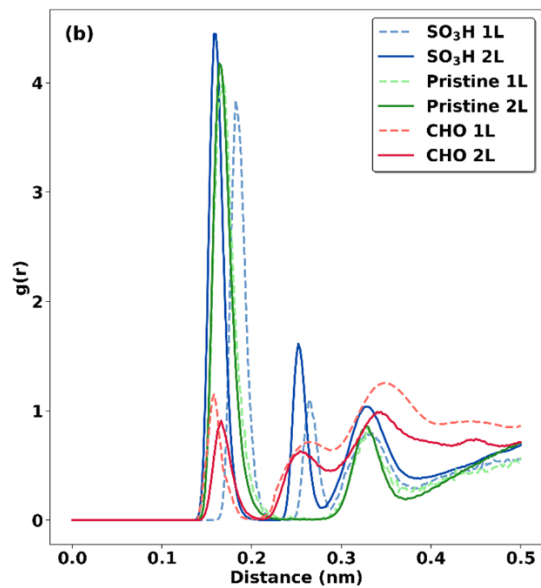
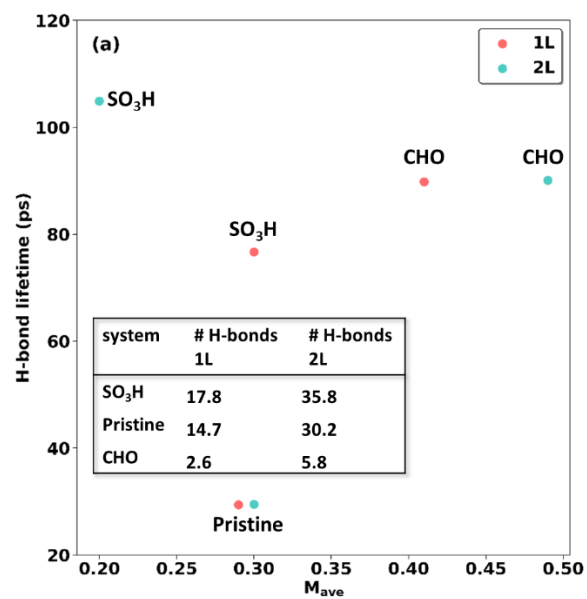


Figure 6.4. (a) Relationship between the H-bond lifetime and  $M_{ave}$  in 1L and 2L configurations. Inset shows the number of H-bonds formed with DMF. (b) RDF curves of oxygen in DMF around COM of functional groups for both 1L and 2L configurations. (c) RDF curves of oxygen in DMF around interior atoms of the nanosheet for both 1L and 2L configurations.

To delve deeper into the intricate interactions between DMF and the functional groups, the RDF of DMF oxygen around the COM of the functional groups is examined. For pristine g- $C_3N_4$ , which lacks functional groups, the hydrogens around the nanosheet are considered as reference points for RDF calculation. Figure 6.4b presents these results, focusing on distances up to 0.5 nm to highlight where most interactions occur. The first peak in the RDFs (near 0.2 nm) corresponds to the H-bonds between DMF molecules and the nanosheet periphery. It is significantly higher for  $SO_3H$  and pristine systems compared to CHO, consistent with the greater number of H-bonds seen in Figure 6.4a. The second peak (between 0.2 and 0.3 nm) represents the interaction of DMF oxygen with oxygen atoms in the  $SO_3H$  or CHO groups (absent in pristine). The third peak in the RDFs of  $SO_3H$  and CHO systems (between 0.3 and 0.4 nm) corresponds to the interactions of DMF oxygen with sulfur in  $SO_3H$  or carbon in CHO groups.

When comparing 1L to 2L systems, an intriguing pattern emerges. As  $\alpha$  increases, for  $SO_3H$ , all RDF peaks shift to a closer distance, and the peak values increase. This suggests that the presence of more  $SO_3H$  functional groups enhance the intensity of solvent-nanosheet interactions by reducing the distance between oxygen in DMF and all atoms in the functional groups. Conversely, for CHO, the first RDF peak moves to a larger distance, and the RDF values at all peaks decrease, indicating diminished solvent-nanosheet interactions. These results correlate well with the observed changes in  $M_{ave}$  for  $SO_3H$  and CHO in the periphery region (Figure 6.3), suggesting the significance of the interactions between oxygen in DMF and the functional groups.

Referring back to Figure 6.3, as we transition from 1L to 2L systems,  $M_{ave}$  in the interior region decreases for both  $SO_3H$  and CHO. To elucidate the underlying interactions, we examine the RDF curves of DMF oxygen around the atoms in the interior part of the nanosheets (Figure 6.4c). The systems with functional groups ( $SO_3H$ , CHO) not only produce higher first peaks, these peaks also appear at shorter distances compared to the pristine system, which indicates stronger attraction between DMF molecules and the interior of the functionalized nanosheets through lone-pair  $\pi$  interactions. As  $\alpha$  increases, a clear rise in the first RDF peak is observed for both  $SO_3H$  and

CHO systems, which aligns with the decrease in  $M_{ave}$  in the interior region. Despite the reduced  $M_{ave}$  in the interior region of CHO, the overall  $M_{ave}$  increases because of the significantly greater number of solvent molecules in the periphery region (Figure 6.3). The pristine system does not exhibit significant changes in peak positions or values regardless of the region of investigation (Figure 6.4b for periphery, Figure 6.4c for interior), which aligns with its lack of change in  $M_{ave}$  as  $\alpha$  varies (Figure 6.2b and Figure 6.3).

The above findings illustrate that the functional groups not only affect solvent-nanosheet interactions in the periphery where they are located but also influence these interactions in the interior. To further understand these observations, we need to examine solvent-solvent interactions and how functional groups might modify these interactions.

#### 6.3.4. Functionalization impacts nanosheet dipole moment

In DMF, a dipole moment exists which directs from N towards O, with N being partially positive and O being partially negative [30]. The dipole-dipole interactions among DMF molecules are closely related to the formation of H-bonds between DMF molecules. Previous research by Zhang et al. [31] supports the existence of such interactions, showing that weak H-bonds form between the oxygen of one DMF molecule and the hydrogen of another. On the other hand, the dipole moment of pristine g-C<sub>3</sub>N<sub>4</sub> is relatively low, limiting its interaction capabilities with surrounding DMF molecules. For instance, Kamel et al. reported that pristine g-C<sub>3</sub>N<sub>4</sub> has a dipole moment of 3.43 Debye in an aqueous media, but functionalization with chitosan boosts it to 9.14 Debye in aqueous media [32]. Similarly, Asif et al. [33] demonstrated that pristine g-C<sub>3</sub>N<sub>4</sub> has a dipole moment of 1.82 Debye in the gas phase, which increases to 3.58 Debye with Ionidamine functionalization in the gas phase and further to 6.14 Debye in water. By examining the partial atomic charges (PACs) in SO<sub>3</sub>H, we find that the PACs of S, O, and N (connected to S) are 0.82, -0.52, and -0.71, respectively. This significant charge difference propagates throughout the nanosheet, leading to average PACs of the C and N atoms in the nanosheet being 0.70 and -0.57, respectively. In contrast, the CHO system has much lower PACs, with C, O, and N (connected to C) having values of 0.28, -0.34, and -0.28, respectively, and average PACs of the C and N atoms in the nanosheet being only 0.40 and -0.35. For the pristine system, these values are 0.43 and -0.39, respectively.

The presence of dipole moment in the nanosheet can be analogized to the creation of an internal electric field within the material. Bai et al. discussed how fluorine doping in g-C<sub>3</sub>N<sub>4</sub>

induces a surface polarization electric field due to the high electronegativity of fluorine, which causes a redistribution of electron density and creates an uneven charge distribution [34]. This phenomenon is not unique to fluorine; similar effects can be expected from other functional groups that introduce polar bonds and non-uniform charge distributions. An internal electric field, akin to an externally applied one, influences the interactions within the surrounding medium. Gao et al. explored the response of DMF to an external electric field and found that such a field disrupt H-bonding between DMF molecules, leading to significant structural reorganization [30]. This reorganization includes changes in O–O, O–C, O–H, and C–C interactions, with the electric field causing the H-bonds, particularly the already weak  $\text{-HC=O}\cdots\text{HC=O}$  interactions, to become further weakened or eliminated.

To assess the effect of dipole moment of the nanosheet on DMF dipole-dipole interactions, the distance distribution between nitrogen atoms from one DMF molecule and oxygen atoms from another DMF molecule is analyzed, focusing on DMF molecules in the first solvation layer as compared to those in the bulk. Figure 6.5a-c depict these distance distributions for  $\text{SO}_3\text{H}$ , pristine, and CHO systems, respectively, with both 1L and 2L configurations. As  $\alpha$  increases, the  $\text{SO}_3\text{H}$  system shows more pronounced peaks in shorter distances, suggesting stronger interactions among DMF molecules in the 2L configuration. Conversely, for the CHO system, the distributions show more peaks at larger N-O distances in the 2L configuration, indicating weaker solvent-solvent interactions as  $\alpha$  increases. For the pristine system, no significant change is observed with an increase in  $\alpha$ . The bulk distribution, shown for comparison in each figure, remains consistent. These results suggest that increasing  $\alpha$  enhances the electrostatic interactions among DMF molecules in the first solvation layer in the presence of the  $\text{SO}_3\text{H}$  functional groups.

The orientation of DMF molecules under the influence of the nanosheet dipole moment is examined by calculating the angle between the dipole moment (from N to O) of each DMF molecule within 0.5 nm of the COM of each functional group and the closest dipole moment in the periphery region of the nanosheet. The dipole moments from the nanosheet are S to N for  $\text{SO}_3\text{H}$ , H to N for pristine, and C (from the functional group) to N for CHO, where the N atoms are bonded to the S, H, or C atoms as illustrated in inset of Figure 6.6a. The results, shown in Figure 6.6a, reveal a bimodal angle distribution for  $\text{SO}_3\text{H}$  and pristine systems, with peaks around  $30^\circ$  and  $90^\circ$ . Conversely, the CHO system exhibits a normal distribution with a peak around  $90^\circ$ . For angles greater than  $110^\circ$ , all systems display similar distributions, but differences emerge at

smaller angles. Smaller angles correspond to the alignment of the DMF dipole moment with the dipole moment of the nanosheet, and stronger dipole-dipole interactions. As  $\alpha$  increases, the peak at  $30^\circ$  for  $\text{SO}_3\text{H}$  becomes more pronounced, indicating enhanced alignment of DMF and nanosheet dipole moments. In contrast, CHO shows a decreased probability at smaller angles as  $\alpha$  increases, suggesting weaker dipole-dipole interactions. Pristine g- $\text{C}_3\text{N}_4$  experiences minimal change with increasing  $\alpha$ .

Consolidating the results on both distance and angle distributions, it becomes evident that adding more layers of  $\text{SO}_3\text{H}$  functionalized nanosheets enhances dipole-dipole interactions in DMF while adding more layers of CHO functionalized nanosheets weakens these interactions. The  $\text{SO}_3\text{H}$  groups, with their strong polar nature, also attract DMF molecules. This, along with the increased solvent-solvent attraction, leads to a more stable solvation layer not only around the periphery but also near the interior, reducing the free energy required for exfoliation. Conversely, the functionalization with CHO groups results in a lower dipole moment compared to pristine g- $\text{C}_3\text{N}_4$ , which diminishes the dipole-dipole interactions among DMF molecules as well as between DMF and the nanosheet, making the solvation layer less stable and thereby increasing the free energy required for exfoliation.

Then enhanced dipole-dipole interaction by  $\text{SO}_3\text{H}$  functionalization is manifested through the rotation of the top layer relative to the bottom layer, as depicted in Figure 6.6b. This rotation facilitates the alignment of DMF and nanosheet dipole moments, and even causes the two nanosheets to slightly separate near the vertices, allowing DMF molecules to intercalate between them (Figure 6.6c). This intercalation is beneficial for exfoliation as it generates a steric effect, shielding nanosheet-nanosheet interactions. Such a behavior is absent in both pristine and CHO systems.

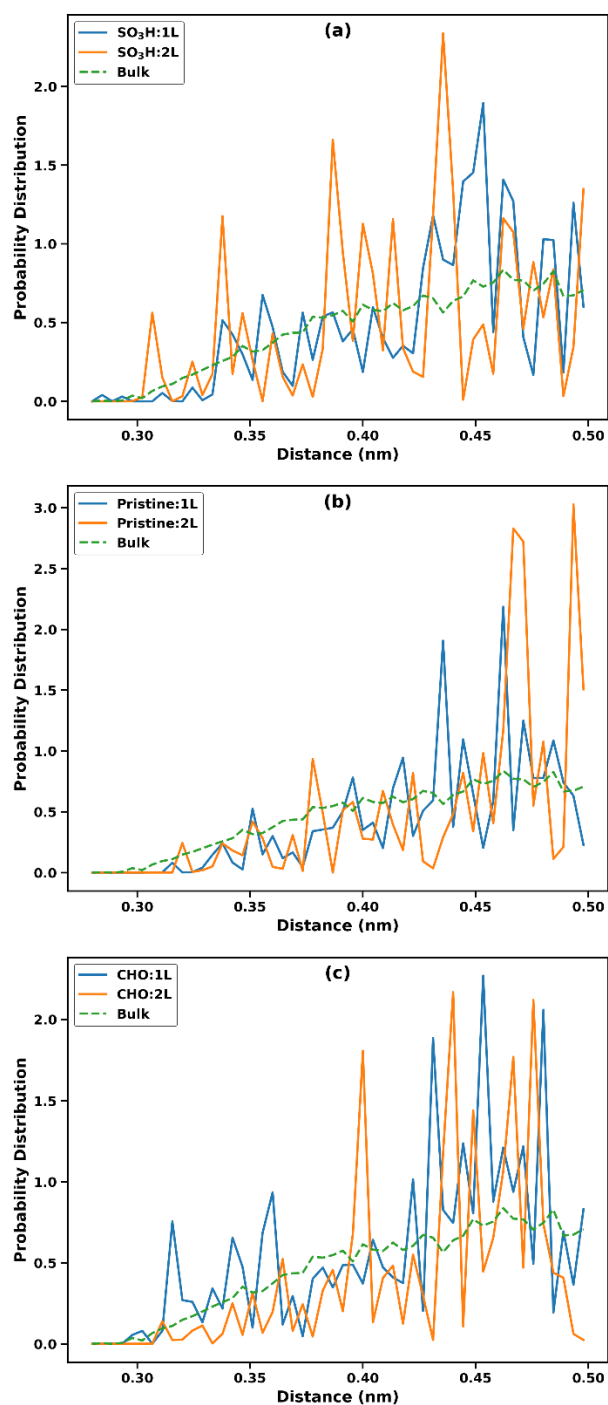


Figure 6.5. Probability distribution for the distance between nitrogen and oxygen atoms of different DMF molecules in the first solvation layer and bulk for (a) SO<sub>3</sub>H, (b) Pristine, and (c) CHO systems.



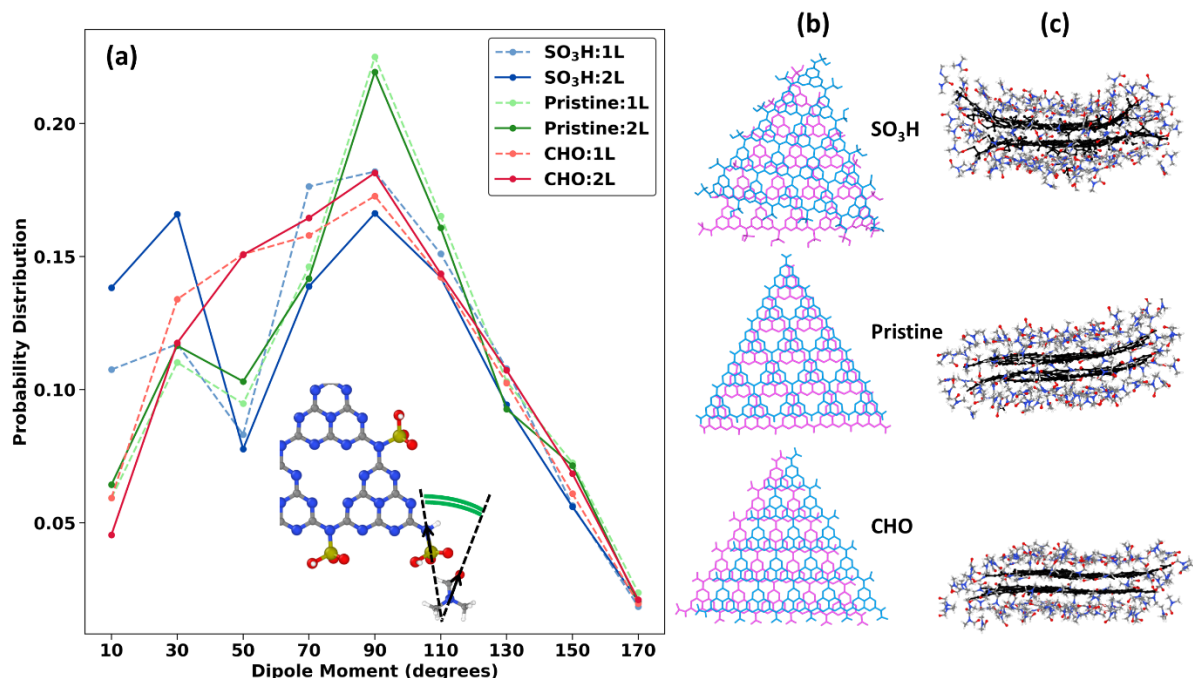


Figure 6.6. (a) Probability distribution for the angle between DMF dipole moments and the nearest dipole moment in the periphery region of the nanosheet. (b) Top view of equilibrated nanosheets shown in (c) - DMF molecules are omitted for clarity. (c) Molecular configurations for 2L systems along with the first solvation layer for SO<sub>3</sub>H, pristine, and CHO. Inset in (a) shows an example for the angle definition.

## 6.4. Discussion

Building on the trend observed when transitioning from 1L to 2L systems, we extended our PMF calculations to explore the effect of increasing the number of layers from 2 to 3. This approach involved exfoliating the top layer from the two bottom layers, enabling us to investigate whether the observed trend continues and to assess the impact of additional functional groups on LPE. The results are presented in Figure 6.7a. As anticipated, the SO<sub>3</sub>H system demonstrates significant improvements when increasing from 2L to 3L. The  $\Delta G_{\text{exf}}$  values decreases by approximately 30%, indicating enhanced exfoliation efficiency with more layers. In contrast, the pristine system shows no significant change, while the CHO system exhibits a 22% increase in  $\Delta G_{\text{exf}}$ , highlighting a deterioration in exfoliation performance. These results underscore the pivotal

role of functionalization in improving the exfoliation process. Functional groups such as  $\text{SO}_3\text{H}$  not only enhance exfoliation efficiency but also contribute to the stabilization of the nanosheets, as previously reported in the literature [7].

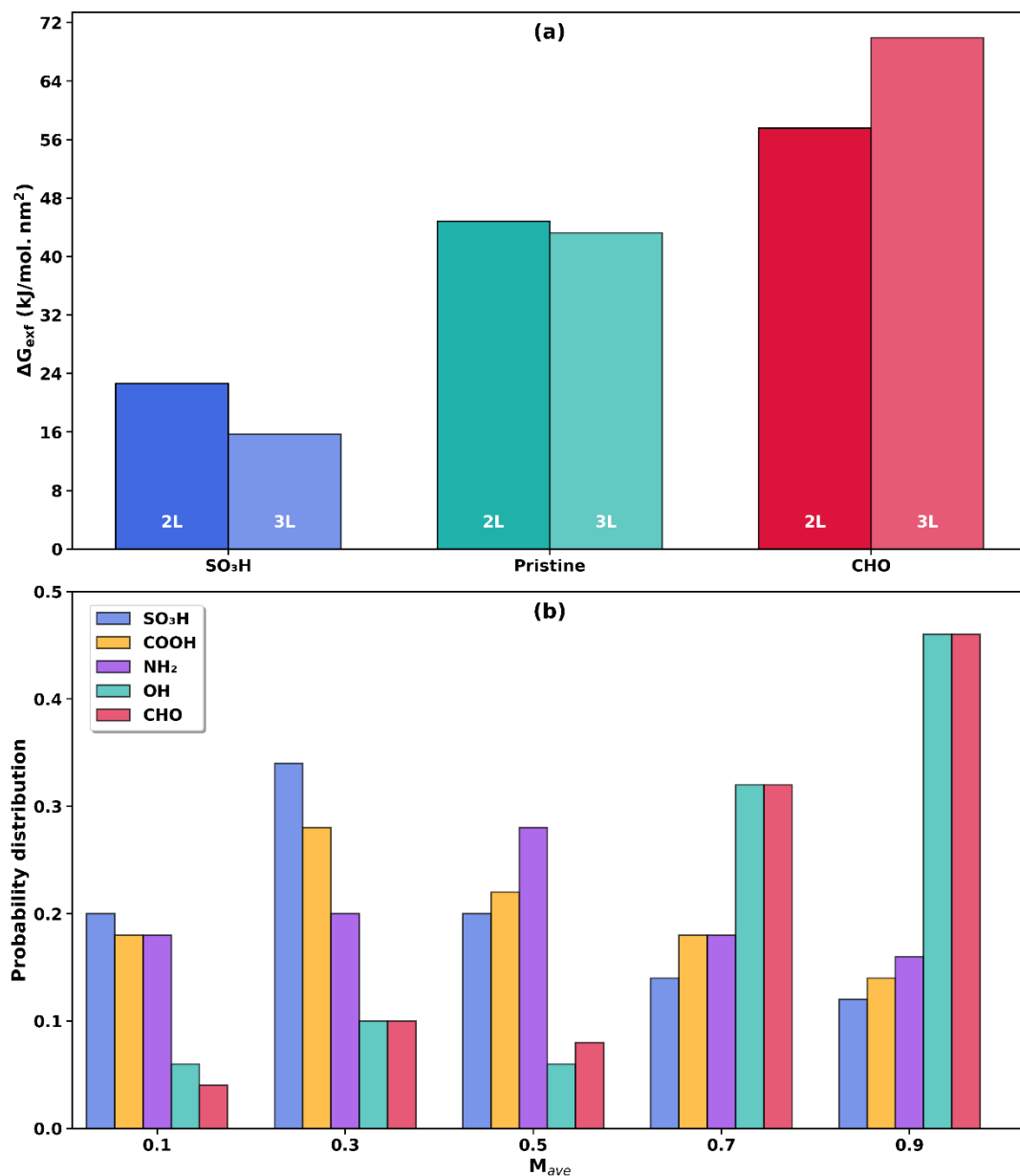


Figure 6.7. (a)  $\Delta G_{\text{exf}}$  values for 2L and 3L  $\text{SO}_3\text{H}$ , Pristine, and CHO systems. (b) Histogram of  $M_{\text{ave}}$  values for 50 common organic solvents tested with different functional groups in a 2L system.

Given the limited research on solvent selection for LPE in the presence of functional groups, and the absence of comprehensive computational studies, we attempt to provide recommendations that may assist experimentalists in choosing the optimal solvent for their specific functionalized g-C<sub>3</sub>N<sub>4</sub> systems. Based on the insights gained earlier, we selected 50 common organic solvents (listed in Appendix D Section D2), equilibrated the 2L system for each functionalization over 20 ns, and calculated  $M_{ave}$  for the last 10 ns. The histogram of these results is presented in Figure 6.7b. SO<sub>3</sub>H, COOH, and NH<sub>2</sub> functional groups have a larger number of solvents with lower  $M_{ave}$  values, suggesting their greater interaction with a range of organic solvents. In contrast, the CHO and OH functional groups show a narrower range of effective solvents, highlighting the challenges in finding suitable solvents for these systems. To provide practical recommendations, we identified the top five solvents with the lowest  $M_{ave}$  for each functional group, as summarized in Table 6.1. These solvents are expected to enhance the exfoliation efficiency for the respective functional groups.

Table 6.1. Best five solvents based on  $M_{ave}$  results for each functional group.

Functional group	Best five solvents
SO <sub>3</sub> H	NVP, Phenoxybenzene, Triethyl phosphate, NMP, m-Cresol
COOH	Quinoline, BD, NVP, Acetamide, Formamide
NH <sub>2</sub>	NMP, NVP, 1,2-Dibromopropane, BD, Triethyl phosphate
OH	Acetamide, Cyclohexanone, NMP, MSA, BD
CHO	Phenoxybenzene, BD, NMP, Formamide, MSA

## 6.5. Conclusion

This study explores the impact of chemical functionalization on the liquid-phase exfoliation (LPE) efficiency of graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) nanosheets through molecular dynamics (MD) simulations. Functional groups sulfonic (SO<sub>3</sub>H), carboxyl (COOH), amine (NH<sub>2</sub>), hydroxyl (OH), and aldehyde (CHO) are evaluated. Our findings reveal that functionalization can either enhance or hinder exfoliation efficiency compared with pristine g-C<sub>3</sub>N<sub>4</sub> nanosheet. Specifically, SO<sub>3</sub>H and COOH groups significantly improve exfoliation by reducing the free

energy of exfoliation ( $\Delta G_{\text{exf}}$ ) and promoting stable solvent-nanosheet interactions.  $\text{SO}_3\text{H}$  groups exhibited the most substantial enhancement, as they create strong dipole-dipole interactions and stable hydrogen bonds, decrease solvent mobility, and increase interaction strength around the nanosheets.  $\text{COOH}$  and  $\text{NH}_2$  groups also show positive effects, though to a lesser extent. On the other hand,  $\text{CHO}$  functionalization increases the free energy of exfoliation by weakening solvent-nanosheet interactions and increasing solvent mobility. These varied outcomes underscore the complexity of solvent-functional group interactions. By understanding the interactions between functionalized  $\text{g-C}_3\text{N}_4$  and solvents, we propose strategies for selecting solvents that can potentially facilitate scalable production and broaden the application of  $\text{g-C}_3\text{N}_4$ -based materials in various technological fields.

## 6.6. Acknowledgement

Digital Research Alliance of Canada is gratefully acknowledged for providing the computing resources and technical support. TT acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Grant number: RGPIN-2018-04281) and the Canada Research Chair program (CRC; Grant number: TIER1 2021-00023). KS thanks NSERC for financial support (Grant number: RGPIN-2020-04620). ES acknowledges support from Alberta Innovates Graduate scholarship.

## 6.7. References

- [1] J. Fu, J. Yu, C. Jiang, B. Cheng,  $\text{g-C}_3\text{N}_4$ -Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.
- [2] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [3] J. Fernández-Catalá, R. Greco, M. Navlani-García, W. Cao, Á. Berenguer-Murcia, D. Cazorla-Amorós,  $\text{g-C}_3\text{N}_4$ -Based Direct Z-Scheme Photocatalysts for Environmental Applications, *Catalysts* 2022, Vol. 12, Page 1137 12 (2022) 1137. <https://doi.org/10.3390/CATAL12101137>.

- [4] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.
- [5] S. Das, A. Chowdhury, Recent advancements of g-C<sub>3</sub>N<sub>4</sub>-based magnetic photocatalysts towards the degradation of organic pollutants: a review, *Nanotechnology* 33 (2021) 072004. <https://doi.org/10.1088/1361-6528/AC3614>.
- [6] M. Majdoub, Z. Anfar, A. Amedlous, Emerging chemical functionalization of g-C<sub>3</sub>N<sub>4</sub>: Covalent/noncovalent modifications and applications, *ACS Nano* 14 (2020) 12390–12469. [https://doi.org/10.1021/ACSNANO.0C06116/ASSET/IMAGES/MEDIUM/NN0C06116\\_0046.GIF](https://doi.org/10.1021/ACSNANO.0C06116/ASSET/IMAGES/MEDIUM/NN0C06116_0046.GIF).
- [7] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906. [https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL\\_FILE/LA7B02441\\_SI\\_007.AVI](https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL_FILE/LA7B02441_SI_007.AVI).
- [8] S. Xu, Z. Zhang, D. Wang, J. Lu, Y. Guo, S. Kang, X. Chang, Ultrafast plasma method allows rapid immobilization of monatomic copper on carboxyl-deficient g-C<sub>3</sub>N<sub>4</sub> for efficient photocatalytic hydrogen production, *Front Chem* 10 (2022) 972496. <https://doi.org/10.3389/FCHEM.2022.972496/BIBTEX>.
- [9] F. Li, H. Zhou, J. Fan, Q. Xiang, Amine-functionalized graphitic carbon nitride decorated with small-sized Au nanoparticles for photocatalytic CO<sub>2</sub> reduction, *J Colloid Interface Sci* 570 (2020) 11–19. <https://doi.org/10.1016/J.JCIS.2020.02.108>.
- [10] M. Yousefi, S. Villar-Rodil, J.I. Paredes, A.Z. Moshfegh, Oxidized graphitic carbon nitride nanosheets as an effective adsorbent for organic dyes and tetracycline for water remediation, *J Alloys Compd* 809 (2019) 151783. <https://doi.org/10.1016/J.JALLCOM.2019.151783>.
- [11] J. Feng, D. Zhang, J. Li, S. Bi, Y. Ma, Graphitic carbon nitride nanodots: electronic structure and its influence factors, *J Mater Sci* 55 (2020) 5488–5498. <https://doi.org/10.1007/S10853-020-04396-X/METRICS>.

- [12] J. Yang, X. Yang, Y. Li, Molecular simulation perspective of liquid-phase exfoliation, dispersion, and stabilization for graphene, *Curr Opin Colloid Interface Sci* 20 (2015) 339–345. <https://doi.org/10.1016/J.COCIS.2015.10.005>.
- [13] S. Haar, M. Bruna, J.X. Lian, F. Tomarchio, Y. Olivier, R. Mazzaro, V. Morandi, J. Moran, A.C. Ferrari, D. Beljonne, A. Ciesielski, P. Samorì, Liquid-Phase Exfoliation of Graphite into Single- and Few-Layer Graphene with  $\alpha$ -Functionalized Alkanes, *Journal of Physical Chemistry Letters* 7 (2016) 2714–2721. [https://doi.org/10.1021/ACS.JPCLETT.6B01260/SUPPL\\_FILE/JZ6B01260\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCLETT.6B01260/SUPPL_FILE/JZ6B01260_SI_001.PDF).
- [14] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910. <https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [15] E. Shahini, N. Chaulagain, K. Shankar, T. Tang, Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning, *ACS Appl Mater Interfaces* 15 (2023) 53786–53801. [https://doi.org/10.1021/ACSAMI.3C09347/SUPPL\\_FILE/AM3C09347\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.3C09347/SUPPL_FILE/AM3C09347_SI_001.PDF).
- [16] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [17] Y. Xu, S.P. Gao, Band gap of C<sub>3</sub>N<sub>4</sub> in the GW approximation, *Int J Hydrogen Energy* 37 (2012) 11072–11080. <https://doi.org/10.1016/J.IJHYDENE.2012.04.138>.
- [18] H. Li, Z. Zhang, Y. Liu, W. Cen, X. Luo, Functional Group Effects on the HOMO–LUMO Gap of g-C<sub>3</sub>N<sub>4</sub>, *Nanomaterials* 2018, Vol. 8, Page 589 8 (2018) 589. <https://doi.org/10.3390/NANO8080589>.
- [19] Z. Chen, S. Li, Y. Peng, C. Hu, Tailoring aromatic ring-terminated edges of g-C<sub>3</sub>N<sub>4</sub> nanosheets for efficient photocatalytic hydrogen evolution with simultaneous antibiotic removal, *Catal Sci Technol* 10 (2020) 5470–5479. <https://doi.org/10.1039/D0CY00898B>.
- [20] L. Ma, H. Fan, M. Li, H. Tian, J. Fang, G. Dong, A simple melamine-assisted exfoliation of polymeric graphitic carbon nitrides for highly efficient hydrogen production from water under

visible light, *J Mater Chem A Mater* 3 (2015) 22404–22412.

<https://doi.org/10.1039/C5TA05850C>.

- [21] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017). <https://doi.org/10.1093/nar/gkx312>.
- [22] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a. <https://doi.org/10.2477/jccjie.2018-0034>.
- [23] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).
- [24] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [25] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [26] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [27] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [28] A. Jain, V. Ramanathan, R. Sankararamakrishnan, Lone pair  $\cdots \pi$  interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds, *Protein Sci* 18 (2009) 595. <https://doi.org/10.1002/PRO.67>.
- [29] N. Grinberg, E. Grushka, *Advances in chromatography. Volume 54*, (n.d.).

- [30] W. Gao, H. Niu, T. Lin, X. Wang, L. Kong, Molecular dynamics study of response of liquid N,N-dimethylformamide to externally applied electric field using a polarizable force field, *Journal of Chemical Physics* 140 (2014) 44501. <https://doi.org/10.1063/1.4861893/72908>.
- [31] C. Zhang, Z. Ren, L. Liu, Z. Yin, Modelling hydrogen bonds in NN-dimethylformamide, *Mol Simul* 39 (2013) 875–881. <https://doi.org/10.1080/08927022.2013.775438>.
- [32] M. Kamel, M. Mohammadi, K. Mohammadifard, E.A. Mahmood, M.R. Poor Heravi, A. Heshmati J.M., Z. Hossaini, Comprehensive theoretical prediction of the stability and electronic properties of hydroxyurea and carmustine drugs on pristine and Chitosan-functionalized graphitic carbon nitride in vacuum and aqueous environment, *Vacuum* 207 (2023) 111565. <https://doi.org/10.1016/J.VACUUM.2022.111565>.
- [33] K. Asif, M. Perveen, R.A. Khera, S. Nazir, A. Raza Ayub, T. Asif, M. Shabbir, J. Iqbal, Computational and theoretical study of graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) as a drug delivery carrier for lonidamine drug to treat cancer, *Comput Theor Chem* 1206 (2021) 113459. <https://doi.org/10.1016/J.COMPTC.2021.113459>.
- [34] X. Bai, X. Wang, X. Lu, T. Jia, B. Sun, C. Wang, S. Hou, R. Zong, A fluorine induced enhancement of the surface polarization and crystallization of g-C<sub>3</sub>N<sub>4</sub> for an efficient charge separation, *New Journal of Chemistry* 45 (2021) 9334–9345. <https://doi.org/10.1039/D1NJ00668A>.



# Chapter 7: Conclusion and Future Perspective

## 7.1. Overall conclusions

This thesis presents a comprehensive investigation into the liquid-phase exfoliation (LPE) of pristine and functionalized graphitic carbon nitride ( $\text{g-C}_3\text{N}_4$ ) nanosheets, emphasizing the molecular mechanisms and optimization strategies through advanced simulations and machine learning techniques. The research is structured into four main chapters, each addressing different aspects of the exfoliation process, from foundational principles to practical applications and solvent recommendations.

**Chapter 3** explores the LPE of  $\text{g-C}_3\text{N}_4$  using molecular dynamics (MD) simulations. The study identifies the most probable path for the exfoliation process and evaluates the free energy of exfoliation in nine common solvents categorized into three groups. Results demonstrate that the solvation free energy of a single  $\text{g-C}_3\text{N}_4$  nanosheet correlates directly with the free energy of exfoliation. Solvents such as NMP and DMF produce high magnitude of solvation free energies and stable adsorption layers, making them particularly effective for LPE. This chapter establishes the foundational principles for solvent selection, emphasizing the critical role of solvent-nanosheet interactions in determining exfoliation efficiency.

**Chapter 4** integrates MD simulations with machine learning (ML) to predict the free energies of exfoliation and solvation for  $\text{g-C}_3\text{N}_4$  in a broader range of solvents. A comprehensive dataset of 49 solvents is created, and various ML algorithms are employed to identify key solvent descriptors influencing the LPE process. The extra-trees regressor emerges as the most accurate model. The ML model recommends several promising solvents, including benzyl alcohol (Bn) and methanesulfonic acid (MSA), which had not previously been used for LPE. Experimental validation confirms the efficacy of these solvents, demonstrating the practical utility of the ML model in guiding solvent selection. This chapter showcases the potential of ML to streamline the solvent selection process, enhancing efficiency and reducing reliance on extensive MD simulations.

**Chapter 5** investigates the performance of binary solvent mixtures in the LPE of g-C<sub>3</sub>N<sub>4</sub>, revealing that one component can dominate the exfoliation process. MD simulations and ML predictions examine the free energy of exfoliation in 171 pure solvents and 14,535 binary mixtures. Certain binary mixtures, such as NMP:Cyclohexane (NMP:CH) and Methanol:Dichloromethane (MET:DCM), demonstrate a distinct behavior where one solvent significantly influences the LPE performance. NMP dominates in the NMP:CH mixture, showing good LPE performance, while MET dominates in the MET:DCM mixture, resulting in poor performance. Additionally, the study finds that certain solvents like IPA, when mixed with water at 25%, can achieve free energies of exfoliation close to those of pure organic solvents. This insight provides experimentalists with valuable guidance for selecting solvent mixtures, potentially reducing costs and environmental impact by incorporating water into the mixtures.

**Chapter 6** examines the impact of chemical functionalization on the LPE efficiency of g-C<sub>3</sub>N<sub>4</sub> nanosheets. MD simulations evaluate the effects of various functional groups, including sulfonic (SO<sub>3</sub>H), carboxyl (COOH), amine (NH<sub>2</sub>), hydroxyl (OH), and aldehyde (CHO), on the exfoliation process. The findings reveal that functionalization can significantly enhance or hinder exfoliation. SO<sub>3</sub>H and COOH groups improve exfoliation by reducing the free energy of exfoliation and promoting stable solvent-sheet interactions, while CHO groups hinder the process by weakening these interactions. The chapter also recommends effective solvents for each functional group based on solvent mobility calculations. This research discusses the complexity of solvent-functional group interactions and offers practical guidelines for optimizing the production of functionalized g-C<sub>3</sub>N<sub>4</sub> nanosheets.

**Overall, this thesis** contributes to the field of nanomaterial synthesis by providing a detailed understanding of molecular interactions in the LPE of g-C<sub>3</sub>N<sub>4</sub> nanosheets. The integration of MD simulations and ML techniques offers a powerful approach to facilitate solvent selection, enabling the identification of optimal solvents and solvent mixtures for efficient and scalable production. The insights gained can be applied to other two-dimensional materials, broadening the scope of applications and advancing the development of high-performance nanomaterials for energy conversion, environmental remediation, and other technological fields.

## **7.2. Future perspectives**

### **7.2.1. Enhancing realism in simulations**

To bring simulations closer to the actual LPE process, future research could explore areas beyond potential of mean force (PMF) calculations for two or three nanosheets. Simulating scenarios with multiple layers of g-C<sub>3</sub>N<sub>4</sub> of larger size and/or incorporating the mechanical effects of sonication could provide more realistic insights. Less computationally expensive approaches, such as coarse-grained molecular dynamics or mesoscale modeling, could also be considered to expand the database of solvent efficacy. These simulations can be calibrated to correlate with experimental metrics like nanosheet concentration and yield, enabling a more comprehensive evaluation of solvent performance.

### **7.2.2. Advancing computational and experimental integration**

Integrating computational and experimental approaches has shown great promise in this research. Future efforts could aim to create a seamless feedback loop between simulation and experimentation. Real-time data sharing can enable immediate refinement of models and experimental protocols, fostering iterative improvements. The development of high-throughput experimental setups, combined with automated data analysis pipelines, can accelerate the identification of optimal solvents and functionalization methods. Multi-scale modeling techniques could be further developed to link molecular-level interactions with macroscopic behaviors, providing a holistic understanding of the LPE process.

### **7.2.3. Exploring solvent mixtures**

The exploration of binary and tertiary solvent mixtures offers significant potential for enhancing LPE efficiency. Creating a comprehensive dataset of these mixtures will enable the development of robust ML models to predict solvent performance. Identifying solvents with synergistic effects or those that overpower water in mixtures can pave the way for more environmentally friendly "green" LPE processes. Future research could focus on systematically studying the interactions within these mixtures to uncover new insights and optimize solvent combinations.

#### **7.2.4. Exploring functional group synergies**

The influence of a few functional groups on g-C<sub>3</sub>N<sub>4</sub> exfoliation has been discussed. Future studies could investigate the synergistic effects of having multiple functional groups on the same nanosheet. Co-functionalization with groups that enhance both exfoliation efficiency and suspension stability could lead to significant improvements. Advanced characterization techniques, such as in-situ spectroscopy and high-resolution electron microscopy, can provide detailed observations of these synergies at the molecular level. Additionally, exploring the impact of functional group density and distribution will offer valuable insights for optimizing functionalization strategies.

#### **7.2.5. Extending methodologies to other 2D materials**

The methodologies developed for g-C<sub>3</sub>N<sub>4</sub> can be extended to other 2D materials. Investigating the LPE processes of materials such as graphene, boron nitride, and transition metal dichalcogenides using similar ML-assisted simulations and experimental integrations can broaden the applicability of the findings. This comparative approach will help identify universal principles as well as material-specific strategies, advancing the field of 2D material synthesis and application.

# Unified Bibliography

- [1.1] M.H. Ahmadi, M. Ghazvini, M.A. Nazari, M.A. Ahmadi, F. Pourfayaz, G. Lorenzini, T. Ming, Renewable energy harvesting with the application of nanotechnology: A review, *Int J Energy Res* 43 (2019) 1387–1410. <https://doi.org/10.1002/ER.4282>.
- [1.2] G. Ali Mansoori, T.R. Bastami, A. Ahmadpour, Z. Eshaghi, ENVIRONMENTAL APPLICATION OF NANOTECHNOLOGY, (2008) 439–493. [https://doi.org/10.1142/9789812790248\\_0010](https://doi.org/10.1142/9789812790248_0010).
- [1.3] P.K. Sharma, S. Dorlikar, P. Rawat, V. Malik, N. Vats, M. Sharma, J.S. Rhyee, A.K. Kaushik, Nanotechnology and its application: a review, *Nanotechnology in Cancer Management: Precise Diagnostics toward Personalized Health Care* (2021) 1–33. <https://doi.org/10.1016/B978-0-12-818154-6.00010-X>.
- [1.4] M.L. Cohen, Nanotubes, nanoscience, and nanotechnology, *Materials Science and Engineering: C* 15 (2001) 1–11.
- [1.5] M.T. Bohr, Nanotechnology goals and challenges for electronic applications, *IEEE Trans Nanotechnol* 1 (2002) 56–62. <https://doi.org/10.1109/TNANO.2002.1005426>.
- [1.6] T. Kida, T. Oka, M. Nagano, Y. Ishiwata, X.G. Zheng, Synthesis and Application of Stable Copper Oxide Nanoparticle Suspensions for Nanoparticulate Film Fabrication, *Journal of the American Ceramic Society* 90 (2007) 107–110. <https://doi.org/10.1111/J.1551-2916.2006.01402.X>.
- [1.7] D.O. Hutchins, T. Weidner, J. Baio, B. Polishak, O. Acton, N. Cernetic, H. Ma, A.K.Y. Jen, Effects of self-assembled monolayer structural order, surface homogeneity and surface energy on pentacene morphology and thin film transistor device performance, *J Mater Chem C Mater* 1 (2013) 101–113. <https://doi.org/10.1039/c2tc00378c>.
- [1.8] R.N. Kostoff, R.G. Koytcheff, C.G.Y. Lau, Global nanotechnology research literature overview, *Technol Forecast Soc Change* 74 (2007) 1733–1747. <https://doi.org/10.1016/j.techfore.2007.04.004>.

- [1.9] J. Wen, W. Li, S. Chen, J. Ma, Simulations of molecular self-assembled monolayers on surfaces: Packing structures, formation processes and functions tuned by intermolecular and interfacial interactions, *Physical Chemistry Chemical Physics* 18 (2016) 22757–22771. <https://doi.org/10.1039/c6cp01049k>.
- [1.10] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.
- [1.11] C. George, M. Ammann, B. D’Anna, D.J. Donaldson, S.A. Nizkorodov, Heterogeneous Photochemistry in the Atmosphere, *Chem Rev* 115 (2015) 4218–4258. <https://doi.org/10.1021/cr500648z>.
- [1.12] M. Kapilashrami, Y. Zhang, Y.S. Liu, A. Hagfeldt, J. Guo, Probing the optical property and electronic structure of TiO<sub>2</sub>nanomaterials for renewable energy applications, *Chem Rev* 114 (2014) 9662–9707. <https://doi.org/10.1021/cr5000893>.
- [1.13] A.J. Bard, M.A. Fox, Artificial Photosynthesis: Solar Splitting of Water to Hydrogen and Oxygen, *Acc Chem Res* 28 (1995) 141–145. <https://doi.org/10.1021/ar00051a007>.
- [1.14] Z. Zou, J. Ye, K. Sayama, H. Arakawa, Direct splitting of water under visible light irradiation with an oxide semiconductor photocatalyst, *Nature* 414 (2001) 625–627. <https://doi.org/10.1038/414625a>.
- [1.15] P. Kumar, P. Kar, A.P. Manuel, S. Zeng, U.K. Thakur, K.M. Alam, Y. Zhang, R. Kisslinger, K. Cui, G.M. Bernard, V.K. Michaelis, K. Shankar, Noble Metal Free, Visible Light Driven Photocatalysis Using TiO<sub>2</sub> Nanotube Arrays Sensitized by P-Doped C<sub>3</sub>N<sub>4</sub> Quantum Dots, *Adv Opt Mater* 1901275 (2019) 1–15. <https://doi.org/10.1002/adom.201901275>.
- [1.16] Q. Wang, T. Hisatomi, Q. Jia, H. Tokudome, M. Zhong, C. Wang, Z. Pan, T. Takata, M. Nakabayashi, N. Shibata, Y. Li, I.D. Sharp, A. Kudo, T. Yamada, K. Domen, Scalable water splitting on particulate photocatalyst sheets with a solar-to-hydrogen energy conversion efficiency exceeding 1%, (2016). <https://doi.org/10.1038/NMAT4589>.

- [1.17] J. Wang, T. Heil, B. Zhu, C.W. Tung, J. Yu, H.M. Chen, M. Antonietti, S. Cao, A Single Cu-Center Containing Enzyme-Mimic Enabling Full Photosynthesis under CO<sub>2</sub> Reduction, *ACS Nano* 14 (2020) 8584–8593. <https://doi.org/10.1021/acsnano.0c02940>.
- [1.18] P. Kumar, E. Vahidzadeh, U.K. Thakur, P. Kar, K.M. Alam, A. Goswami, N. Mahdi, K. Cui, G.M. Bernard, V.K. Michaelis, K. Shankar, C<sub>3</sub>N<sub>5</sub>: A Low Bandgap Semiconductor Containing an Azo-Linked Carbon Nitride Framework for Photocatalytic, Photovoltaic and Adsorbent Applications, *J Am Chem Soc* 141 (2019) 5415–5436. <https://doi.org/10.1021/jacs.9b00144>.
- [1.19] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [1.20] Y. Xu, S.P. Gao, Band gap of C<sub>3</sub>N<sub>4</sub> in the GW approximation, *Int J Hydrogen Energy* 37 (2012) 11072–11080. <https://doi.org/10.1016/J.IJHYDENE.2012.04.138>.
- [1.21] S. Cao, J. Low, J. Yu, M. Jaroniec, Polymeric Photocatalysts Based on Graphitic Carbon Nitride, *Advanced Materials* 27 (2015) 2150–2176. <https://doi.org/10.1002/adma.201500033>.
- [1.22] J. Liu, Q. Huang, Z. Chen, L. Yao, L. Huang, C. Dong, L. Niu, Y. Zhang, Electrochemical impedance spectroscopy study on polysulfone/g-C<sub>3</sub>N<sub>4</sub> composite membrane during the separation process, (2020). <https://doi.org/10.5004/dwt.2020.24997>.
- [1.23] Y. Yuan, L. Zhang, J. Xing, M.I.B. Utama, X. Lu, K. Du, Y. Li, X. Hu, S. Wang, A. Genç, R. Dunin-Borkowski, J. Arbiol, Q. Xiong, High-yield synthesis and optical properties of g-C<sub>3</sub>N<sub>4</sub>, *Nanoscale* 7 (2015) 12343–12350. <https://doi.org/10.1039/c5nr02905h>.
- [1.24] Y. Zheng, L. Lin, B. Wang, X. Wang, Graphitic Carbon Nitride Polymers toward Sustainable Photoredox Catalysis, *Angewandte Chemie International Edition* 54 (2015) 12868–12884. <https://doi.org/10.1002/anie.201501788>.
- [1.25] A. Amiri, M. Naraghi, G. Ahmadi, M. Soleymaniha, M. Shanbedi, A review on liquid-phase exfoliation for scalable production of pure graphene, wrinkled, crumpled and

- functionalized graphene and challenges, *FlatChem* 8 (2018) 40–71.  
<https://doi.org/10.1016/J.FLATC.2018.03.004>.
- [1.26] F. Goni, A. Chemelli, F. Uhlig, High-Yield Production of Selected 2D Materials by Understanding Their Sonication-Assisted Liquid-Phase Exfoliation, *Nanomaterials* 2021, Vol. 11, Page 3253 11 (2021) 3253. <https://doi.org/10.3390/NANO11123253>.
- [1.27] Y. Xu, H. Cao, Y. Xue, B. Li, W. Cai, Liquid-Phase Exfoliation of Graphene: An Overview on Exfoliation Media, Techniques, and Challenges, *Nanomaterials* 8 (2018). <https://doi.org/10.3390/NANO8110942>.
- [1.28] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenis, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045.  
[https://doi.org/10.1021/ACSAMI.5B06974/SUPPL\\_FILE/AM5B06974\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.5B06974/SUPPL_FILE/AM5B06974_SI_001.PDF).
- [1.29] C. Backes, D. Hanlon, B.M. Szydłowska, A. Harvey, R.J. Smith, T.M. Higgins, J.N. Coleman, Preparation of Liquid-exfoliated Transition Metal Dichalcogenide Nanosheets with Controlled Size and Thickness: A State of the Art Protocol, *J Vis Exp* 2016 (2016) 54806. <https://doi.org/10.3791/54806>.
- [1.30] M. V. Bracamonte, G.I. Lacconi, S.E. Urreta, L.E.F. Foa Torres, On the nature of defects in liquid-phase exfoliated graphene, *Journal of Physical Chemistry C* 118 (2014) 15455–15459. [https://doi.org/10.1021/JP501930A/SUPPL\\_FILE/JP501930A\\_SI\\_001.PDF](https://doi.org/10.1021/JP501930A/SUPPL_FILE/JP501930A_SI_001.PDF).
- [1.31] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenis, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045. [https://doi.org/10.1021/ACSAMI.5B06974/ASSET/IMAGES/LARGE/AM-2015-06974S\\_0012.JPEG](https://doi.org/10.1021/ACSAMI.5B06974/ASSET/IMAGES/LARGE/AM-2015-06974S_0012.JPEG).



- [1.32] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571.  
<https://doi.org/10.1126/science.1194975>.
- [1.33] J. Zhang, Y. Chen, X. Wang, Two-dimensional covalent carbon nitride nanosheets: Synthesis, functionalization, and applications, *Energy Environ Sci* 8 (2015) 3092–3108.  
<https://doi.org/10.1039/c5ee01895a>.
- [1.34] X. Zhang, X. Xie, H. Wang, J. Zhang, B. Pan, Y. Xie, Enhanced photoresponsive ultrathin graphitic-phase C<sub>3</sub>N<sub>4</sub> nanosheets for bioimaging, *J Am Chem Soc* 135 (2013) 18–21. <https://doi.org/10.1021/ja308249k>.
- [1.35] S. Yang, Y. Gong, J. Zhang, L. Zhan, L. Ma, Z. Fang, R. Vajtai, X. Wang, P.M. Ajayan, Exfoliated graphitic carbon nitride nanosheets as efficient catalysts for hydrogen evolution under visible light, *Advanced Materials* 25 (2013) 2452–2456.  
<https://doi.org/10.1002/adma.201204453>.
- [1.36] J. Shen, Y. He, J. Wu, C. Gao, K. Keyshar, X. Zhang, Y. Yang, M. Ye, R. Vajtai, J. Lou, P.M. Ajayan, Liquid Phase Exfoliation of Two-Dimensional Materials by Directly Probing and Matching Surface Tension Components, *Nano Lett* 15 (2015) 5449–5454.  
[https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E\\_0005.JPEG](https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E_0005.JPEG).
- [1.37] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [1.38] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906.

[https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL\\_FILE/LA7B02441\\_SI\\_007.AVI](https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL_FILE/LA7B02441_SI_007.AVI).

- [1.39] C.Y. Wong, B.L. Cotts, H. Wu, N.S. Ginsberg, Exciton dynamics reveal aggregates with intermolecular order at hidden interfaces in solution-cast organic semiconducting films, *Nat Commun* 6 (2015) 1–7. <https://doi.org/10.1038/ncomms6946>.
- [1.40] T.K. Mukhopadhyay, A. Datta, Disentangling the liquid phase exfoliation of two-dimensional materials: an “in silico” perspective, *Physical Chemistry Chemical Physics* 22 (2020) 22157–22179. <https://doi.org/10.1039/D0CP03128C>.
- [1.41] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648. <https://doi.org/10.1021/ja1064284>.
- [1.42] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268. <https://doi.org/10.1021/acsnano.5b02683>.
- [1.43] V. Sresht, A. Govind Rajan, E. Bordes, M.S. Strano, A.A.H. Pádua, D. Blankschtein, Quantitative Modeling of MoS<sub>2</sub>-Solvent Interfaces: Predicting Contact Angles and Exfoliation Performance using Molecular Dynamics, *Journal of Physical Chemistry C* 121 (2017) 9022–9031. [https://doi.org/10.1021/ACS.JPCC.7B00484/SUPPL\\_FILE/JP7B00484\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCC.7B00484/SUPPL_FILE/JP7B00484_SI_001.PDF).
- [1.44] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.
- [1.45] A. Gotzias, Y.G. Lazarou, Graphene Exfoliation in Binary NMP/Water Mixtures by Molecular Dynamics Simulations, *Chempluschem* 89 (2024) e202300758. <https://doi.org/10.1002/CPLU.202300758>.
- [1.46] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight

- from molecular dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954.  
<https://doi.org/10.1007/s10934-021-01047-7>.
- [2.1] C.L. Brooks, Computer simulation of liquids, *J Solution Chem* 18 (1989) 99–99.  
<https://doi.org/10.1007/BF00646086>.
- [2.2] H. Kamberaj, *Molecular Dynamics Simulations in Statistical Physics: Theory and Applications*, Springer International Publishing, Cham, 2020. <https://doi.org/10.1007/978-3-030-35702-3>.
- [2.3] M.E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, 2023.
- [2.4] D. Frenkel, B. Smit, *Understanding Molecular Simulation: Monte Carlo Simulations in Various Ensembles*, Academic Press, London (2002) 111–137.  
<https://www.sciencedirect.com/science/article/pii/B9780122673511500079> (accessed August 5, 2024).
- [2.5] W.F. Van Gunsteren, H.J.C. Berendsen, A Leap-frog Algorithm for Stochastic Dynamics, *Mol Simul* 1 (1988) 173–185. <https://doi.org/10.1080/08927028808080941>.
- [2.6] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *Journal of Chemical Physics* 126 (2007). <https://doi.org/10.1063/1.2408420>/186581.
- [2.7] M. Bernetti, G. Bussi, Pressure control using stochastic cell rescaling, *Journal of Chemical Physics* 153 (2020). <https://doi.org/10.1063/5.0020514>/199610.
- [2.8] W. Damm, A. Frontera, J. Tirado-Rives, W.L. Jorgensen, OPLS all-atom force field for carbohydrates, *J Comput Chem* 18 (1997) 1955–1970. [https://doi.org/10.1002/\(SICI\)1096-987X\(199712\)18:16<1955::AID-JCC1>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1096-987X(199712)18:16<1955::AID-JCC1>3.0.CO;2-L).
- [2.9] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J Chem Phys* 103 (1995) 8577–8593.  
<https://doi.org/10.1063/1.470117>.

- [2.10] R.P. Futrelle, D.J. McGinty, Calculation of spectra and correlation functions from molecular dynamics data using the fast Fourier transform, *Chem Phys Lett* 12 (1971) 285–287. [https://doi.org/10.1016/0009-2614\(71\)85065-0](https://doi.org/10.1016/0009-2614(71)85065-0).
- [2.11] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [2.12] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [2.13] N.E. Jackson, B.M. Savoie, A. Statt, M.A. Webb, Introduction to Machine Learning for Molecular Simulation, *J Chem Theory Comput* 19 (2023) 4335–4337. <https://doi.org/10.1021/ACS.JCTC.3C00735>.
- [2.14] E. Prašnikar, M. Ljubič, A. Perdih, J. Borišek, Machine learning heralding a new development phase in molecular dynamics simulations, *Artificial Intelligence Review* 2024 57:4 57 (2024) 1–36. <https://doi.org/10.1007/S10462-024-10731-4>.
- [2.15] L. Breiman, Random Forests, *Machine Learning* 2001 45:1 45 (2001) 5–32. <https://doi.org/10.1023/A:1010933404324>.
- [2.16] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, *Machine Learning* 2006 63:1 63 (2006) 3–42. <https://doi.org/10.1007/S10994-006-6226-1>.
- [2.17] H.D.- ICML, undefined 1997, Improving regressors using boosting techniques, Citeseer (n.d.). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.21.5683&rep=rep1&type=pdf> (accessed August 8, 2022).
- [2.18] M.T.-J. of machine learning research, undefined 2001, Sparse Bayesian learning and the relevance vector machine, *Jmlr.Org* (n.d.). <https://www.jmlr.org/papers/volume1/tipping01a/tipping01a.pdf?ref=https://githubhelp.com> (accessed August 8, 2022).

- [2.19] W. Bin How, B. Wang, W. Chu, A. Tkatchenko, O. V. Prezhdo, Significance of the Chemical Environment of an Element in Nonadiabatic Molecular Dynamics: Feature Selection and Dimensionality Reduction with Machine Learning, *Journal of Physical Chemistry Letters* 12 (2021) 12026–12032.  
[https://doi.org/10.1021/ACS.JPCLETT.1C03469/SUPPL\\_FILE/JZ1C03469\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCLETT.1C03469/SUPPL_FILE/JZ1C03469_SI_001.PDF).
- [2.20] M.A. Hall, L.A. Smith, Practical feature subset selection for machine learning, *Computer Science* Volume 20 No 1 (1998) 181–191.  
<https://researchcommons.waikato.ac.nz/handle/10289/1512> (accessed August 13, 2022).
- [2.21] D.M. Belete, M.D. Huchaiah, Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results, *International Journal of Computers and Applications* 44 (2022) 875–886.  
<https://doi.org/10.1080/1206212X.2021.1974663>.
- [3.1] J. Fu, J. Yu, C. Jiang, B. Cheng, g-C<sub>3</sub>N<sub>4</sub>-Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.
- [3.2] X. Ma, J. Hu, H. He, S. Dong, C. Huang, X. Chen, New Understanding on Enhanced Photocatalytic Activity of g-C<sub>3</sub>N<sub>4</sub>/BiPO<sub>4</sub> Heterojunctions by Effective Interfacial Coupling, *ACS Appl Nano Mater* 1 (2018) 5507–5515.  
<https://doi.org/10.1021/acsanm.8b01012>.
- [3.3] Y. Yuan, L. Zhang, J. Xing, M.I.B. Utama, X. Lu, K. Du, Y. Li, X. Hu, S. Wang, A. Genç, R. Dunin-Borkowski, J. Arbiol, Q. Xiong, High-yield synthesis and optical properties of g-C<sub>3</sub>N<sub>4</sub>, *Nanoscale* 7 (2015) 12343–12350. <https://doi.org/10.1039/c5nr02905h>.
- [3.4] J. Xu, L. Zhang, R. Shi, Y. Zhu, Chemical exfoliation of graphitic carbon nitride for efficient heterogeneous photocatalysis, *J Mater Chem A Mater* 1 (2013) 14766–14772.  
<https://doi.org/10.1039/C3TA13188B>.
- [3.5] Y. Liu, D. Xie, M. Song, L. Jiang, G. Fu, L. Liu, J. Li, Water desalination across multilayer graphitic carbon nitride membrane: Insights from non-equilibrium molecular dynamics simulations, *Carbon N Y* 140 (2018) 131–138.  
<https://doi.org/10.1016/j.carbon.2018.08.043>.

- [3.6] A. Alaghmandfard, K. Ghandi, A Comprehensive Review of Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)–Metal Oxide-Based Nanocomposites: Potential for Photocatalysis and Sensing, *Nanomaterials* 2022, Vol. 12, Page 294 12 (2022) 294.  
<https://doi.org/10.3390/NANO12020294>.
- [3.7] Y. Dong, Q. Wang, H. Wu, Y. Chen, C.H. Lu, Y. Chi, H.H. Yang, Graphitic Carbon Nitride Materials: Sensing, Imaging and Therapy, *Small* 12 (2016) 5376–5393.  
<https://doi.org/10.1002/SMLL.201602056>.
- [3.8] S. Patnaik, D.P. Sahoo, K. Parida, An overview on Ag modified g-C<sub>3</sub>N<sub>4</sub> based nanostructured materials for energy and environmental applications, *Renewable and Sustainable Energy Reviews* 82 (2018) 1297–1312.  
<https://doi.org/10.1016/J.RSER.2017.09.026>.
- [3.9] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [3.10] X. Dong, F. Cheng, Recent development in exfoliated two-dimensional g-C<sub>3</sub>N<sub>4</sub> nanosheets for photocatalytic applications, *J Mater Chem A Mater* 3 (2015) 23642–23652.  
<https://doi.org/10.1039/c5ta07374j>.
- [3.11] S. Cao, J. Low, J. Yu, M. Jaroniec, Polymeric Photocatalysts Based on Graphitic Carbon Nitride, *Advanced Materials* 27 (2015) 2150–2176.  
<https://doi.org/10.1002/adma.201500033>.
- [3.12] Y. Zheng, L. Lin, B. Wang, X. Wang, Graphitic Carbon Nitride Polymers toward Sustainable Photoredox Catalysis, *Angewandte Chemie International Edition* 54 (2015) 12868–12884. <https://doi.org/10.1002/anie.201501788>.
- [3.13] K.M. Alam, C.E. Jensen, P. Kumar, R.W. Hooper, G.M. Bernard, A. Patidar, A.P. Manuel, N. Amer, A. Palmgren, D.N. Purschke, N. Chaulagain, J. Garcia, P.S. Kirwin, L.C.T. Shoute, K. Cui, S. Gusarov, A.E. Kobryn, V.K. Michaelis, F.A. Hegmann, K. Shankar, Photocatalytic Mechanism Control and Study of Carrier Dynamics in CdS@C<sub>3</sub>N<sub>5</sub>Core-Shell Nanowires, *ACS Appl Mater Interfaces* 13 (2021) 47418–47439.

[https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550\\_0011.JPEG](https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550_0011.JPEG).

- [3.14] S. Angizi, M. Ali Akbar, M. Darestani-Farahani, al -, S. Dhanraj Nehate, S. Sundares, R. Peale, P. Kumar, S. Mulmi, D. Laishram, K.M. Alam, U.K. Thakur, V. Thangadurai, K. Shankar, Water-splitting photoelectrodes consisting of heterojunctions of carbon nitride with a p-type low bandgap double perovskite oxide, *Nanotechnology* 32 (2021) 485407. <https://doi.org/10.1088/1361-6528/ABEDEC>.
- [3.15] S. Cao, H. Chen, F. Jiang, Z. Hu, X. Wang, Construction of Acetaldehyde-Modified g-C<sub>3</sub>N<sub>4</sub> Ultrathin Nanosheets via Ethylene Glycol-Assisted Liquid Exfoliation for Selective Fluorescence Sensing of Ag<sup>+</sup>, *ACS Appl Mater Interfaces* 10 (2018) 44624–44633. [https://doi.org/10.1021/ACSAMI.8B15501/ASSET/IMAGES/LARGE/AM-2018-15501V\\_0002.JPEG](https://doi.org/10.1021/ACSAMI.8B15501/ASSET/IMAGES/LARGE/AM-2018-15501V_0002.JPEG).
- [3.16] B. Abreu, B. Almeida, P. Ferreira, R. M. F. Fernandes, D.M. Fernandes, E.F. Marques, A critical assessment of the role of ionic surfactants in the exfoliation and stabilization of 2D nanosheets: The case of the transition metal dichalcogenides MoS<sub>2</sub>, WS<sub>2</sub> and MoSe<sub>2</sub>, *J Colloid Interface Sci* 626 (2022) 167–177. <https://doi.org/10.1016/J.JCIS.2022.06.097>.
- [3.17] H. Zhao, H. Wu, J. Wu, J. Li, Y. Wang, Y. Zhang, H. Liu, Preparation of MoS<sub>2</sub>/WS<sub>2</sub> nanosheets by liquid phase exfoliation with assistance of epigallocatechin gallate and study as an additive for high-performance lithium-sulfur batteries, *J Colloid Interface Sci* 552 (2019) 554–562. <https://doi.org/10.1016/J.JCIS.2019.05.080>.
- [3.18] Q. Wan, H. Wang, S. Li, J. Wang, Efficient liquid-phase exfoliation of few-layer graphene in aqueous 1, 1, 3, 3-tetramethylurea solution, *J Colloid Interface Sci* 526 (2018) 167–173. <https://doi.org/10.1016/J.JCIS.2018.04.110>.
- [3.19] C. Huo, Z. Yan, X. Song, H. Zeng, 2D materials via liquid exfoliation: a review on fabrication and applications, *Sci Bull (Beijing)* 60 (2015) 1994–2008. <https://doi.org/10.1007/S11434-015-0936-3>.

- [3.20] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon* N Y 141 (2019) 580–607. <https://doi.org/10.1016/J.CARBON.2018.09.082>.
- [3.21] Y. Hernandez, M. Lotya, D. Rickard, S.D. Bergin, J.N. Coleman, Measurement of multicomponent solubility parameters for graphene facilitates solvent discovery, *Langmuir* 26 (2010) 3208–3213. <https://doi.org/10.1021/la903188a>.
- [3.22] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. v. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/SCIENCE.1194975>.
- [3.23] L. Ma, H. Fan, J. Wang, Y. Zhao, H. Tian, G. Dong, Water-assisted ions in situ intercalation for porous polymeric graphitic carbon nitride nanosheets with superior photocatalytic hydrogen evolution performance, *Appl Catal B* 190 (2016) 93–102. <https://doi.org/10.1016/j.apcatb.2016.03.002>.
- [3.24] C. Wu, S. Xue, Z. Qin, M. Nazari, G. Yang, S. Yue, T. Tong, H. Ghasemi, F.C.R. Hernandez, S. Xue, D. Zhang, H. Wang, Z.M. Wang, S. Pu, J. Bao, Making g-C<sub>3</sub>N<sub>4</sub> ultra-thin nanosheets active for photocatalytic overall water splitting, *Appl Catal B* 282 (2021) 119557. <https://doi.org/10.1016/j.apcatb.2020.119557>.
- [3.25] S.P. Pattnaik, A. Behera, S. Martha, R. Acharya, K. Parida, Facile synthesis of exfoliated graphitic carbon nitride for photocatalytic degradation of ciprofloxacin under solar irradiation, *J Mater Sci* 54 (2019) 5726–5742. <https://doi.org/10.1007/s10853-018-03266-x>.
- [3.26] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.



- [3.27] M. Ayán-Varela, S. Villar-Rodil, J.I. Paredes, J.M. Munuera, A. Pagán, A.A. Lozano-Pérez, J.L. Cenis, A. Martínez-Alonso, J.M.D. Tascón, Investigating the Dispersion Behavior in Solvents, Biocompatibility, and Use as Support for Highly Efficient Metal Catalysts of Exfoliated Graphitic Carbon Nitride, *ACS Appl Mater Interfaces* 7 (2015) 24032–24045. <https://doi.org/10.1021/acsami.5b06974>.
- [3.28] S. Yang, Y. Gong, J. Zhang, L. Zhan, L. Ma, Z. Fang, R. Vajtai, X. Wang, P.M. Ajayan, Exfoliated graphitic carbon nitride nanosheets as efficient catalysts for hydrogen evolution under visible light, *Advanced Materials* 25 (2013) 2452–2456. <https://doi.org/10.1002/adma.201204453>.
- [3.29] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [3.30] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuvsen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/science.1194975>.
- [3.31] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight from molecular dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954. <https://doi.org/10.1007/s10934-021-01047-7>.
- [3.32] Y. Zhang, C. Shen, X. Lu, X. Mu, P. Song, Effects of defects in g-C<sub>3</sub>N<sub>4</sub> on excited-state charge distribution and transfer: Potential for improved photocatalysis, *Spectrochim Acta A Mol Biomol Spectrosc* 227 (2020) 117687. <https://doi.org/10.1016/J.SAA.2019.117687>.
- [3.33] X. Zhou, M. Zhu, L. Kang, Single-Atom X/g-C<sub>3</sub>N<sub>4</sub> (X = Au<sup>1</sup>, Pd<sup>1</sup>, and Ru<sup>1</sup>) Catalysts for Acetylene Hydrochlorination: A Density Functional Theory Study, *Catalysts* 2019, Vol. 9, Page 808 9 (2019) 808. <https://doi.org/10.3390/CATAL9100808>.

- [3.34] X. Chen, R. Hu, DFT-based study of single transition metal atom doped g-C<sub>3</sub>N<sub>4</sub> as alternative oxygen reduction reaction catalysts, *Int J Hydrogen Energy* 44 (2019) 15409–15416. <https://doi.org/10.1016/J.IJHYDENE.2019.04.057>.
- [3.35] Z. Bonakchi, A. Nakhaei Pour, S. Soheili, Molecular simulation of methane on various g-C<sub>3</sub>N<sub>4</sub> isomers: collision, adsorption, desorption, and diffusion studies, *Journal of the Iranian Chemical Society* 19 (2022) 3649–3657. <https://doi.org/10.1007/S13738-022-02562-3/TABLES/2>.
- [3.36] L. Kang, M. Zhu, Y. Zhao, A DFT Study of Acetylene Hydrogenation Catalyzed by S-Doped Pd 1 /g-C 3 N 4, (n.d.). <https://doi.org/10.3390/catal9110887>.
- [3.37] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906. <https://doi.org/10.1021/acs.langmuir.7b02441>.
- [3.38] W.L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, *Proc Natl Acad Sci U S A* 102 (2005) 6665–6670. <https://doi.org/10.1073/pnas.0408037102>.
- [3.39] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017). <https://doi.org/10.1093/nar/gkx312>.
- [3.40] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a. <https://doi.org/10.2477/jccjie.2018-0034>.
- [3.41] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [3.42] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).

- [3.43] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [3.44] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268. <https://doi.org/10.1021/acsnano.5b02683>.
- [3.45] A. Gotzias, Binding Free Energy Calculations of Bilayer Graphenes Using Molecular Dynamics, *J Chem Inf Model* 61 (2021) 1164–1171. <https://doi.org/10.1021/acs.jcim.1c00043>.
- [3.46] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.
- [3.47] C. Fu, X. Yang, Molecular simulation of interfacial mechanics for solvent exfoliation of graphene from graphite, *Carbon N Y* 55 (2013) 350–360. <https://doi.org/10.1016/j.carbon.2012.12.083>.
- [3.48] G. Zhou, P. Rajak, S. Susarla, P.M. Ajayan, R.K. Kalia, A. Nakano, P. Vashishta, Molecular Simulation of MoS<sub>2</sub> Exfoliation, *Sci Rep* 8 (2018) 1–9. <https://doi.org/10.1038/s41598-018-35008-z>.
- [3.49] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).
- [3.50] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1–2 (2015) 19–25. <https://doi.org/10.1016/J.SOFTX.2015.06.001>.
- [3.51] V. Georgakilas, J.N. Tiwari, K.C. Kemp, J.A. Perman, A.B. Bourlinos, K.S. Kim, R. Zboril, Noncovalent Functionalization of Graphene and Graphene Oxide for Energy Materials, Biosensing, Catalytic, and Biomedical Applications, *Chem Rev* 116 (2016) 5464–5519. <https://doi.org/10.1021/ACS.CHEMREV.5B00620>.

- [3.52] A.M. Silva, M.I. Rojas, Electric and structural properties of polymeric graphite carbon nitride (g-C<sub>3</sub>N<sub>4</sub>): A Density Functional Theory study, *Comput Theor Chem* (2016). <https://doi.org/10.1016/j.comptc.2016.11.004>.
- [3.53] B. Mortazavi, Ultra high stiffness and thermal conductivity of graphene like C<sub>3</sub>N, *Carbon N Y* 118 (2017) 25–34. <https://doi.org/10.1016/J.CARBON.2017.03.029>.
- [3.54] D. Wu, D.A. Kofke, Phase-space overlap measures. II. Design and implementation of staging methods for free-energy calculations, *J Chem Phys* 123 (2005) 084109. <https://doi.org/10.1063/1.2011391>.
- [3.55] H.K. Choi, Y. Oh, H. Jung, H. Hong, B.C. Ku, N.H. You, Y.K. Kim, E.S. Shin, J. Yu, Influences of carboxyl functionalization of intercalators on exfoliation of graphite oxide: A molecular dynamics simulation, *Physical Chemistry Chemical Physics* 20 (2018) 28616–28622. <https://doi.org/10.1039/c8cp05436c>.
- [3.56] H.A. Yu, M. Karplus, A thermodynamic analysis of solvation, *J Chem Phys* 89 (1998) 2366. <https://doi.org/10.1063/1.455080>.
- [3.57] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648. <https://doi.org/10.1021/ja1064284>.
- [3.58] R. Biswas, Molecular dynamics studies on the exfoliation of graphene in room temperature ionic liquids, *J Mol Liq* 337 (2021) 116592. <https://doi.org/10.1016/j.molliq.2021.116592>.
- [3.59] A. Luzar, D. Chandler, Hydrogen-bond kinetics in liquid water, *Nature* 1996 379:6560 379 (1996) 55–57. <https://doi.org/10.1038/379055a0>.
- [3.60] M. Shakourian-Fard, G. Kamath, Z. Jamshidi, Trends in Physisorption of Ionic Liquids on Boron-Nitride Sheets, (2014). <https://doi.org/10.1021/jp506277n>.
- [3.61] T. Chen, M. Li, J. Liu,  $\pi$ - $\pi$  Stacking Interaction: A Nondestructive and Facile Means in Material Engineering for Bioapplications, *Cryst Growth Des* 18 (2018) 2765–2783.

[https://doi.org/10.1021/ACS.CGD.7B01503/ASSET/IMAGES/LARGE/CG-2017-01503A\\_0004.JPEG](https://doi.org/10.1021/ACS.CGD.7B01503/ASSET/IMAGES/LARGE/CG-2017-01503A_0004.JPEG).

- [3.62] C.A. Hunter, J.K.M. Sanders, The nature of  $\pi$ - $\pi$  interactions, *J Am Chem Soc* 112 (2002) 5525–5534. <https://doi.org/10.1021/JA00170A016>.
- [3.63] A.L. Ringer, M.O. Sinnokrot, R.P. Lively, C.D. Sherrill, The effect of multiple substituents on sandwich and T-shaped  $\pi$ - $\pi$  interactions, *Chemistry* 12 (2006) 3821–3828. <https://doi.org/10.1002/CHEM.200501316>.
- [3.64] M. J. Rashkin, M. L. Waters, Unexpected Substituent Effects in Offset  $\pi$ - $\pi$  Stacked Interactions in Water, *J Am Chem Soc* 124 (2002) 1860–1861. <https://doi.org/10.1021/ja016508z>.
- [3.65] M. Peleg, M.D. Normand, M.G. Corradini, The Arrhenius Equation Revisited, <https://doi.org/10.1080/10408398.2012.667460> 52 (2012) 830–851. <https://doi.org/10.1080/10408398.2012.667460>.
- [3.66] L. Sun, Y. Qi, C.J. Jia, Z. Jin, W. Fan, Enhanced visible-light photocatalytic activity of g-C<sub>3</sub>N<sub>4</sub>/Zn<sub>2</sub>GeO<sub>4</sub> heterojunctions with effective interfaces based on band match, *Nanoscale* 6 (2014) 2649–2659. <https://doi.org/10.1039/C3NR06104C>.
- [3.67] E. Bordes, B. Morcos, D. Bourgoigne, J.M. Andanson, P.O. Bussière, C.C. Santini, A. Benayad, M.C. Gomes, A.A.H. Pádua, Dispersion and stabilization of exfoliated graphene in ionic liquids, *Front Chem* 7 (2019) 223. <https://doi.org/10.3389/FCHEM.2019.00223/BIBTEX>.
- [3.68] Y.J. Yuan, Z. Shen, S. Wu, Y. Su, L. Pei, Z. Ji, M. Ding, W. Bai, Y. Chen, Z.T. Yu, Z. Zou, Liquid exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets to construct 2D-2D MoS<sub>2</sub>/g-C<sub>3</sub>N<sub>4</sub> photocatalyst for enhanced photocatalytic H<sub>2</sub> production activity, *Appl Catal B* 246 (2019) 120–128. <https://doi.org/10.1016/j.apcatb.2019.01.043>.
- [3.69] Q. Hao, Y. Song, H. Ji, Z. Mo, X. She, J. Deng, T. Muhmood, X. Wu, S. Yuan, H. Xu, H. Li, Surface N modified 2D g-C<sub>3</sub>N<sub>4</sub> nanosheets derived from DMF for photocatalytic H<sub>2</sub> evolution, *Appl Surf Sci* 459 (2018) 845–852. <https://doi.org/10.1016/j.apsusc.2018.07.154>.

- [3.70] Y. Wang, L. Li, Y. Wei, J. Xue, H. Chen, L. Ding, J. Caro, H. Wang, Water Transport with Ultralow Friction through Partially Exfoliated g-C<sub>3</sub>N<sub>4</sub> Nanosheet Membranes with Self-Supporting Spacers, *Angewandte Chemie - International Edition* 56 (2017) 8974–8980. <https://doi.org/10.1002/anie.201701288>.
- [3.71] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [3.72] L. Ma, H. Fan, M. Li, H. Tian, J. Fang, G. Dong, A simple melamine-assisted exfoliation of polymeric graphitic carbon nitrides for highly efficient hydrogen production from water under visible light, *J Mater Chem A Mater* 3 (2015) 22404–22412. <https://doi.org/10.1039/C5TA05850C>.
- [3.73] K. Koschek, V. Durmaz, O. Krylova, M. Wieczorek, S. Gupta, M. Richter, A. Bujotzek, C. Fischer, R. Haag, C. Freund, M. Weber, J. Rademann, Peptide–polymer ligands for a tandem WW-domain, an adaptive multivalent protein–protein interaction: lessons on the thermodynamic fitness of flexible ligands, *Beilstein Journal of Organic Chemistry* 11:93 11 (2015) 837–847. <https://doi.org/10.3762/BJOC.11.93>.
- [3.74] F. Jiménez-Ángeles, H.-K. Kwon, K. Sadman, T. Wu, K. R. Shull, M. Olvera de la Cruz, Self-Assembly of Charge-Containing Copolymers at the Liquid–Liquid Interface, *ACS Cent Sci* 5 (2019) 688–699. <https://doi.org/10.1021/acscentsci.9b00084>.
- [4.1] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [4.2] G.Q. Zhao, J. Zou, J. Hu, X. Long, F.P. Jiao, A critical review on graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>)-based composites for environmental remediation, *Sep Purif Technol* 279 (2021) 119769. <https://doi.org/10.1016/J.SEPPUR.2021.119769>.
- [4.3] Y. Liu, D. Xie, M. Song, L. Jiang, G. Fu, L. Liu, J. Li, Water desalination across multilayer graphitic carbon nitride membrane: Insights from non-equilibrium molecular dynamics

- simulations, *Carbon* N Y 140 (2018) 131–138.  
<https://doi.org/10.1016/j.carbon.2018.08.043>.
- [4.4] A. Alaghmandfard, K. Ghandi, A Comprehensive Review of Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)–Metal Oxide-Based Nanocomposites: Potential for Photocatalysis and Sensing, *Nanomaterials* 2022, Vol. 12, Page 294 12 (2022) 294.  
<https://doi.org/10.3390/NANO12020294>.
- [4.5] Y. Dong, Q. Wang, H. Wu, Y. Chen, C.H. Lu, Y. Chi, H.H. Yang, Graphitic Carbon Nitride Materials: Sensing, Imaging and Therapy, *Small* 12 (2016) 5376–5393.  
<https://doi.org/10.1002/SMLL.201602056>.
- [4.6] L. Bai, H. Huang, S. Yu, D. Zhang, H. Huang, Y. Zhang, Role of transition metal oxides in g-C<sub>3</sub>N<sub>4</sub>-based heterojunctions for photocatalysis and supercapacitors, *Journal of Energy Chemistry* 64 (2022) 214–235. <https://doi.org/10.1016/J.JECHEM.2021.04.057>.
- [4.7] K.M. Alam, N. Chaulagain, E. Shahini, M. Masud Rana, J. Garcia, N. Kumar, A.E. Kobryn, S. Gusarov, T. Tang, K. Shankar, Low bandgap carbon nitride nanoparticles incorporated in titania nanotube arrays by in situ electrophoretic anodization for photocatalytic CO<sub>2</sub> reduction, *Chemical Engineering Journal* 456 (2023) 141067.  
<https://doi.org/10.1016/J.CEJ.2022.141067>.
- [4.8] K.M. Alam, C.E. Jensen, P. Kumar, R.W. Hooper, G.M. Bernard, A. Patidar, A.P. Manuel, N. Amer, A. Palmgren, D.N. Purschke, N. Chaulagain, J. Garcia, P.S. Kirwin, L.C.T. Shoute, K. Cui, S. Gusarov, A.E. Kobryn, V.K. Michaelis, F.A. Hegmann, K. Shankar, Photocatalytic Mechanism Control and Study of Carrier Dynamics in CdS@C<sub>3</sub>N<sub>5</sub>Core-Shell Nanowires, *ACS Appl Mater Interfaces* 13 (2021) 47418–47439.  
[https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550\\_0011.JPEG](https://doi.org/10.1021/ACSAMI.1C08550/ASSET/IMAGES/LARGE/AM1C08550_0011.JPEG).
- [4.9] S. Angizi, M. Ali Akbar, M. Darestani-Farahani, al -, S. Dhanraj Nehate, S. Sundaresh, R. Peale, P. Kumar, S. Mulmi, D. Laishram, K.M. Alam, U.K. Thakur, V. Thangadurai, K. Shankar, Water-splitting photoelectrodes consisting of heterojunctions of carbon nitride with a p-type low bandgap double perovskite oxide, *Nanotechnology* 32 (2021) 485407.  
<https://doi.org/10.1088/1361-6528/ABEDEC>.

- [4.10] B. Abreu, B. Almeida, P. Ferreira, R. M. F. Fernandes, D.M. Fernandes, E.F. Marques, A critical assessment of the role of ionic surfactants in the exfoliation and stabilization of 2D nanosheets: The case of the transition metal dichalcogenides MoS<sub>2</sub>, WS<sub>2</sub> and MoSe<sub>2</sub>, *J Colloid Interface Sci* 626 (2022) 167–177. <https://doi.org/10.1016/J.JCIS.2022.06.097>.
- [4.11] I. Benabdallah, A. Kara, M. Benaissa, Exfoliation and re-aggregation mechanisms of black phosphorus: A molecular dynamics study, *Appl Surf Sci* 507 (2020). <https://doi.org/10.1016/j.apsusc.2019.144826>.
- [4.12] R. Biswas, Molecular dynamics studies on the exfoliation of graphene in room temperature ionic liquids, *J Mol Liq* 337 (2021) 116592. <https://doi.org/10.1016/j.molliq.2021.116592>.
- [4.13] T.K. Mukhopadhyay, A. Datta, Deciphering the role of solvents in the liquid phase exfoliation of hexagonal boron nitride: A molecular dynamics simulation study, *Journal of Physical Chemistry C* 121 (2017) 811–822. <https://doi.org/10.1021/acs.jpcc.6b09446>.
- [4.14] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268. <https://doi.org/10.1021/acsnano.5b02683>.
- [4.15] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648. <https://doi.org/10.1021/ja1064284>.
- [4.16] G. Zhou, P. Rajak, S. Susarla, P.M. Ajayan, R.K. Kalia, A. Nakano, P. Vashishta, Molecular Simulation of MoS<sub>2</sub> Exfoliation, *Sci Rep* 8 (2018) 1–9. <https://doi.org/10.1038/s41598-018-35008-z>.
- [4.17] Y. Hernandez, M. Lotya, D. Rickard, S.D. Bergin, J.N. Coleman, Measurement of multicomponent solubility parameters for graphene facilitates solvent discovery, *Langmuir* 26 (2010) 3208–3213. <https://doi.org/10.1021/la903188a>.
- [4.18] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. v. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T.



- Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571.  
<https://doi.org/10.1126/SCIENCE.1194975>.
- [4.19] C. Wu, S. Xue, Z. Qin, M. Nazari, G. Yang, S. Yue, T. Tong, H. Ghasemi, F.C.R. Hernandez, S. Xue, D. Zhang, H. Wang, Z.M. Wang, S. Pu, J. Bao, Making g-C<sub>3</sub>N<sub>4</sub> ultra-thin nanosheets active for photocatalytic overall water splitting, *Appl Catal B* 282 (2021) 119557. <https://doi.org/10.1016/j.apcatb.2020.119557>.
- [4.20] S.P. Pattnaik, A. Behera, S. Martha, R. Acharya, K. Parida, Facile synthesis of exfoliated graphitic carbon nitride for photocatalytic degradation of ciprofloxacin under solar irradiation, *J Mater Sci* 54 (2019) 5726–5742. <https://doi.org/10.1007/s10853-018-03266-x>.
- [4.21] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910. <https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [4.22] J.S. Delaney, ESOL: Estimating aqueous solubility directly from molecular structure, *J Chem Inf Comput Sci* 44 (2004) 1000–1005.  
[https://doi.org/10.1021/CI034243X/SUPPL\\_FILE/CI034243XSI20040112\\_053635.TXT](https://doi.org/10.1021/CI034243X/SUPPL_FILE/CI034243XSI20040112_053635.TXT).
- [4.23] C. Reichardt, T. Welton, *Solvents and Solvent Effects in Organic Chemistry: Fourth Edition*, Solvents and Solvent Effects in Organic Chemistry: Fourth Edition (2010).  
<https://doi.org/10.1002/9783527632220>.
- [4.24] A.P. Sgouros, C.J. Revelas, A.T. Lakkas, D.N. Theodorou, Solvation Free Energy of Dilute Grafted (Nano)Particles in Polymer Melts via the Self-Consistent Field Theory, *J Phys Chem B* (2022). <https://doi.org/10.1021/ACS.JPCB.2C05306>.
- [4.25] L. Pereyaslavets, G. Kamath, O. Butin, A. Illarionov, M. Olevanov, I. Kurnikov, S. Sakipov, I. Leontyev, E. Voronina, T. Gannon, G. Nawrocki, M. Darkhovskiy, I. Ivahnenko, A. Kostikov, J. Scaranto, M.G. Kurnikova, S. Banik, H. Chan, M.G. Sternberg, S.K.R.S. Sankaranarayanan, B. Crawford, J. Potoff, M. Levitt, R.D. Kornberg, B. Fain,

- Accurate determination of solvation free energies of neutral organic compounds from first principles, *Nature Communications* 2022 13:1 13 (2022) 1–7.  
<https://doi.org/10.1038/s41467-022-28041-0>.
- [4.26] F.H. Vermeire, W.H. Green, Transfer learning for solvation free energies: From quantum chemistry to experiments, *Chemical Engineering Journal* 418 (2021) 129307.  
<https://doi.org/10.1016/J.CEJ.2021.129307>.
- [4.27] Z. Wan, Q. de Wang, D. Liu, J. Liang, Accelerating the optimization of enzyme-catalyzed synthesis conditions via machine learning and reactivity descriptors, *Org Biomol Chem* 19 (2021) 6267–6273. <https://doi.org/10.1039/D1OB01066B>.
- [4.28] S. Perumal, R. Atchudan, T.N.J. Immanuel Edison, J.J. Shim, Y.R. Lee, Exfoliation and Noncovalent Functionalization of Graphene Surface with Poly-N-Vinyl-2-Pyrrolidone by In Situ Polymerization, *Molecules* 2021, Vol. 26, Page 1534 26 (2021) 1534.  
<https://doi.org/10.3390/MOLECULES26061534>.
- [4.29] Y. Saito, K. Shin, K. Terayama, S. Desai, M. Onga, Y. Nakagawa, Y.M. Itahashi, Y. Iwasa, M. Yamada, K. Tsuda, Deep-learning-based quality filtering of mechanically exfoliated 2D crystals, *Npj Computational Materials* 2019 5:1 5 (2019) 1–6.  
<https://doi.org/10.1038/s41524-019-0262-4>.
- [4.30] E.M.D. Siriwardane, R.P. Joshi, N. Kumar, D. Çakır, Revealing the Formation Energy-Exfoliation Energy-Structure Correlation of MAB Phases Using Machine Learning and DFT, *ACS Appl Mater Interfaces* 12 (2020) 29424–29431.  
[https://doi.org/10.1021/ACSAMI.0C03536/ASSET/IMAGES/LARGE/AM0C03536\\_0003.JPEG](https://doi.org/10.1021/ACSAMI.0C03536/ASSET/IMAGES/LARGE/AM0C03536_0003.JPEG).
- [4.31] Z. Wan, Q. de Wang, Machine Learning Prediction of the Exfoliation Energies of Two-Dimension Materials via Data-Driven Approach, *Journal of Physical Chemistry Letters* 12 (2021) 11470–11475.  
[https://doi.org/10.1021/ACS.JPCLETT.1C03335/ASSET/IMAGES/LARGE/JZ1C03335\\_0006.JPEG](https://doi.org/10.1021/ACS.JPCLETT.1C03335/ASSET/IMAGES/LARGE/JZ1C03335_0006.JPEG).

- [4.32] W.K. Darkwah, Y. Ao, Mini Review on the Structure and Properties (Photocatalysis), and Preparation Techniques of Graphitic Carbon Nitride Nano-Based Particle, and Its Applications, *Nanoscale Research Letters* 2018 13:1 13 (2018) 1–15.  
<https://doi.org/10.1186/S11671-018-2702-3>.
- [4.33] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon N Y* 141 (2019) 580–607.  
<https://doi.org/10.1016/J.CARBON.2018.09.082>.
- [4.34] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [4.35] X. Chen, R. Hu, DFT-based study of single transition metal atom doped g-C<sub>3</sub>N<sub>4</sub> as alternative oxygen reduction reaction catalysts, *Int J Hydrogen Energy* 44 (2019) 15409–15416. <https://doi.org/10.1016/J.IJHYDENE.2019.04.057>.
- [4.36] Z. Bonakchi, A. Nakhaei Pour, S. Soheili, Molecular simulation of methane on various g-C<sub>3</sub>N<sub>4</sub> isomers: collision, adsorption, desorption, and diffusion studies, *Journal of the Iranian Chemical Society* 19 (2022) 3649–3657. <https://doi.org/10.1007/S13738-022-02562-3/TABLES/2>.
- [4.37] W.L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, *Proc Natl Acad Sci U S A* (2005).  
<https://doi.org/10.1073/pnas.0408037102>.
- [4.38] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).
- [4.39] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282.  
<https://doi.org/10.1021/acs.jctc.5b00414>.

- [4.40] J.N. Coleman, M. Lotya, A. O'Neill, S.D. Bergin, P.J. King, U. Khan, K. Young, A. Gaucher, S. De, R.J. Smith, I. V. Shvets, S.K. Arora, G. Stanton, H.Y. Kim, K. Lee, G.T. Kim, G.S. Duesberg, T. Hallam, J.J. Boland, J.J. Wang, J.F. Donegan, J.C. Grunlan, G. Moriarty, A. Shmeliov, R.J. Nicholls, J.M. Perkins, E.M. Grieveson, K. Theuwissen, D.W. McComb, P.D. Nellist, V. Nicolosi, Two-dimensional nanosheets produced by liquid exfoliation of layered materials, *Science* (1979) 331 (2011) 568–571. <https://doi.org/10.1126/science.1194975>.
- [4.41] C. Caleman, P.J. Van Maaren, M. Hong, J.S. Hub, L.T. Costa, D. Van Der Spoel, Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant, *J Chem Theory Comput* 8 (2012) 61–74. <https://doi.org/10.1021/ct200731v>.
- [4.42] V. Consonni, R. Todeschini, Molecular descriptors, *Challenges and Advances in Computational Chemistry and Physics* 8 (2010) 29–102. [https://doi.org/10.1007/978-1-4020-9783-6\\_3](https://doi.org/10.1007/978-1-4020-9783-6_3).
- [4.43] J.P.M. Jcambeck, F. Mocci, A.P. Lyubartsev, A. Laaksonen, Partial atomic charges and their impact on the free energy of solvation, *J Comput Chem* 34 (2013) 187–197. <https://doi.org/10.1002/JCC.23117>.
- [4.44] J. Tukey, *Exploratory data analysis*, 1977. [http://theta.edu.pl/wp-content/uploads/2012/10/exploratorydataanalysis\\_tukey.pdf](http://theta.edu.pl/wp-content/uploads/2012/10/exploratorydataanalysis_tukey.pdf) (accessed September 2, 2022).
- [4.45] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, *Machine Learning* 2006 63:1 63 (2006) 3–42. <https://doi.org/10.1007/S10994-006-6226-1>.
- [4.46] J.H. Friedman, Stochastic gradient boosting, *Comput Stat Data Anal* 38 (2002) 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2).
- [4.47] R.E. Schapire, Explaining adaboost, *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik* (2013) 37–52. [https://doi.org/10.1007/978-3-642-41136-6\\_5/COVER](https://doi.org/10.1007/978-3-642-41136-6_5/COVER).
- [4.48] F.A. da Silva, A.P. Viana, C.C.G. Correa, E.A. Santos, J.A.V.S. de Oliveira, J.D.G. Andrade, R.M. Ribeiro, L.S. Glória, Bayesian ridge regression shows the best fit for SSR

- markers in *Psidium guajava* among Bayesian models, *Scientific Reports* 2021 11:1 11 (2021) 1–11. <https://doi.org/10.1038/s41598-021-93120-z>.
- [4.49] C.A. Ramezan, T.A. Warner, A.E. Maxwell, Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification, *Remote Sensing* 2019, Vol. 11, Page 185 11 (2019) 185. <https://doi.org/10.3390/RS11020185>.
- [4.50] J.M. Sutter, J.H. Kalivas, Comparison of Forward Selection, Backward Elimination, and Generalized Simulated Annealing for Variable Selection, *Microchemical Journal* 47 (1993) 60–66. <https://doi.org/10.1006/MCHJ.1993.1012>.
- [4.51] M.A. Hall, L.A. Smith, Practical feature subset selection for machine learning, *Computer Science* Volume 20 No 1 (1998) 181–191. <https://researchcommons.waikato.ac.nz/handle/10289/1512> (accessed August 13, 2022).
- [4.52] K.P. Shroff, H.H. Maheta, A comparative study of various feature selection techniques in high-dimensional data set to improve classification accuracy, 2015 International Conference on Computer Communication and Informatics, ICCCI 2015 (2015). <https://doi.org/10.1109/ICCCI.2015.7218098>.
- [4.53] F. Pedregosa FABIANPEDREGOSA, V. Michel, O. Grisel OLIVIERGRISEL, M. Blondel, P. Prettenhofer, R. Weiss, J. Vanderplas, D. Cournapeau, F. Pedregosa, G. Varoquaux, A. Gramfort, B. Thirion, O. Grisel, V. Dubourg, A. Passos, M. Brucher, M. Perrot and Édouard and, and Édouard Duchesnay, Fré. Duchesnay EDOUARD DUCHESNAY, Scikit-learn: Machine learning in Python, *Jmlr.Org* 12 (2011) 2825–2830. <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://> (accessed October 12, 2023).
- [4.54] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).

- [4.55] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [4.56] N. Ye, Z. Wang, S. Wang, H. Fang, D. Wang, Aqueous aggregation and stability of graphene nanoplatelets, graphene oxide, and reduced graphene oxide in simulated natural environmental conditions: complex roles of surface and solution chemistry, *Environmental Science and Pollution Research* 25 (2018) 10956–10965. <https://doi.org/10.1007/S11356-018-1326-6/FIGURES/7>.
- [4.57] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).
- [4.58] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [4.59] M. Neumann, O. Steinhauser, G. Stuart Pawley, Consistent calculation of the static and frequency-dependent dielectric constant in computer simulations, *https://Doi.Org/10.1080/00268978400101081* 52 (2006) 97–113. <https://doi.org/10.1080/00268978400101081>.
- [4.60] M.P. Allen, D.J. Tildesley, *Computer simulation of liquids*, 2017. <https://books.google.com/books?hl=en&lr=&id=WFExDwAAQBAJ&oi=fnd&pg=PP1&ots=VGTF41KD4V&sig=M1bd3YNU5iqGsmQa-Bt47iBd6Tg> (accessed August 21, 2022).
- [4.61] S.T. Lin, P.K. Maiti, W.A. Goddard, Two-phase thermodynamic model for efficient and accurate absolute entropy of water from molecular dynamics simulations, *Journal of Physical Chemistry B* 114 (2010) 8191–8198. [https://doi.org/10.1021/JP103120Q/ASSET/IMAGES/LARGE/JP-2010-03120Q\\_0004.JPEG](https://doi.org/10.1021/JP103120Q/ASSET/IMAGES/LARGE/JP-2010-03120Q_0004.JPEG).
- [4.62] S.T. Lin, M. Blanco, W.A. Goddard, The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase

- diagram of Lennard-Jones fluids, *J Chem Phys* 119 (2003) 11792.  
<https://doi.org/10.1063/1.1624057>.
- [4.63] D. van der Spoel, P.J. van Maaren, The origin of layer structure artifacts in simulations of liquid water, *J Chem Theory Comput* 2 (2006) 1–11. <https://doi.org/10.1021/CT0502256>.
- [4.64] C.-Y. Chang, M.-T. Hsu, E.X. Esposito, Y.J. Tseng, Oversampling to Overcome Overfitting: Exploring the Relationship between Data Set Composition, Molecular Descriptors, and Predictive Modeling Methods, (2013). <https://doi.org/10.1021/ci4000536>.
- [4.65] B. Efron, R. Tibshirani, An introduction to the bootstrap, 1994.  
[https://books.google.com/books?hl=en&lr=&id=gLlpIUxRntoC&oi=fnd&pg=PR14&ots=AaAuZ9QdF7&sig=x-o2EoZMRM8w7J\\_-BTLo8GSY3TI](https://books.google.com/books?hl=en&lr=&id=gLlpIUxRntoC&oi=fnd&pg=PR14&ots=AaAuZ9QdF7&sig=x-o2EoZMRM8w7J_-BTLo8GSY3TI) (accessed September 19, 2022).
- [4.66] Y. Zhang, C. Ling, A strategy to apply machine learning to small datasets in materials science, *Npj Computational Materials* 2018 4:1 4 (2018) 1–8.  
<https://doi.org/10.1038/s41524-018-0081-z>.
- [4.67] Y. Liu, T. Zhao, W. Ju, S. Shi, Materials discovery and design using machine learning, *Journal of Materiomics* 3 (2017) 159–177. <https://doi.org/10.1016/J.JMAT.2017.08.002>.
- [4.68] Y. Nademi, T. Tang, H. Uludağ, Modeling Uptake of Polyethylenimine/Short Interfering RNA Nanoparticles in Breast Cancer Cells Using Machine Learning, *Adv Nanobiomed Res* 1 (2021) 2000106. <https://doi.org/10.1002/ANBR.202000106>.
- [4.69] I. Tsamardinos, E. Greasidou, G. Borboudakis, Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation, *Mach Learn* 107 (2018) 1895–1922.  
<https://doi.org/10.1007/S10994-018-5714-4/FIGURES/7>.
- [4.70] R. Singh, C.C. Tripathi, Enhancing Liquid-Phase Exfoliation of Graphene with Addition of Anthracene in Organic Solvents, *Arab J Sci Eng* 42 (2017) 2417–2424.  
<https://doi.org/10.1007/S13369-016-2379-9/METRICS>.
- [4.71] B. Liang, K. Liu, P. Liu, L. Qian, G. Zhao, W. Pan, C. Chen, Organic salt-assisted liquid-phase shear exfoliation of expanded graphite into graphene nanosheets, *Journal of Materiomics* 7 (2021) 1181–1189. <https://doi.org/10.1016/J.JMAT.2021.03.007>.

- [4.72] D. Sahoo, B. Kumar, J. Sinha, S. Ghosh, S.S. Roy, B. Kaviraj, Cost effective liquid phase exfoliation of MoS<sub>2</sub> nanosheets and photocatalytic activity for wastewater treatment enforced by visible light, *Scientific Reports* 2020 10:1 10 (2020) 1–12.  
<https://doi.org/10.1038/s41598-020-67683-2>.
- [4.73] Y.J. Yuan, Z. Shen, S. Wu, Y. Su, L. Pei, Z. Ji, M. Ding, W. Bai, Y. Chen, Z.T. Yu, Z. Zou, Liquid exfoliation of g-C<sub>3</sub>N<sub>4</sub> nanosheets to construct 2D-2D MoS<sub>2</sub>/g-C<sub>3</sub>N<sub>4</sub> photocatalyst for enhanced photocatalytic H<sub>2</sub> production activity, *Appl Catal B* 246 (2019) 120–128. <https://doi.org/10.1016/J.APCATB.2019.01.043>.
- [4.74] W. Du, X. Jiang, L. Zhu, From graphite to graphene: direct liquid-phase exfoliation of graphite to produce single- and few-layered pristine graphene, *J Mater Chem A Mater* 1 (2013) 10592–10606. <https://doi.org/10.1039/C3TA12212C>.
- [4.75] Y. Wang, Z. Shi, J. Yin, Boron nitride nanosheets: large-scale exfoliation in methanesulfonic acid and their composites with polybenzimidazole, *J Mater Chem* 21 (2011) 11371–11377. <https://doi.org/10.1039/C1JM10342C>.
- [4.76] S.M. Lundberg, P.G. Allen, S.-I. Lee, A Unified Approach to Interpreting Model Predictions, (n.d.). <https://doi.org/10.5555/3295222.3295230>.
- [5.1] J. Fu, J. Yu, C. Jiang, B. Cheng, g-C<sub>3</sub>N<sub>4</sub>-Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.
- [5.2] X. Ma, J. Hu, H. He, S. Dong, C. Huang, X. Chen, New Understanding on Enhanced Photocatalytic Activity of g-C<sub>3</sub>N<sub>4</sub>/BiPO<sub>4</sub> Heterojunctions by Effective Interfacial Coupling, *ACS Appl Nano Mater* 1 (2018) 5507–5515.  
<https://doi.org/10.1021/acsanm.8b01012>.
- [5.3] S. Patnaik, D.P. Sahoo, K. Parida, An overview on Ag modified g-C<sub>3</sub>N<sub>4</sub> based nanostructured materials for energy and environmental applications, *Renewable and Sustainable Energy Reviews* 82 (2018) 1297–1312.  
<https://doi.org/10.1016/J.RSER.2017.09.026>.
- [5.4] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are



- We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329.  
<https://doi.org/10.1021/acs.chemrev.6b00075>.
- [5.5] V. Sresht, A.A.H. Pádua, D. Blankschtein, Liquid-Phase Exfoliation of Phosphorene: Design Rules from Molecular Dynamics Simulations, *ACS Nano* 9 (2015) 8255–8268.  
[https://doi.org/10.1021/ACSNANO.5B02683/ASSET/IMAGES/MEDIUM/NN-2015-026836\\_0009.GIF](https://doi.org/10.1021/ACSNANO.5B02683/ASSET/IMAGES/MEDIUM/NN-2015-026836_0009.GIF).
- [5.6] C.J. Shih, S. Lin, M.S. Strano, D. Blankschtein, Understanding the stabilization of liquid-phase-exfoliated graphene in polar solvents: Molecular dynamics simulations and kinetic theory of colloid aggregation, *J Am Chem Soc* 132 (2010) 14638–14648.  
<https://doi.org/10.1021/ja1064284>.
- [5.7] Q. Lin, L. Li, S. Liang, M. Liu, J. Bi, L. Wu, Efficient synthesis of monolayer carbon nitride 2D nanosheet with tunable concentration and enhanced visible-light photocatalytic activities, *Appl Catal B* 163 (2015) 135–142. <https://doi.org/10.1016/j.apcatb.2014.07.053>.
- [5.8] J. Shen, Y. He, J. Wu, C. Gao, K. Keyshar, X. Zhang, Y. Yang, M. Ye, R. Vajtai, J. Lou, P.M. Ajayan, Liquid Phase Exfoliation of Two-Dimensional Materials by Directly Probing and Matching Surface Tension Components, *Nano Lett* 15 (2015) 5449–5454.  
[https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E\\_0005.JPEG](https://doi.org/10.1021/ACS.NANOLETT.5B01842/ASSET/IMAGES/LARGE/NL-2015-01842E_0005.JPEG).
- [5.9] X. Zou, Y. Zhao, M. Li, S. Zhou, C. Chen, Construction of graphitic carbon nitride nanosheets via an improved solvent exfoliation strategy and interfacial mechanics insight from molecular dynamics simulations, *Journal of Porous Materials* 28 (2021) 943–954.  
<https://doi.org/10.1007/s10934-021-01047-7>.
- [5.10] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910.  
<https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [5.11] E. Shahini, N. Chaulagain, K. Shankar, T. Tang, Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning, *ACS Appl Mater*

Interfaces 15 (2023) 53786–53801.

[https://doi.org/10.1021/ACSAMI.3C09347/SUPPL\\_FILE/AM3C09347\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.3C09347/SUPPL_FILE/AM3C09347_SI_001.PDF).

- [5.12] W.K. Darkwah, Y. Ao, Mini Review on the Structure and Properties (Photocatalysis), and Preparation Techniques of Graphitic Carbon Nitride Nano-Based Particle, and Its Applications, *Nanoscale Research Letters* 2018 13:1 13 (2018) 1–15.  
<https://doi.org/10.1186/S11671-018-2702-3>.
- [5.13] M. Inagaki, T. Tsumura, T. Kinumoto, M. Toyoda, Graphitic carbon nitrides (g-C<sub>3</sub>N<sub>4</sub>) with comparative discussion to carbon materials, *Carbon N Y* 141 (2019) 580–607.  
<https://doi.org/10.1016/J.CARBON.2018.09.082>.
- [5.14] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [5.15] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017).  
<https://doi.org/10.1093/nar/gkx312>.
- [5.16] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a.  
<https://doi.org/10.2477/jccjie.2018-0034>.
- [5.17] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).
- [5.18] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282.  
<https://doi.org/10.1021/acs.jctc.5b00414>.

- [5.19] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [5.20] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [5.21] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J Comput Phys* 22 (1976) 245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4).
- [5.22] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [5.23] A. Gupta, V. Arunachalam, S. Vasudevan, Liquid-Phase Exfoliation of MoS<sub>2</sub> Nanosheets: The Critical Role of Trace Water, *Journal of Physical Chemistry Letters* 7 (2016) 4884–4890. [https://doi.org/10.1021/ACS.JPCLETT.6B02405/ASSET/IMAGES/LARGE/JZ-2016-02405F\\_0005.JPEG](https://doi.org/10.1021/ACS.JPCLETT.6B02405/ASSET/IMAGES/LARGE/JZ-2016-02405F_0005.JPEG).
- [5.24] G.A. DiLabio, E.R. Johnson, Lone pair- $\pi$  and  $\pi$ - $\pi$  interactions play an important role in proton-coupled electron transfer reactions, *J Am Chem Soc* 129 (2007) 6199–6203. [https://doi.org/10.1021/JA068090G/SUPPL\\_FILE/JA068090GSI20070312\\_120206.PDF](https://doi.org/10.1021/JA068090G/SUPPL_FILE/JA068090GSI20070312_120206.PDF).
- [5.25] T.J. Mooibroek, P. Gamez, J. Reedijk, Lone pair- $\pi$  interactions: a new supramolecular bond?, *CrystEngComm* 10 (2008) 1501–1515. <https://doi.org/10.1039/B812026A>.
- [5.26] A. Jain, V. Ramanathan, R. Sankararamakrishnan, Lone pair  $\cdots \pi$  interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds, *Protein Sci* 18 (2009) 595. <https://doi.org/10.1002/PRO.67>.
- [5.27] Y.N. Imai, Y. Inoue, I. Nakanishi, K. Kitaura, Cl- $\pi$  interactions in protein-ligand complexes, *Protein Science* 17 (2008) 1129–1137. <https://doi.org/10.1110/PS.033910.107>.

- [5.28] C. Amsaraj, R. Bharathikannan, P. Muthuraja, M. Rajkumar, Chlorine directed C–Cl ...  $\pi$  and C–H...Cl interactions in acridinium 3,5-dichlorosalicylate: Synthesis, X-ray diffraction and theoretical analysis, *J Mol Struct* 1220 (2020) 128759.  
<https://doi.org/10.1016/J.MOLSTRUC.2020.128759>.
- [5.29] K.E. Riley, K.A. Tran, Strength and Character of R–X... $\pi$  Interactions Involving Aromatic Amino Acid Sidechains in Protein-Ligand Complexes Derived from Crystal Structures in the Protein Data Bank, *Crystals* 2017, Vol. 7, Page 273 7 (2017) 273.  
<https://doi.org/10.3390/CRYST7090273>.
- [5.30] A. Jain, C.S. Purohit, S. Verma, R. Sankararamakrishnan, Close contacts between carbonyl oxygen atoms and aromatic centers in protein structures:  $\pi$ ... $\pi$  or lone-pair... $\pi$  interactions?, *Journal of Physical Chemistry B* 111 (2007) 8680–8683.  
[https://doi.org/10.1021/JP072742L/SUPPL\\_FILE/JP072742LSI20070604\\_010130.PDF](https://doi.org/10.1021/JP072742L/SUPPL_FILE/JP072742LSI20070604_010130.PDF).
- [5.31] T. Korenaga, H. Tanaka, T. Ema, T. Sakai, Intermolecular oxygen atom... $\pi$  interaction in the crystal packing of chiral amino alcohol bearing a pentafluorophenyl group, *J Fluor Chem* 122 (2003) 201–205. [https://doi.org/10.1016/S0022-1139\(03\)00089-7](https://doi.org/10.1016/S0022-1139(03)00089-7).
- [5.32] U. Halim, C.R. Zheng, Y. Chen, Z. Lin, S. Jiang, R. Cheng, Y. Huang, X. Duan, A rational design of cosolvent exfoliation of layered materials by directly probing liquid–solid interaction, *Nature Communications* 2013 4:1 4 (2013) 1–7.  
<https://doi.org/10.1038/ncomms3213>.
- [5.33] U. Halim, C.R. Zheng, Y. Chen, Z. Lin, S. Jiang, R. Cheng, Y. Huang, X. Duan, A rational design of cosolvent exfoliation of layered materials by directly probing liquid–solid interaction, *Nature Communications* 2013 4:1 4 (2013) 1–7.  
<https://doi.org/10.1038/ncomms3213>.
- [5.34] K.-G. Zhou, N.-N. Mao, H.-X. Wang, Y. Peng, H.-L. Zhang, K.-G. Zhou, N.-N. Mao, H.-X.H. Wang, -L Zhang, Y. Peng, A Mixed-Solvent Strategy for Efficient Exfoliation of Inorganic Graphene Analogues, *Angewandte Chemie International Edition* 50 (2011) 10839–10842. <https://doi.org/10.1002/ANIE.201105364>.

- [5.35] X. Gu, Y. Zhao, K. Sun, C.L.Z. Vieira, Z. Jia, C. Cui, Z. Wang, A. Walsh, S. Huang, Method of ultrasound-assisted liquid-phase exfoliation to prepare graphene, *Ultrason Sonochem* 58 (2019) 104630. <https://doi.org/10.1016/J.ULTSONCH.2019.104630>.
- [6.1] J. Fu, J. Yu, C. Jiang, B. Cheng, g-C<sub>3</sub>N<sub>4</sub>-Based Heterostructured Photocatalysts, *Adv Energy Mater* (2018). <https://doi.org/10.1002/aenm.201701503>.
- [6.2] Y. Wang, L. Liu, T. Ma, Y. Zhang, H. Huang, Y.H. Wang, L.Z. Liu, Y.H. Zhang, H.W. Huang, T.Y. Ma, 2D Graphitic Carbon Nitride for Energy Conversion and Storage, *Adv Funct Mater* 31 (2021) 2102540. <https://doi.org/10.1002/ADFM.202102540>.
- [6.3] J. Fernández-Catalá, R. Greco, M. Navlani-García, W. Cao, Á. Berenguer-Murcia, D. Cazorla-Amorós, g-C<sub>3</sub>N<sub>4</sub>-Based Direct Z-Scheme Photocatalysts for Environmental Applications, *Catalysts* 2022, Vol. 12, Page 1137 12 (2022) 1137. <https://doi.org/10.3390/CATAL12101137>.
- [6.4] W.J. Ong, L.L. Tan, Y.H. Ng, S.T. Yong, S.P. Chai, Graphitic Carbon Nitride (g-C<sub>3</sub>N<sub>4</sub>)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer to Achieving Sustainability?, *Chem Rev* 116 (2016) 7159–7329. <https://doi.org/10.1021/acs.chemrev.6b00075>.
- [6.5] S. Das, A. Chowdhury, Recent advancements of g-C<sub>3</sub>N<sub>4</sub>-based magnetic photocatalysts towards the degradation of organic pollutants: a review, *Nanotechnology* 33 (2021) 072004. <https://doi.org/10.1088/1361-6528/AC3614>.
- [6.6] M. Majdoub, Z. Anfar, A. Amedlous, Emerging chemical functionalization of g-C<sub>3</sub>N<sub>4</sub>: Covalent/noncovalent modifications and applications, *ACS Nano* 14 (2020) 12390–12469. [https://doi.org/10.1021/ACSNANO.0C06116/ASSET/IMAGES/MEDIUM/NN0C06116\\_0046.GIF](https://doi.org/10.1021/ACSNANO.0C06116/ASSET/IMAGES/MEDIUM/NN0C06116_0046.GIF).
- [6.7] B. Kumru, M. Antonietti, B.V.K.J. Schmidt, Enhanced Dispersibility of Graphitic Carbon Nitride Particles in Aqueous and Organic Media via a One-Pot Grafting Approach, *Langmuir* 33 (2017) 9897–9906. [https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL\\_FILE/LA7B02441\\_SI\\_007.AVI](https://doi.org/10.1021/ACS.LANGMUIR.7B02441/SUPPL_FILE/LA7B02441_SI_007.AVI).

- [6.8] S. Xu, Z. Zhang, D. Wang, J. Lu, Y. Guo, S. Kang, X. Chang, Ultrafast plasma method allows rapid immobilization of monatomic copper on carboxyl-deficient g-C<sub>3</sub>N<sub>4</sub> for efficient photocatalytic hydrogen production, *Front Chem* 10 (2022) 972496. <https://doi.org/10.3389/FCHEM.2022.972496/BIBTEX>.
- [6.9] F. Li, H. Zhou, J. Fan, Q. Xiang, Amine-functionalized graphitic carbon nitride decorated with small-sized Au nanoparticles for photocatalytic CO<sub>2</sub> reduction, *J Colloid Interface Sci* 570 (2020) 11–19. <https://doi.org/10.1016/J.JCIS.2020.02.108>.
- [6.10] M. Yousefi, S. Villar-Rodil, J.I. Paredes, A.Z. Moshfegh, Oxidized graphitic carbon nitride nanosheets as an effective adsorbent for organic dyes and tetracycline for water remediation, *J Alloys Compd* 809 (2019) 151783. <https://doi.org/10.1016/J.JALLCOM.2019.151783>.
- [6.11] J. Feng, D. Zhang, J. Li, S. Bi, Y. Ma, Graphitic carbon nitride nanodots: electronic structure and its influence factors, *J Mater Sci* 55 (2020) 5488–5498. <https://doi.org/10.1007/S10853-020-04396-X/METRICS>.
- [6.12] J. Yang, X. Yang, Y. Li, Molecular simulation perspective of liquid-phase exfoliation, dispersion, and stabilization for graphene, *Curr Opin Colloid Interface Sci* 20 (2015) 339–345. <https://doi.org/10.1016/J.COCIS.2015.10.005>.
- [6.13] S. Haar, M. Bruna, J.X. Lian, F. Tomarchio, Y. Olivier, R. Mazzaro, V. Morandi, J. Moran, A.C. Ferrari, D. Beljonne, A. Ciesielski, P. Samorì, Liquid-Phase Exfoliation of Graphite into Single- and Few-Layer Graphene with  $\alpha$ -Functionalized Alkanes, *Journal of Physical Chemistry Letters* 7 (2016) 2714–2721. [https://doi.org/10.1021/ACS.JPCLETT.6B01260/SUPPL\\_FILE/JZ6B01260\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JPCLETT.6B01260/SUPPL_FILE/JZ6B01260_SI_001.PDF).
- [6.14] E. Shahini, K. Shankar, T. Tang, Liquid-phase exfoliation of graphitic carbon nitrides studied by molecular dynamics simulation, *J Colloid Interface Sci* 630 (2023) 900–910. <https://doi.org/10.1016/J.JCIS.2022.10.150>.
- [6.15] E. Shahini, N. Chaulagain, K. Shankar, T. Tang, Predicting Free Energies of Exfoliation and Solvation for Graphitic Carbon Nitrides Using Machine Learning, *ACS Appl Mater*

- Interfaces 15 (2023) 53786–53801.  
[https://doi.org/10.1021/ACSAMI.3C09347/SUPPL\\_FILE/AM3C09347\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.3C09347/SUPPL_FILE/AM3C09347_SI_001.PDF).
- [6.16] Y. Zheng, J. Liu, J. Liang, M. Jaroniec, S.Z. Qiao, Graphitic carbon nitride materials: controllable synthesis and applications in fuel cells and photocatalysis, *Energy Environ Sci* 5 (2012) 6717–6731. <https://doi.org/10.1039/C2EE03479D>.
- [6.17] Y. Xu, S.P. Gao, Band gap of C<sub>3</sub>N<sub>4</sub> in the GW approximation, *Int J Hydrogen Energy* 37 (2012) 11072–11080. <https://doi.org/10.1016/J.IJHYDENE.2012.04.138>.
- [6.18] H. Li, Z. Zhang, Y. Liu, W. Cen, X. Luo, Functional Group Effects on the HOMO–LUMO Gap of g-C<sub>3</sub>N<sub>4</sub>, *Nanomaterials* 2018, Vol. 8, Page 589 8 (2018) 589. <https://doi.org/10.3390/NANO8080589>.
- [6.19] Z. Chen, S. Li, Y. Peng, C. Hu, Tailoring aromatic ring-terminated edges of g-C<sub>3</sub>N<sub>4</sub> nanosheets for efficient photocatalytic hydrogen evolution with simultaneous antibiotic removal, *Catal Sci Technol* 10 (2020) 5470–5479. <https://doi.org/10.1039/D0CY00898B>.
- [6.20] L. Ma, H. Fan, M. Li, H. Tian, J. Fang, G. Dong, A simple melamine-assisted exfoliation of polymeric graphitic carbon nitrides for highly efficient hydrogen production from water under visible light, *J Mater Chem A Mater* 3 (2015) 22404–22412. <https://doi.org/10.1039/C5TA05850C>.
- [6.21] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017). <https://doi.org/10.1093/nar/gkx312>.
- [6.22] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a. <https://doi.org/10.2477/jccjie.2018-0034>.
- [6.23] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *Journal of Physical Chemistry B* 105 (2001) 6474–6487. [https://doi.org/10.1021/JP003919D/SUPPL\\_FILE/JP003919D\\_S.PDF](https://doi.org/10.1021/JP003919D/SUPPL_FILE/JP003919D_S.PDF).

- [6.24] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282. <https://doi.org/10.1021/acs.jctc.5b00414>.
- [6.25] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23 (1977) 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- [6.26] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J Comput Chem* 13 (1992) 1011–1021. <https://doi.org/10.1002/JCC.540130812>.
- [6.27] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J Comput Chem* (2005). <https://doi.org/10.1002/jcc.20291>.
- [6.28] A. Jain, V. Ramanathan, R. Sankararamakrishnan, Lone pair  $\cdots \pi$  interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds, *Protein Sci* 18 (2009) 595. <https://doi.org/10.1002/PRO.67>.
- [6.29] N. Grinberg, E. Grushka, *Advances in chromatography*. Volume 54, (n.d.).
- [6.30] W. Gao, H. Niu, T. Lin, X. Wang, L. Kong, Molecular dynamics study of response of liquid N,N-dimethylformamide to externally applied electric field using a polarizable force field, *Journal of Chemical Physics* 140 (2014) 44501. <https://doi.org/10.1063/1.4861893/72908>.
- [6.31] C. Zhang, Z. Ren, L. Liu, Z. Yin, Modelling hydrogen bonds in NN-dimethylformamide, *Mol Simul* 39 (2013) 875–881. <https://doi.org/10.1080/08927022.2013.775438>.
- [6.32] M. Kamel, M. Mohammadi, K. Mohammadifard, E.A. Mahmood, M.R. Poor Heravi, A. Heshmati J.M., Z. Hossaini, Comprehensive theoretical prediction of the stability and electronic properties of hydroxyurea and carmustine drugs on pristine and Chitosan-functionalized graphitic carbon nitride in vacuum and aqueous environment, *Vacuum* 207 (2023) 111565. <https://doi.org/10.1016/J.VACUUM.2022.111565>.



- [6.33] K. Asif, M. Perveen, R.A. Khera, S. Nazir, A. Raza Ayub, T. Asif, M. Shabbir, J. Iqbal, Computational and theoretical study of graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>) as a drug delivery carrier for lonidamine drug to treat cancer, *Comput Theor Chem* 1206 (2021) 113459. <https://doi.org/10.1016/J.COMPTC.2021.113459>.
- [6.34] X. Bai, X. Wang, X. Lu, T. Jia, B. Sun, C. Wang, S. Hou, R. Zong, A fluorine induced enhancement of the surface polarization and crystallization of g-C<sub>3</sub>N<sub>4</sub> for an efficient charge separation, *New Journal of Chemistry* 45 (2021) 9334–9345. <https://doi.org/10.1039/D1NJ00668A>.




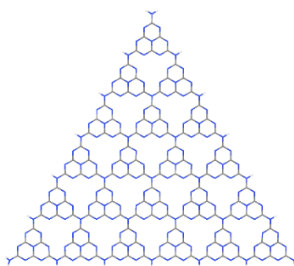
# Appendix A

## Supporting Information for Chapter 3

### A1. Effect of sheet size

A series of simulations were conducted with different sheet size in three solvents (MET, THF and DMF). Table A1 shows the simulated models, which consist of 6, 10, 15 and 21 heptazine units, respectively. Figure A1 a shows the PMF data for THF, where the curves for 15 and 21 units almost overlap with each other. Figure A1b shows that regardless of the type of solvent,  $\Delta G_{\text{exf}}$  reaches a plateau as the number of units increases to 15. This suggests that the effect sheet size is negligible for nanosheets containing over 15 units.

Table A1. Models of g-C<sub>3</sub>N<sub>4</sub> nanosheets simulated to study the effect of sheet size

Number of heptazine units	Surface area (nm <sup>2</sup> )	Structure
6	2.37	
10	4.00	
15	6.12	
21	8.65	

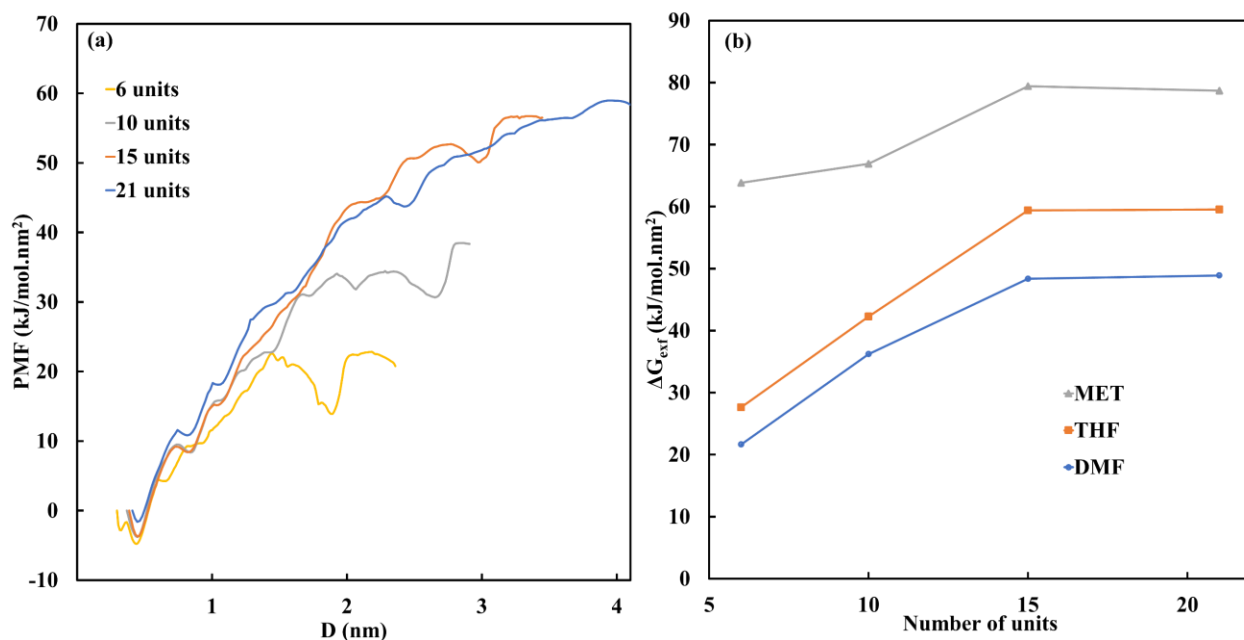


Figure A1. (a) PMF curves in THF for the exfoliation of two g-C<sub>3</sub>N<sub>4</sub> nanosheets with different size. (b)  $\Delta G_{\text{exf}}$  in three different solvents as a function of number of units.

## A2. Validation of force field parameters

Figure A2 and Table A2 compares the bond angle, bond length, surface energy, and hydration free energy of g-C<sub>3</sub>N<sub>4</sub> nanosheet obtained from molecular dynamics (MD) and density functional theory (DFT) calculations. Geometry optimization in DFT was carried out using the Gaussian 16 program, with the B3LYP hybrid exchange–correlation functional and the 6-311++G(d,p) basis set [1,2]. In Table A2, the free energy of hydration evaluated by DFT was obtained using the polarizable continuum model (PCM) and the integral equation formalism [3]. Surface energy and free energy of hydration were also evaluated by MD using different force fields and charge schemes, and are reported separately in Table A2 for comparison. The OPLS-AA force field accompanied by Charge Model 5 (CM5) with a scaling factor of 1.20 produces the best agreement with the DFT results, and is therefore used to model the g-C<sub>3</sub>N<sub>4</sub> nanosheet. Energy minimization using this force field also generated bond and angle distributions that agree well with DFT results (Figure A2).

Table A3 shows the validation of OPLS-AA force field for all the solvents used in this study. MD results on density, dielectric constant, and surface tension are compared with the corresponding experimental values. Good agreement is achieved for the vast majority of the properties, and only the dielectric constants of BD and ACE show relatively poor comparison. Although using other approaches such as a polarizable force fields [4–8] may further improve the comparison, for compatibility (with the g-C<sub>3</sub>N<sub>4</sub> model) and consistency (fair comparison across the solvents) the OPLS-AA force field is used for all components in this work.

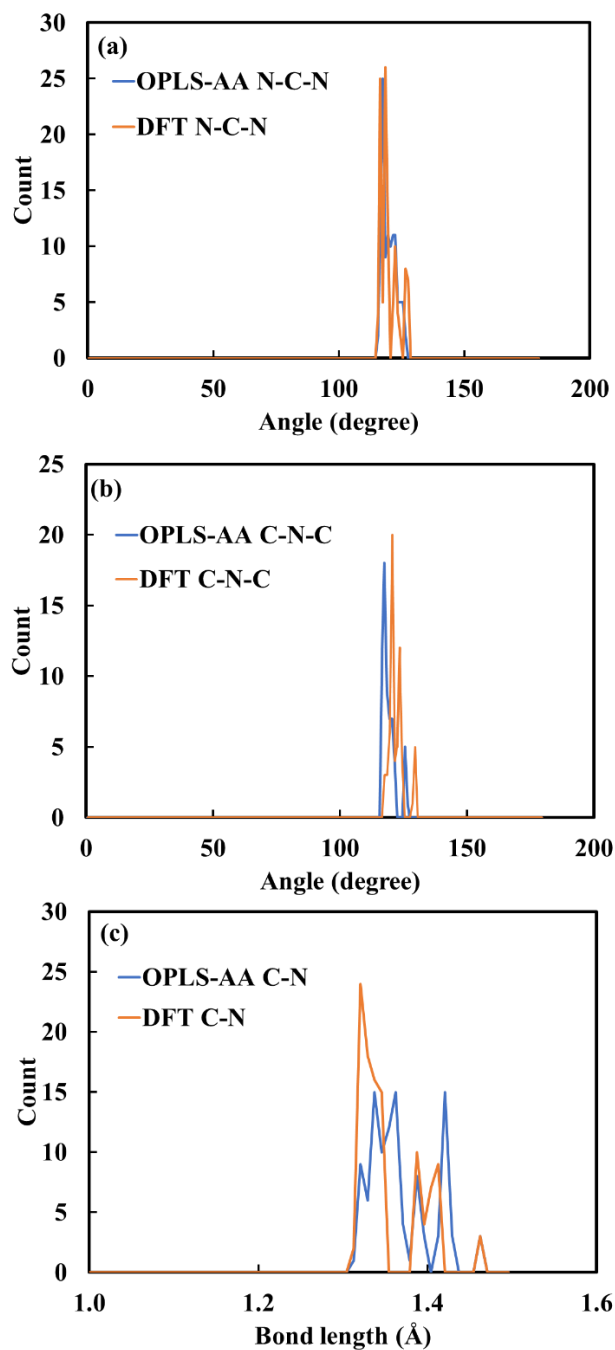


Figure A2. The N-C-N (a) and C-N-C (b) bond angle distributions, and the C-N bond length distribution (c); comparisons made between results from energy minimization using OPLS-AA force field and from geometry optimization using DFT.

Table A2. Surface energy and hydration free energy of a 6-unit g-C<sub>3</sub>N<sub>4</sub> nanosheet evaluated using different methods.

Method/force field	Charge scheme	Surface energy (mJ/m <sup>2</sup> )	Free energy of hydration (kJ/mol)
OPLS-AA	CM5 (ref. [9])	130.2 ± 0.4	-105.2 ± 1.9
OPLS-AA	1.20*CM5	119.3 ± 0.3	-89.6 ± 2.1
OPLS-AA	1.14*CM1A (ref. [10,11])	135.0 ± 0.5	-172.0 ± 3.4
GROMOS	Default (generated by ATB server)	111.0 ± 0.3	-155.4 ± 2.5
GROMOS	CHELPG (generated by Gaussian)	48.2 ± 0.1	-198.4 ± 4.0
DFT	B3LYP/6-311++G(d,p)	115.0 (ref. [12])	-77.1

Table A3. Calculated and experimental (in parenthesis) values of density, dielectric constant, and surface tension at 298.15 K for the nine solvents used in this study.

Solvent	Density (kg/m <sup>3</sup> )	Dielectric constant	Surface tension (mJ/m <sup>2</sup> )	Source of OPLS-AA parameters	Charge scheme
<b>MET</b>	735 ± 0.1 (787 [13])	30.4 ± 0.7 (31.92 [14])	20.0 ± 0.4 (22.1 [14])	Ref. [15]	Ref. [15]
<b>BD</b>	994 ± 0.2 (1013 [14])	15.1 ± 0.9 (32.9 [14])	45.6 ± 0.4 (45.7 [14])	LigParGen	1.14*CM5
<b>ACE</b>	799 ± 0.1 (785 [13])	15.4 ± 0.3 (20.5 [14])	17.2 ± 0.1 (22.7 [14])	Ref. [16]	Ref. [16]
<b>FRM</b>	1126 ± 0.3 (1129 [13])	107.6 ± 0.2 (108.9 [14])	57.2 ± 0.5 (57.0 [14])	Ref. [17]	Ref. [17]
<b>DMF</b>	958 ± 0.1 (943 [13])	37.2 ± 0.6 (37.5 [14])	39.1 ± 0.5 (37.1 [14])	LigParGen	1.20*CM5
<b>CB</b>	1111 ± 0.1 (1110 [14])	4.5 ± 0.0 (5.6 [14])	30.2 ± 0.2 (33.0 [14])	LigParGen	R.E.D server [18]
<b>THF</b>	833 ± 0.2 (880 [13])	5.8 ± 0.0 (6.98 [14])	27.4 ± 0.1 (32.61 [19])	LigParGen	1.14*CM5
<b>NMP</b>	1033 ± 0.2 (1030 [14])	27.1 ± 0.3 (33.0 [14])	38.9 ± 0.2 (41.0 [14])	LigParGen	1.20*CM5
<b>water</b>	1008 ± 0.3 (997 [20])	77.1 ± 0.2 (78.6 [20])	69.5 ± 0.1 (72.0 [20])	TIP4P model [21]	TIP4P model [21]

### A3. Additional data from pulling and US simulations

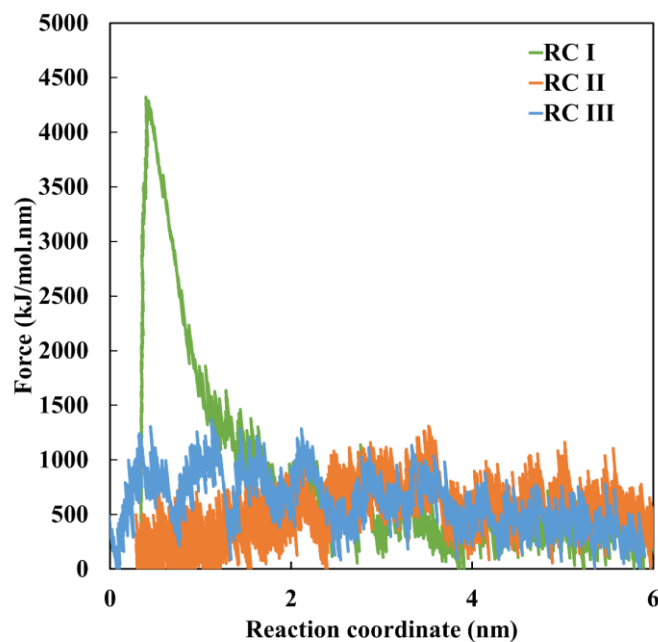


Figure A3. Force-displacement curves for two initially stacked g-C<sub>3</sub>N<sub>4</sub> nanosheets in DMF, pulled apart along the three RCs.

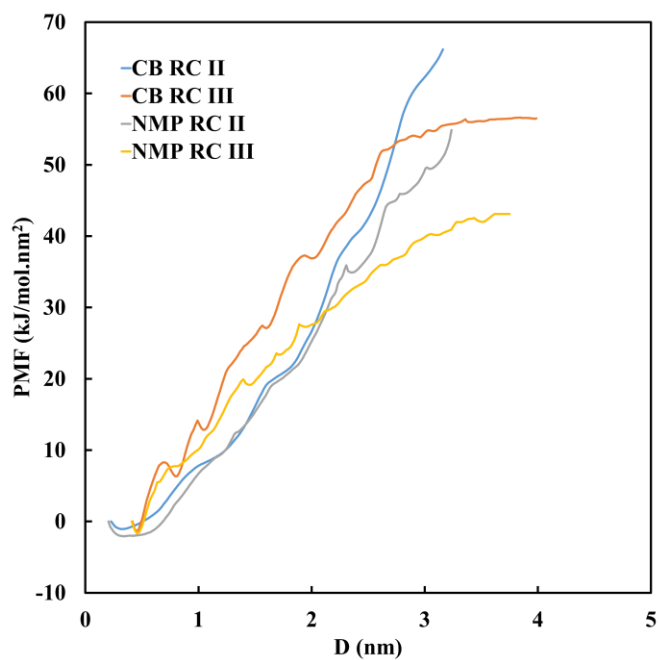
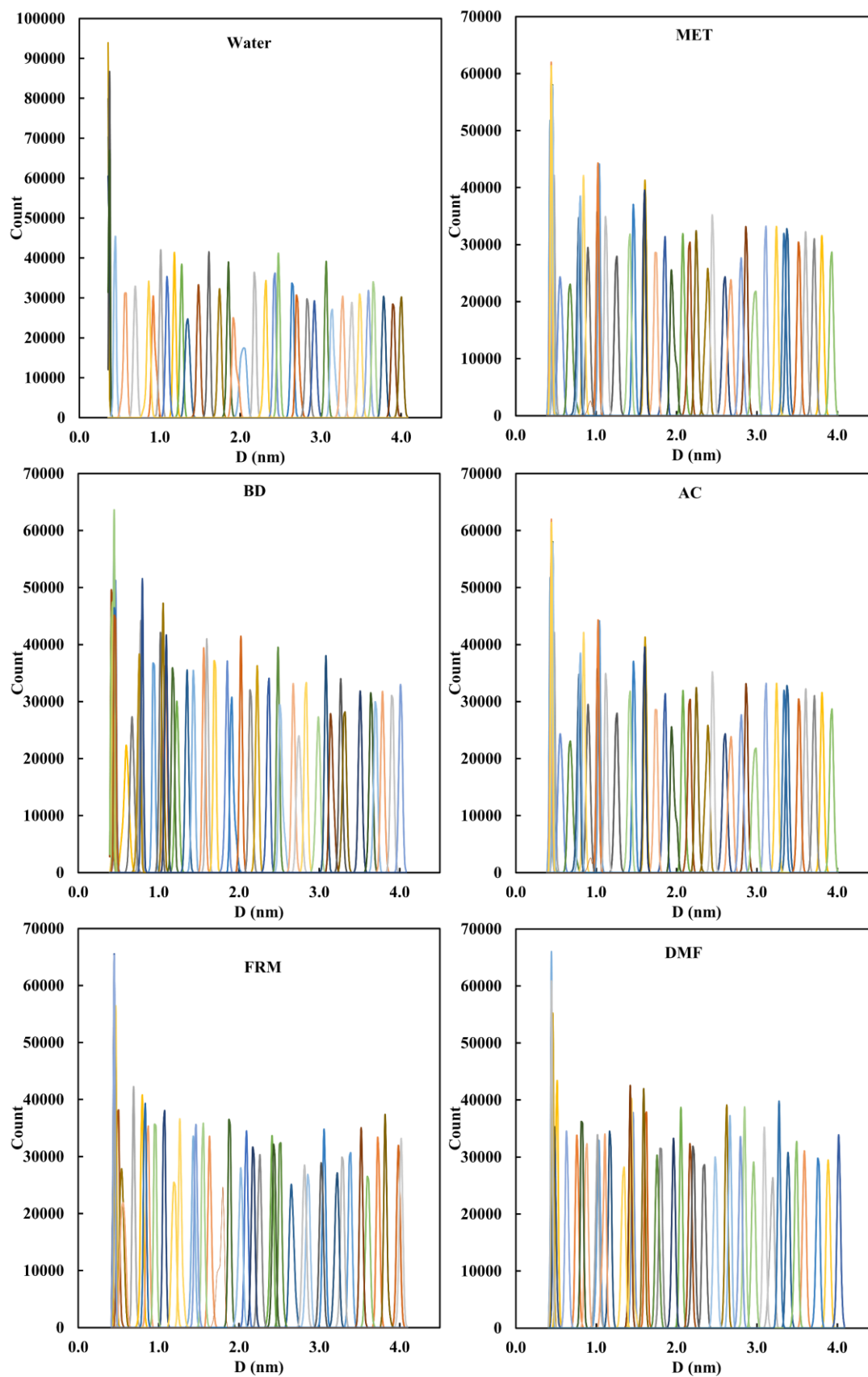


Figure A4. PMF curves for separating two stacked g-C<sub>3</sub>N<sub>4</sub> nanosheets in CB and NMP along RC II and RC III.





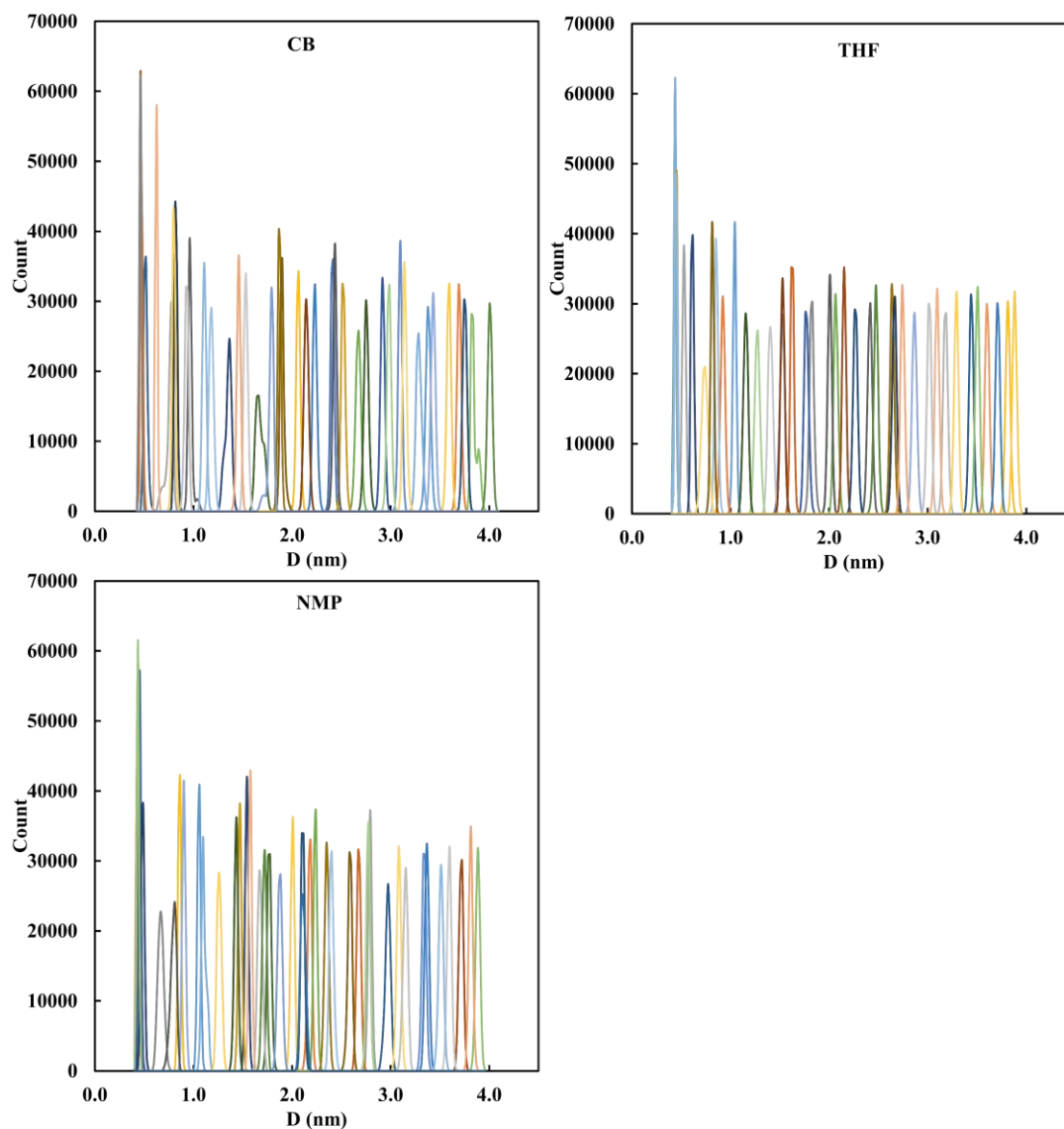


Figure A5. Histograms from US simulations for all solvents.

Table A4.  $\Delta G_{\text{exf}}$  results for DMF from five independent sets of US simulations.

US simulation set #	$\Delta G_{\text{exf}}$ (kJ/mol.nm <sup>2</sup> )	Statistics (kJ/mol.nm <sup>2</sup> )
1	48.98	$48.46 \pm 0.29$
2	48.56	
3	48.79	
4	49.15	
5	49.31	

#### A4. Additional analysis for group 1

From the simulations listed in the last row of Table 1 in the main text, the average numbers of hydrogen bonds (H-bonds),  $N_H$ , were counted for the last 30 ns of the two simulations (solvent alone and solvent with sheet). Without the presence of the sheet, the average  $N_H$  per water molecule is 1.77, the corresponding numbers are 0.97 for MET and 1.70 for BD. This is consistent with the fact that there are two hydroxy groups in each molecule of water and BD, while only one in MET. Upon the insertion of the g-C<sub>3</sub>N<sub>4</sub> nanosheet, H-bonds between solvent molecules are broken, while new H-bonds between the sheet and the solvent are established. The numbers of broken H-bonds, newly established H-bonds and net increase in H-bonds are shown in Figure A6a. The number of broken H-bonds decreases with increasing size of the solvent molecules, and it is almost four times higher for water compared to BD. This is expected since the number of water molecules disrupted by the introduction of the nanosheet is much greater due to its smaller size. The number of newly generated H-bonds between solvent molecules and the sheet is almost the same for MET and BD while water has twice the value. Despite the differences in the numbers of H-bonds broken by the sheet and newly generated H-bonds with the sheet, the net increase in  $N_H$  is similar among the three solvents (Figure A6a).

The types of H-bonds that exist in the simulated systems are shown in Figure A6b. When forming H-bonds with the sheet, MET and BD almost always act as acceptors and the majority of the H-bonds are located on the periphery of the sheet where the 21 H atoms are placed (Figure A7a, middle and right subfigures). On the contrary, water molecules can act both as an acceptor and as a donor (Figure A6b), and H-bonds are formed not only on the periphery but also in the interior of the sheet (Figure A7a, left subfigure). The lack of H-bonds in the interior of the sheet for MET and BD could be related to the size of the solvent molecules. For solvents with larger molecular size, the number of molecules that could form close contact with the surface decreases, which is depicted in Figure A7b by showing the first solvation layer around the sheet and the numbers of solvent molecules in this layer. Here the first solvation layer is defined as solvent molecules whose center of mass (COM) is within 0.5 nm of the sheet, based on the solvent distribution around the sheet (Figure A8). In addition, not all solvent molecules in the first solvation layer form H-bonds with the sheet; the steric hinderance caused by larger molecules reduces the probability of molecular rotation and alignment necessary to form H-bonds in the

interior of the sheet. Although H-bonding controls some of the phenomena associated with carbon nitrides, the contribution of H-bonding to  $\Delta G_{\text{sol}}$  is small in all three solvents Table A5), and H-bonding alone cannot explain the difference in  $\Delta G_{\text{sol}}$ .

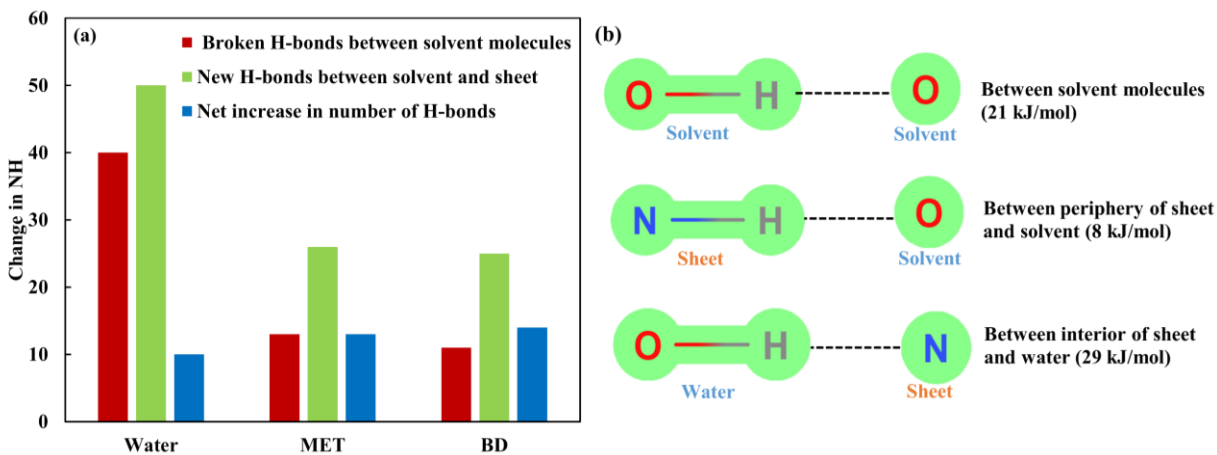


Figure A6. (a) H-bond analysis for solvents from group 1. (b) Different types of H-bonds and their corresponding strength [22].

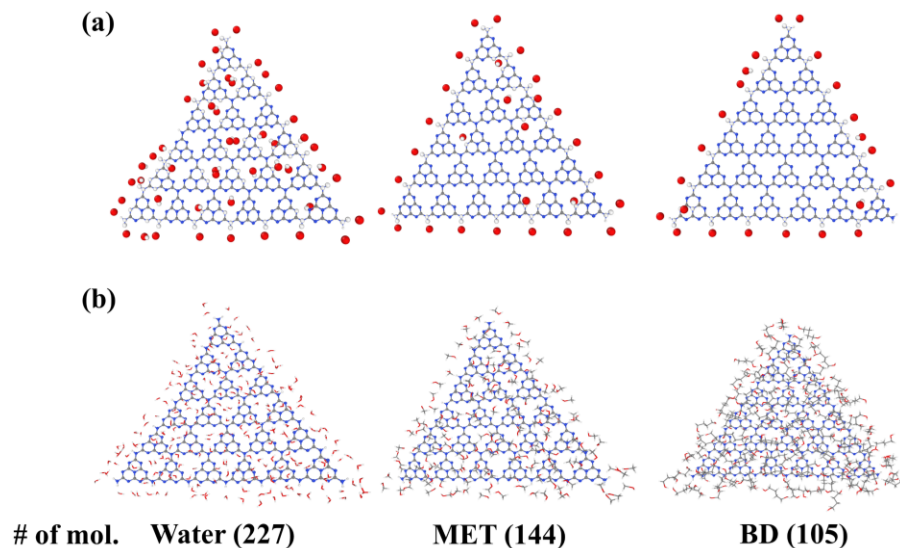


Figure A7. (a) Snapshots of the sheet and solvent atoms involved in H-bonding with the sheet. (b) Snapshots of the sheet and solvent molecules within 0.5 nm of the sheet, which constitute the first solvation layer around the sheet. The average number of solvent molecules in this solvation layer is given below each subfigure. All snapshots are extracted from the simulations for “ $\Delta H_{\text{sol}}$  and  $\Delta S_{\text{sol}}$  calculation” (refer to Table 1 in main text). Color scheme: blue, N; gray, C; white, H; red, O.

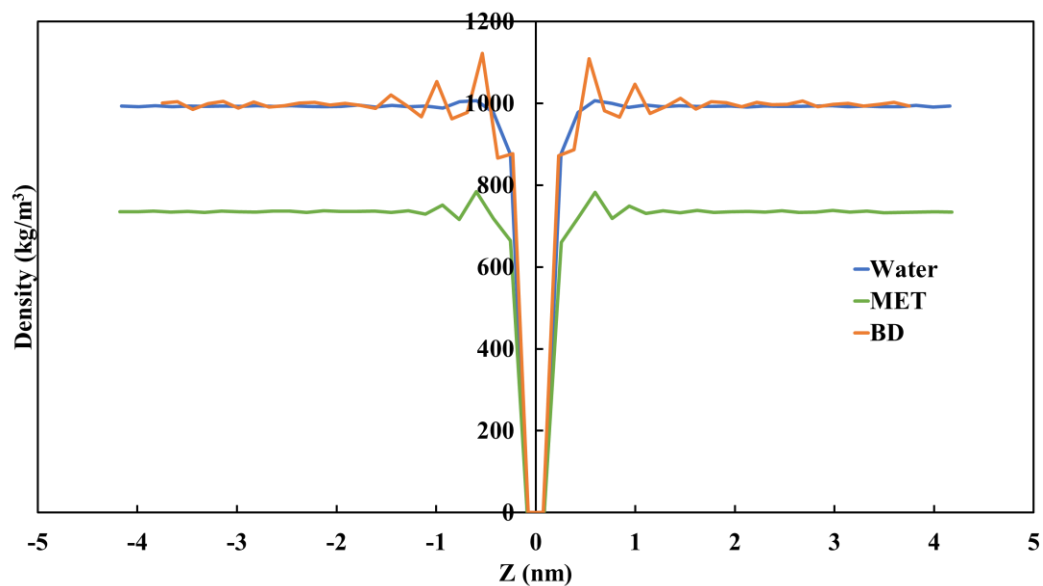


Figure A8. The density profile of solvents as a function of  $Z$ , the coordinate perpendicular to the nanosheet with  $Z = 0$  corresponding to the location of the nanosheet. For all three solvents, the first peak is located at  $\sim 0.5$  nm from the sheet. The solvent molecules within this distance are considered as the first solvation layer.

Table A5. Contribution of H-bonding to  $\Delta G_{\text{sol}}$  for solvents from group 1.

Solvent	Net increase in # of H-bonds (average)	Energy difference due to H-bonds (kJ/mol)	$\Delta G_{\text{sol}}$ (kJ/mol)	% contribution of H-bonding to $\Delta G_{\text{sol}}$
water	10	-131	-1047.0	12%
MET	13	-40	-690.2	5%
BD	14	-53	-551.6	9%

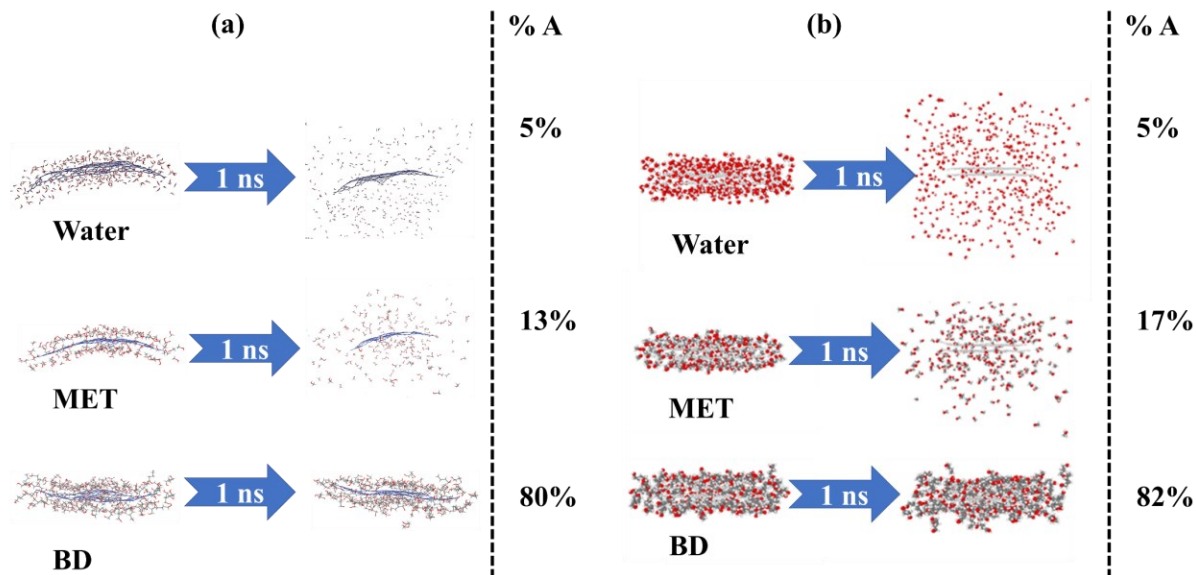


Figure A9. (a) The snapshots of the first solvation layer around the sheet at 30 ns and the same molecules at 31 ns in the simulation with one nanosheet. (b) The snapshots of the first solvation layer around the sheets at the beginning and after 1 ns of the US simulation where the COM distance between the two sheets is restrained to around 0.34 nm. %A: % of initially adhering solvent molecules that remained in the first solvation layer after 1 ns.

## A5. Energy analysis for the solvation process

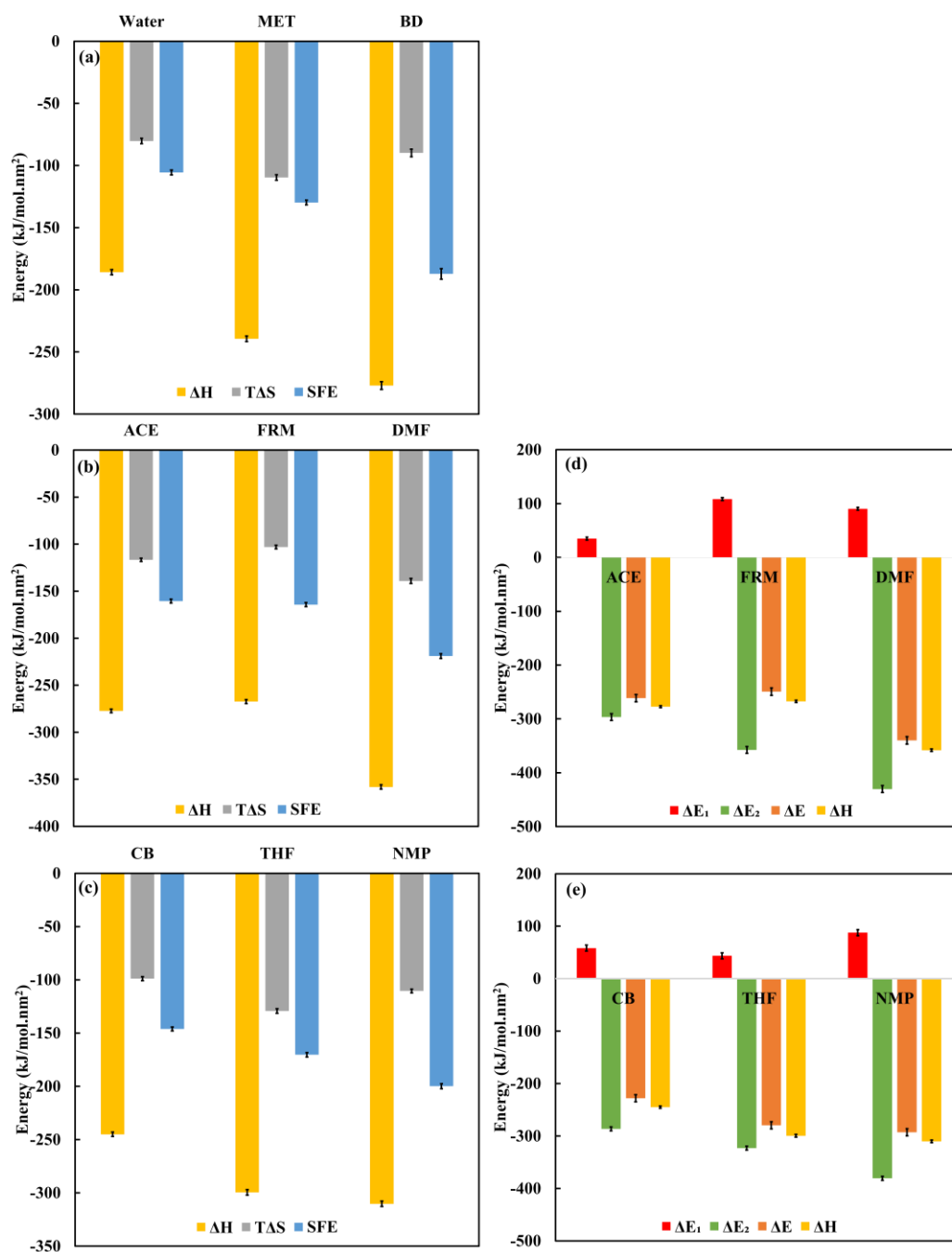


Figure A10. Enthalpic and entropic contributions to  $\Delta G_{\text{sol}}$  for (a) group 1, (b) group 2, and (c) group 3. Energy change in the solvation process associated with cavity generation ( $\Delta E_1$ ) and sheet insertion ( $\Delta E_2$ ) for (d) group 2 and (e) group 3. The sum of the two terms ( $\Delta E = \Delta E_1 + \Delta E_2$ ) along with  $\Delta H_{\text{sol}}$  are also shown. Corresponding data for group 1 is given in the main text. Statistics are based on the last 30 ns of the 60 ns simulations.

## A6. RDF plots

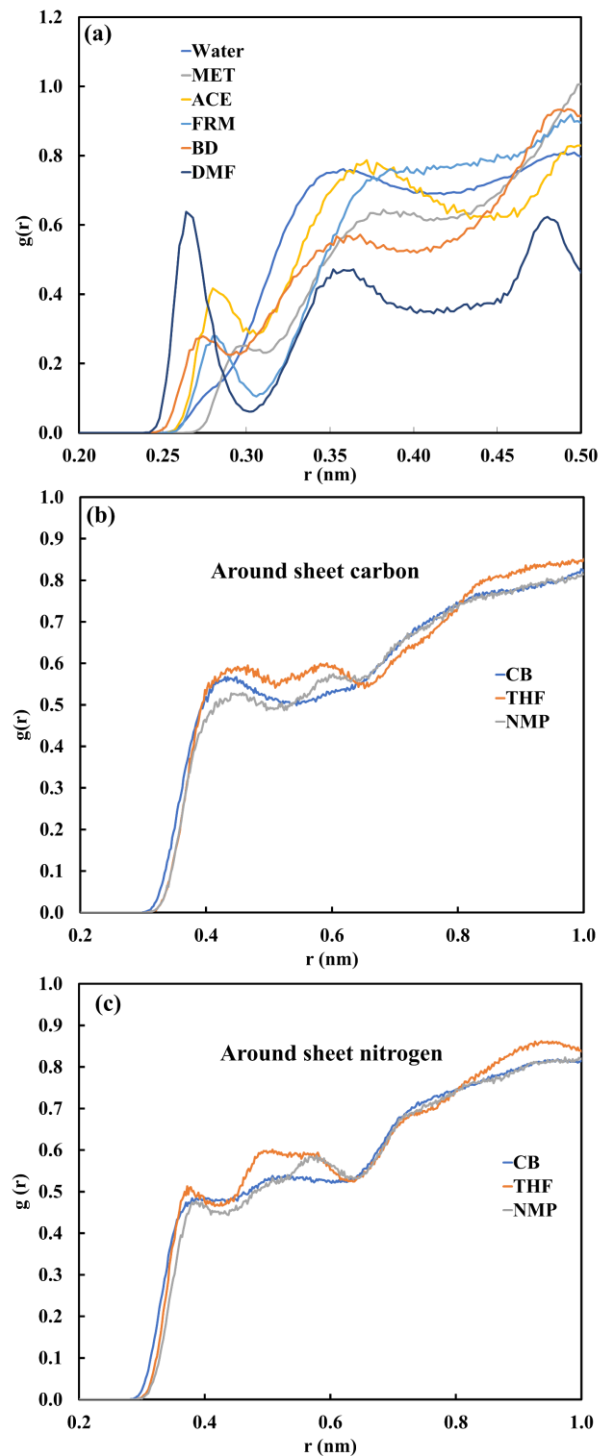


Figure A11. (a) RDF of solvent oxygen around nitrogen atoms of the  $g\text{-C}_3\text{N}_4$  sheet for group 1 and 2 solvents. RDF of the aromatic center of group 3 solvent molecules around carbon (b) and nitrogen (c) atoms of the sheet.



## A7. Fraction of molecules that departed from the first solvation layer

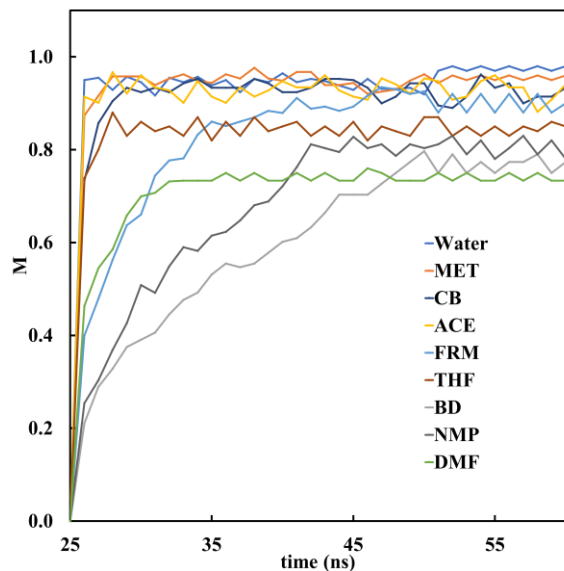


Figure A12. The fraction of molecules that departed from the first solvation layer ( $M$ ) versus time of the simulation.

Table A6. Statistics for  $M$  in the last 10 ns of Figure A12.

Solvent	Statistics of $M$
Water	$0.970 \pm 0.018$
MET	$0.955 \pm 0.006$
BD	$0.770 \pm 0.018$
ACE	$0.930 \pm 0.024$
FRM	$0.901 \pm 0.021$
DMF	$0.739 \pm 0.008$
CB	$0.922 \pm 0.023$
THF	$0.848 \pm 0.015$
NMP	$0.805 \pm 0.019$

## References

- [1] M.F. gaussian. com/, undefined 2004, Gaussian 03 Rev. E. 01, Ci.Nii.Ac.Jp (n.d.).  
<https://ci.nii.ac.jp/naid/10029868106/> (accessed September 28, 2022).
- [2] H.B. Schlegel, Optimization of equilibrium geometries and transition structures, *J Comput Chem* 3 (1982) 214–218. <https://doi.org/10.1002/JCC.540030212>.
- [3] V. Barone, M. Cossi, J. Tomasi, Geometry Optimization of Molecular Structures in Solution by the Polarizable Continuum Model, *J Comput Chem* 19 (1998).  
[https://doi.org/10.1002/\(SICI\)1096-987X\(199803\)19:4](https://doi.org/10.1002/(SICI)1096-987X(199803)19:4).
- [4] I. González-Veloso, N.M. Figueiredo, M.D. Natália S Cordeiro, J.M. C Marques, F. Pirani, F. Vasconcellos Prudente, Unravelling the Interactions of Magnetic Ionic Liquids by Energy Decomposition Schemes: Towards a Transferable Polarizable Force Field, *Molecules* 2021, Vol. 26, Page 5526 26 (2021) 5526. <https://doi.org/10.3390/MOLECULES26185526>.
- [5] M. Jorge, L. Lue, The dielectric constant: Reconciling simulation and experiment, *J Chem Phys* 150 (2019) 084108. <https://doi.org/10.1063/1.5080927>.
- [6] M. Jorge, J.R.B. Gomes, A.W. Milne, Self-consistent electrostatic embedding for liquid phase polarization, *J Mol Liq* 322 (2021) 114550. <https://doi.org/10.1016/J.MOLLIQ.2020.114550>.
- [7] D. Bedrov, J.P. Piquemal, O. Borodin, A.D. MacKerell, B. Roux, C. Schröder, Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields, *Chem Rev* 119 (2019) 7940–7995.  
[https://doi.org/10.1021/ACS.CHEMREV.8B00763/ASSET/IMAGES/MEDIUM/CR-2018-007638\\_M063.GIF](https://doi.org/10.1021/ACS.CHEMREV.8B00763/ASSET/IMAGES/MEDIUM/CR-2018-007638_M063.GIF).
- [8] A.A.H. Pádua, A.A.H. P ' Adua, Resolving dispersion and induction components for polarisable molecular simulations of ionic liquids, *J Chem Phys* 146 (2017) 204501.  
<https://doi.org/10.1063/1.4983687>.
- [9] L.S. Dodda, J.Z. Vilseck, K.J. Cutrona, W.L. Jorgensen, Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling, *J Chem Theory Comput* 11 (2015) 4273–4282.  
<https://doi.org/10.1021/acs.jctc.5b00414>.

- [10] L.S. Dodda, I.C. De Vaca, J. Tirado-Rives, W.L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res* (2017). <https://doi.org/10.1093/nar/gkx312>.
- [11] M. YABE, K. MORI, K. UEDA, M. TAKEDA, Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers, *Journal of Computer Chemistry, Japan -International Edition* 5 (2019) n/a. <https://doi.org/10.2477/jccjie.2018-0034>.
- [12] X. Zhang, X. Xie, H. Wang, J. Zhang, B. Pan, Y. Xie, Enhanced photoresponsive ultrathin graphitic-phase C<sub>3</sub>N<sub>4</sub> nanosheets for bioimaging, *J Am Chem Soc* 135 (2013) 18–21. <https://doi.org/10.1021/ja308249k>.
- [13] Y. Marcus, *Wiley Series in Solution Chemistry, Volume 4 - The Properties of Solvents*, (1999). <https://www.wiley.com/en-ca/The+Properties+of+Solvents-p-9780471983699> (accessed September 28, 2022).
- [14] D. Lide, *CRC handbook of chemistry and physics*, 2004. [https://books.google.com/books?hl=en&lr=&id=WDl18hA006AC&oi=fnd&pg=PA1&dq=Lide,+D.+R.+CRC+Handbook+of+Chemistry+and+Physics+90th+edition%3B+CRC+Press:+Cleveland,+Ohio,+2009&ots=U0nF-POZQq&sig=P2FHhRmUdRaKTQ\\_xMdmGIY1\\_gY](https://books.google.com/books?hl=en&lr=&id=WDl18hA006AC&oi=fnd&pg=PA1&dq=Lide,+D.+R.+CRC+Handbook+of+Chemistry+and+Physics+90th+edition%3B+CRC+Press:+Cleveland,+Ohio,+2009&ots=U0nF-POZQq&sig=P2FHhRmUdRaKTQ_xMdmGIY1_gY) (accessed September 28, 2022).
- [15] R. Zangi, Refinement of the OPLSAA Force-Field for Liquid Alcohols, *ACS Omega* 3 (2018) 18089–18099. [https://doi.org/10.1021/ACSOMEGA.8B03132/SUPPL\\_FILE/AO8B03132\\_SI\\_001.PDF](https://doi.org/10.1021/ACSOMEGA.8B03132/SUPPL_FILE/AO8B03132_SI_001.PDF).
- [16] C. Caleman, P.J. Van Maaren, M. Hong, J.S. Hub, L.T. Costa, D. Van Der Spoel, Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant, *J Chem Theory Comput* 8 (2012) 61–74. <https://doi.org/10.1021/ct200731v>.
- [17] A.P. de La Luz, G.A. Méndez-Maldonado, E. Núñez-Rojas, F. Bresme, J. Alexandre, A New Force Field of Formamide and the Effect of the Dielectric Constant on Miscibility, *J Chem*

Theory Comput 11 (2015) 2792–2800.

[https://doi.org/10.1021/ACS.JCTC.5B00080/SUPPL\\_FILE/CT5B00080\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.JCTC.5B00080/SUPPL_FILE/CT5B00080_SI_001.PDF).

- [18] S. F-Y Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, O.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building, *Physical Chemistry Chemical Physics* 12 (2010) 7821–7839. <https://doi.org/10.1039/C0CP00111B>.
- [19] M.D. Lechner, C. Wohlfarth, *Surface Tension of Pure Liquids and Binary Liquid Mixtures: (Supplement to IV/16) (Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology - New Series)*, Springer, 2008.  
<https://link.springer.com/book/9783540755074> (accessed September 28, 2022).
- [20] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E.E. Bolton, PubChem in 2021: New data content and improved web interfaces, *Nucleic Acids Res* 49 (2021) D1388–D1395. <https://doi.org/10.1093/nar/gkaa971>.
- [21] H.W. Horn, W.C. Swope, J.W. Pitera, J.D. Madura, T.J. Dick, G.L. Hura, T. Head-Gordon, Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew, *J Chem Phys* 120 (2004) 9665. <https://doi.org/10.1063/1.1683075>.
- [22] W.D. Cornell, P. Cieplak, C.I. Bayly, P.A. Kollman, Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation, *J Am Chem Soc* 115 (1993) 9620–9631. [https://doi.org/10.1021/JA00074A030/SUPPL\\_FILE/JA9620.PDF](https://doi.org/10.1021/JA00074A030/SUPPL_FILE/JA9620.PDF).

# Appendix B

## Supporting Information for Chapter 4

### B1. Dataset

Table B1. Numbering, name, Simplified Molecular Input Line Entry System (SMILES) codes, and molecular structure of solvents used to build the dataset.

Name	SMILES	Structure	Name	SMILES	Structure
1. Chloroform	<chem>C1C(Cl)Cl</chem>		26. Pyrimidine	<chem>C1=CN=CN=C1</chem>	
2. Dichlorofluoromethane	<chem>C(Cl)Cl</chem>		28. Dimethylacetamide	<chem>CN(C)C(C)=O</chem>	
3. Dibromomethane	<chem>BrCBr</chem>		27. 1,4-dichlorobutane	<chem>C(CCCl)CCl</chem>	
4. Methanoic acid	<chem>O=CO</chem>		29. 1,2,3,5-Tetrafluorobenzene	<chem>C1=C(C=C(C(=C1)F)F)F</chem>	
5. Formamide	<chem>O=CN</chem>		30. 1,3-Difluorobenzene	<chem>C1=CC(=CC(=C1)F)F</chem>	
6. Methanol	<chem>CO</chem>		31. Benzenethiol	<chem>C1=CC=C(C=C1)S</chem>	
7. Pentachloroethane	<chem>C(C(Cl)(Cl)Cl)(Cl)Cl</chem>		32. 4-Methylpyridine	<chem>CC1=CC=NC=C1</chem>	
8. Acetamide	<chem>CC(=O)N</chem>		33. Cyclohexanone	<chem>C1CCC(=O)CC1</chem>	
9. N-Methylformamide	<chem>CNC=O</chem>		34. Triethyl phosphate	<chem>CCOP(=O)(OCC)OCC</chem>	
10. Ethanol	<chem>OCC</chem>		35. N,N-diethylethanamine	<chem>CCN(CC)CC</chem>	
11. Dimethyl sulfoxide	<chem>CS(=O)C</chem>		36. Benzonitrile	<chem>N#Cc1ccccc1</chem>	
12. Ethanolamine	<chem>NCCO</chem>		37. Toluene	<chem>Cc1ccccc1</chem>	

13. Ethane-1,2-diamine	<chem>C(CN)N</chem>		38. 3-Methylphenol	<chem>CC1=CC(=CC=C1)O</chem>	
14. gamma-Butyrolactone	<chem>O=C1OCC1</chem>		39. 1,2-Dimethoxybenzene	<chem>COC1=CC=CC(OC)=C1</chem>	
15. Propanenitrile	<chem>CCC#N</chem>		40. 1-Butoxybutane	<chem>CCCCOCCCC</chem>	
16. 1,2-Dibromopropane	<chem>CC(CBr)Br</chem>		41. Quinoline	<chem>C1=CC=C2C(=C1)C=CC=N2</chem>	
17. Methyloxirane	<chem>CC1CO1</chem>		42. 2,6-Dimethyl-4-heptanone	<chem>CC(C)CC(=O)CC(C)C</chem>	
18. Acetone	<chem>CC(=O)C</chem>		43. Phenoxybenzene	<chem>O(c1ccccc1)c2ccccc2</chem>	
19. Tetrahydrofuran	<chem>C1CCOC1</chem>		44. Chlorobenzene	<chem>c1ccccc1Cl</chem>	
20. Dimethylformamide	<chem>CN(C)C=O</chem>		45. 1-Cyclohexylpyrrolidin-2-one	<chem>O=C2N(C1CCCCC1)CC2</chem>	
21. N-Methylacetamide	<chem>CC(=O)NC</chem>		46. Dichlorobenzene	<chem>ClC1=CC=C(Cl)C=C1</chem>	
22. Glycerol	<chem>OCC(O)CO</chem>		47. Propan-2-ol	<chem>CC(O)C</chem>	
23. Isobutane	<chem>CC(C)C</chem>		48. N-Vinylpyrrolidone	<chem>O=C1N(C=C)CCC1</chem>	
24. Butane-1,4-diol	<chem>OCCCCO</chem>		49. Water	<chem>O</chem>	<chem>H2O</chem>
25. 2-Methylpropan-2-amine	<chem>CC(C)(C)N</chem>				

Table B2. Dataset: descriptors and target properties for developing ML models.

#	$\Delta H_{vap}$ (kJ/mol)	$C_p$ (J/mol·K)	$\varepsilon$	$\gamma$ (mN/m)	$\kappa_T$ (1/GPa)	$\alpha_p$ ( $10^{-3}/K$ )	$\rho$ (g/l)	$M_w$ (g/mol)	DHBC	AHBC	HAC	Arm	$C_{mean}$	$C_{max}$	$\Delta G_{sol}$ (kJ/mol·nm <sup>2</sup> )	$\Delta G_{ext}$ (kJ/mol·nm <sup>2</sup> )
1	29.21	94.16	5.67	26.04	2.15	1.99	1374.1	119.37	0	0	4	0	0.13	0.33	147.48	59.41
2	23.78	70.12	5.9	17.40	2.23	2.09	1203.4	84.93	0	1	4	0	0.08	0.10	134.79	60.09
3	34.17	104.2	13.97	25.10	0.91	1.30	2475.3	173.84	0	0	3	0	0.09	0.11	152.17	49.56
4	42.48	98.65	42.94	37.31	0.54	1.05	1136.8	46.03	0	0	3	0	0.39	0.53	164.07	55.33
5	25.83	87.11	59.19	57.20	0.36	1.25	1126.0	45.04	1	1	3	0	0.31	0.53	157.00	74.69
6	24.33	75.8	35.64	12.03	2.02	1.47	735.1	32.04	1	1	2	0	0.05	0.05	129.80	78.68
7	47.53	154.8	4.97	34.70	0.77	1.14	1681.7	202.09	0	0	7	0	0.10	0.21	148.82	53.49
8	64.67	137.2	44.65	38.16	0.46	0.79	1018.7	59.07	1	1	4	0	0.32	0.76	180.23	56.76
9	49.85	119.4	35.08	36.91	0.46	0.99	979.9	59.07	1	1	4	0	0.22	0.50	184.03	50.79
10	31.57	109.53	16.60	18.77	1.32	1.33	792.0	46.07	1	1	3	0	0.19	0.68	151.20	67.96
11	55.62	132.17	40.59	42.43	0.58	1.03	1082.0	78.13	0	2	4	0	0.17	0.46	180.70	55.19
12	31.22	106.81	49.88	48.22	0.47	0.96	1038.1	61.08	2	2	4	0	0.29	0.90	174.59	71.98
13	31.44	142.34	19.38	27.67	0.52	1.24	981.5	60.10	2	2	4	0	0.30	0.90	169.93	69.79
14	64.97	142.67	32.78	36.14	0.54	1.12	1104.0	86.09	0	2	6	1	0.20	0.44	172.30	47.55
15	33.57	89.33	20.56	16.26	1.50	1.77	753.6	55.08	0	1	4	0	0.15	0.56	158.86	70.58
16	50.99	147.09	36.08	33.90	0.52	1.04	2096.8	201.89	0	0	5	0	0.11	0.22	164.09	51.66
17	26.93	111.61	51.94	13.27	1.70	1.93	774.70	58.08	0	1	4	0	0.12	0.40	157.35	72.04
18	30.76	117.37	17.69	17.21	1.45	1.57	799.0	58.08	0	1	4	0	0.17	0.47	153.20	69.44

Table B2. Dataset - continued

#	$\Delta H_{vap}$ (kJ/mol)	$C_p$ (J/mol·K)	$\epsilon$	$\gamma$ (mN/m)	$\kappa_T$ (1/GPa)	$\alpha_p$ ( $10^{-3}/K$ )	$\rho$ (g/l)	$M_w$ (g/mol)	DHBC	AHBC	HAC	Arm	$C_{mean}$	$C_{max}$	$\Delta G_{sol}$ (kJ/mol·nm <sup>2</sup> )	$\Delta G_{ext}$ (kJ/mol·nm <sup>2</sup> )
19	32.14	124.23	6.07	18.07	1.03	1.83	833.0	72.11	0	1	5	1	0.13	0.38	163.20	59.41
20	46.08	148.95	41.38	31.50	0.46	1.21	922.2	73.10	0	1	5	0	0.14	0.50	181.83	54.85
21	55.12	150.16	15.01	33.16	0.50	0.90	952.0	73.09	1	1	5	0	0.20	0.50	186.65	53.28
22	89.37	174.12	39.12	51.83	0.22	0.59	1259.7	92.09	3	3	6	0	0.44	0.73	204.60	61.68
23	24.61	80.93	1.01	12.66	1.66	2.12	584.9	58.12	0	0	4	0	0.04	0.04	114.57	77.36
24	39.50	150.01	15.39	45.65	0.34	0.85	994.0	90.12	2	2	6	0	0.20	0.59	187.20	62.83
25	36.68	181.21	4.01	19.37	1.27	1.92	763.4	73.14	1	1	5	0	0.18	0.90	171.88	55.45
26	49.13	142.55	8.01	41.51	0.44	0.67	1091.8	80.09	0	2	6	1	0.48	0.87	201.33	49.34
27	46.92	180.91	46.37	25.49	0.80	1.10	1128.1	127.01	0	0	6	0	0.09	0.20	149.59	60.60
28	51.82	182.32	26.37	32.18	0.55	1.02	924.1	87.12	0	1	6	0	0.15	0.37	167.40	45.95
29	36.87	173.13	8.97	41.22	1.51	1.70	1342.8	150.07	0	4	10	1	0.16	0.22	148.89	52.01
30	53.24	145.22	5.94	14.11	1.44	1.75	1107.1	114.09	0	2	8	1	0.15	0.22	148.89	54.89
31	41.31	153.11	2.42	21.20	0.82	1.33	1052.0	110.18	1	1	7	1	0.12	0.22	147.08	65.01
32	46.44	160.66	6.98	27.97	0.60	0.95	948.4	93.13	0	1	7	1	0.24	0.68	172.04	55.99
33	47.61	179.02	46.89	27.33	0.65	0.99	947.9	98.15	0	1	7	1	0.13	0.47	174.96	63.42
34	77.60	331.29	30.81	28.81	0.58	0.64	1081.1	182.15	0	4	11	0	0.26	2.00	208.28	48.26
35	59.62	228.87	28.29	22.97	1.34	1.89	754.1	101.19	0	1	7	0	0.11	0.63	166.82	47.64
36	53.19	151.21	6.81	40.74	0.48	0.88	1012.7	103.12	0	1	8	1	0.14	0.43	172.62	56.98



Table B2. Dataset - continued

#	$\Delta H_{vap}$ (kJ/mol)	$C_p$ (J/mol·K)	$\varepsilon$	$\gamma$ (mN/m)	$\kappa_T$ (1/GPa)	$\alpha_p$ ( $10^{-3}/K$ )	$\rho$ (g/l)	$M_w$ (g/mol)	DHBC	AHBC	HAC	Arm	$C_{mean}$	$C_{max}$	$\Delta G_{sol}$ (kJ/mol·nm <sup>2</sup> )	$\Delta G_{ext}$ (kJ/mol·nm <sup>2</sup> )
37	40.02	157.18	4.50	20.96	0.88	1.49	872.0	92.14	0	0	7	1	0.10	0.12	148.27	56.91
38	66.28	246.84	10.18	31.44	0.53	0.98	1022.2	108.14	1	1	8	1	0.15	0.59	166.70	50.34
39	63.53	238.11	6.38	30.40	0.53	1.00	1057.0	138.16	0	2	10	1	0.10	0.29	164.37	52.90
40	50.21	289.39	11.97	37.61	1.13	1.52	767.8	130.23	0	1	9	0	0.09	0.40	151.85	48.90
41	61.05	194.05	24.93	33.73	0.48	0.90	1077.7	129.16	0	1	10	1	0.21	0.69	181.91	44.37
42	59.43	316.96	8.43	20.21	0.77	1.33	836.1	142.24	0	1	10	0	0.11	0.47	167.57	53.84
43	72.70	275.17	11.60	31.49	0.44	1.16	1084.2	170.21	0	1	13	1	0.12	0.27	170.26	48.91
44	40.16	150.59	4.45	32.99	0.80	1.32	1111.0	112.55	0	0	7	1	0.11	0.17	138.80	58.04
45	39.34	141.71	4.40	42.30	0.45	0.70	1037.0	167.25	0	1	12	1	0.19	0.93	195.31	62.56
46	49.03	163.78	8.90	36.61	0.85	1.33	1285.0	147.00	0	0	8	1	0.09	0.13	139.00	49.26
47	45.39	163.89	14.84	22.07	1.08	1.10	800.0	60.10	1	1	4	0	0.17	0.68	156.60	59.52
48	55.32	149.07	33.97	40.72	0.40	0.77	1044.7	111.14	0	1	10	1	0.24	0.79	189.87	49.72
49	40.80	70.52	52.95	59.69	0.53	1.00	1004.4	18.01	1	1	1	0	0.80	0.83	94.90	119.38

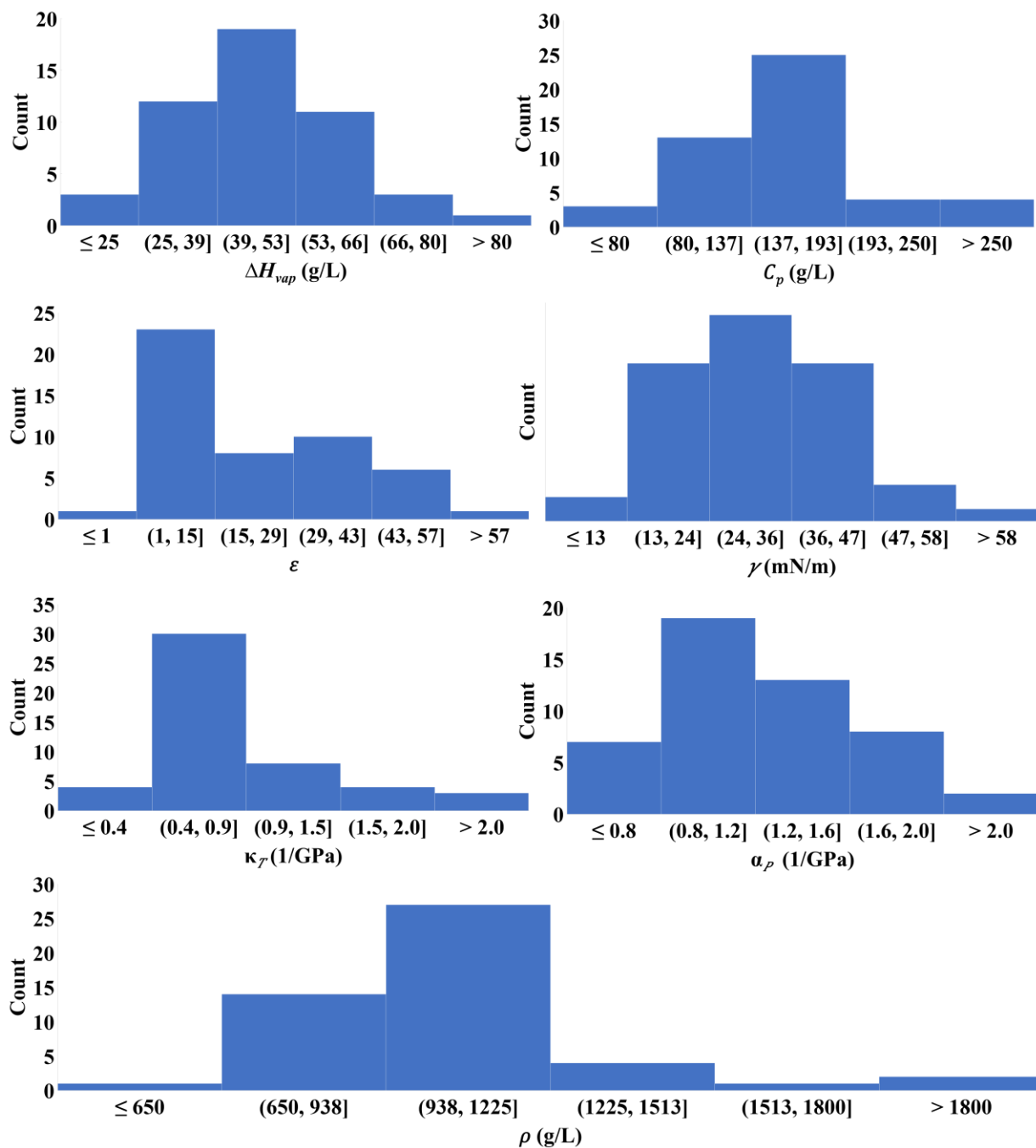


Figure B1. Histogram of macroscopic descriptors calculated for the 49 solvents.

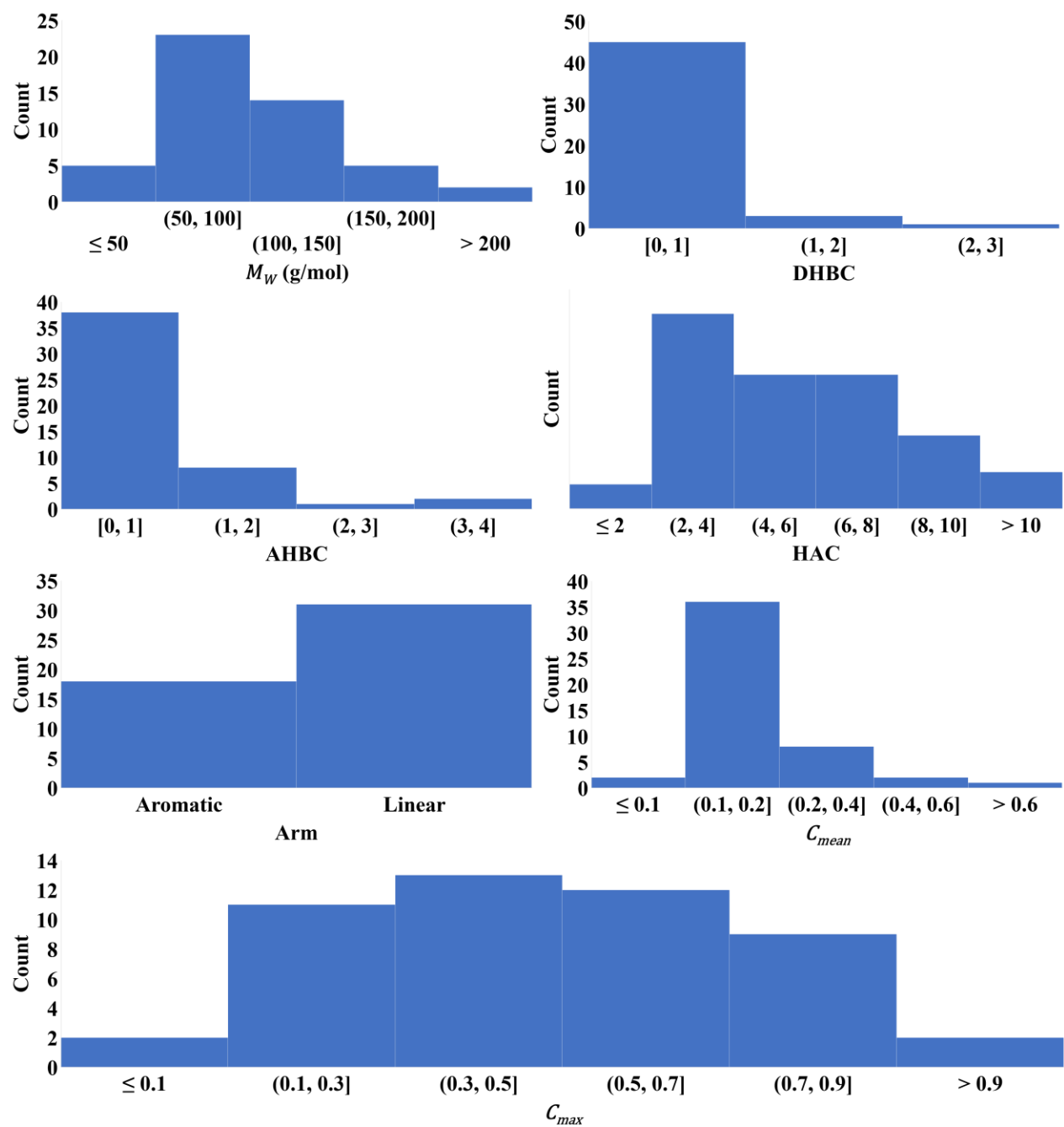


Figure B2. Histogram of molecular descriptors calculated for the 49 solvents.

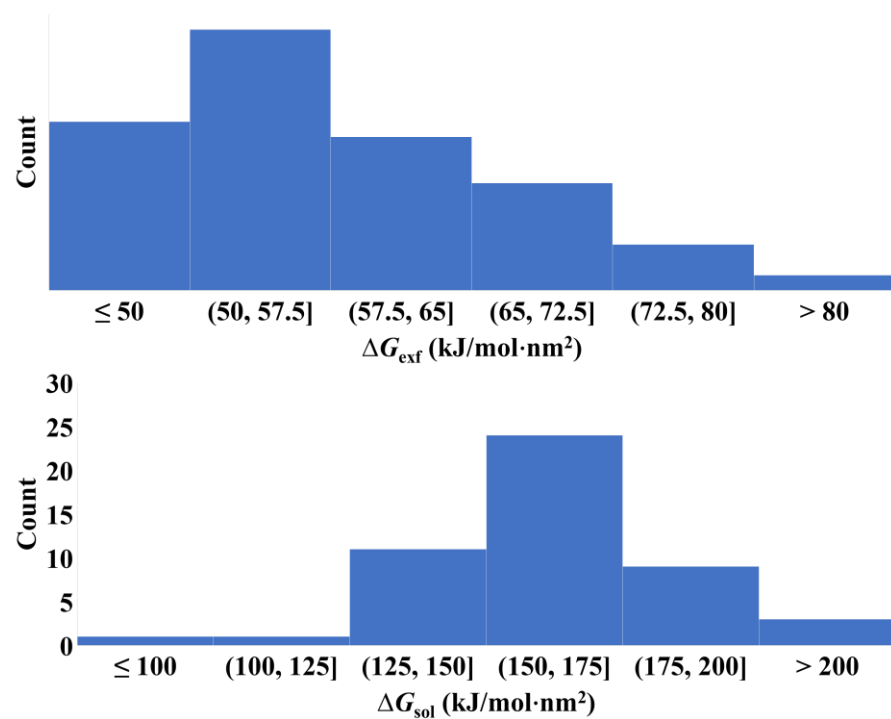


Figure B3. Histogram of target properties calculated for the 49 solvents.

## B2. Determination of outlier

In Tukey's method [1], the inter-quartile range (IQR) is defined as the interval between the first ( $Q_1$ ) and the third ( $Q_3$ ) quartiles. Data less than  $Q_1 - (1.5 \times \text{IQR})$  or more than  $Q_3 + (1.5 \times \text{IQR})$  are considered outliers. In our initial dataset that included all 49 solvents,  $Q_1$  and  $Q_3$  for  $\Delta G_{\text{exf}}$  are respectively 50.79 and 62.83 kJ/mol·nm<sup>2</sup>, and  $\text{IQR} = Q_3 - Q_1 = 12.04$  kJ/mol·nm<sup>2</sup>. Water has a  $\Delta G_{\text{exf}}$  of 119.40 kJ/mol·nm<sup>2</sup> which is higher than  $Q_3 + (1.5 \times \text{IQR}) = 80.89$  kJ/mol·nm<sup>2</sup>.  $Q_1$  and  $Q_3$  for  $\Delta G_{\text{sol}}$  are respectively 149.59 and 174.96 kJ/mol·nm<sup>2</sup>, and  $\text{IQR} = Q_3 - Q_1 = 25.37$  kJ/mol·nm<sup>2</sup>. Water has a  $\Delta G_{\text{sol}}$  of 94.90 kJ/mol·nm<sup>2</sup> which is less than  $Q_1 - (1.5 \times \text{IQR}) = 111.53$  kJ/mol·nm<sup>2</sup>. It is therefore identified as an outlier and removed from ML modeling.

### B3. Descriptor selection

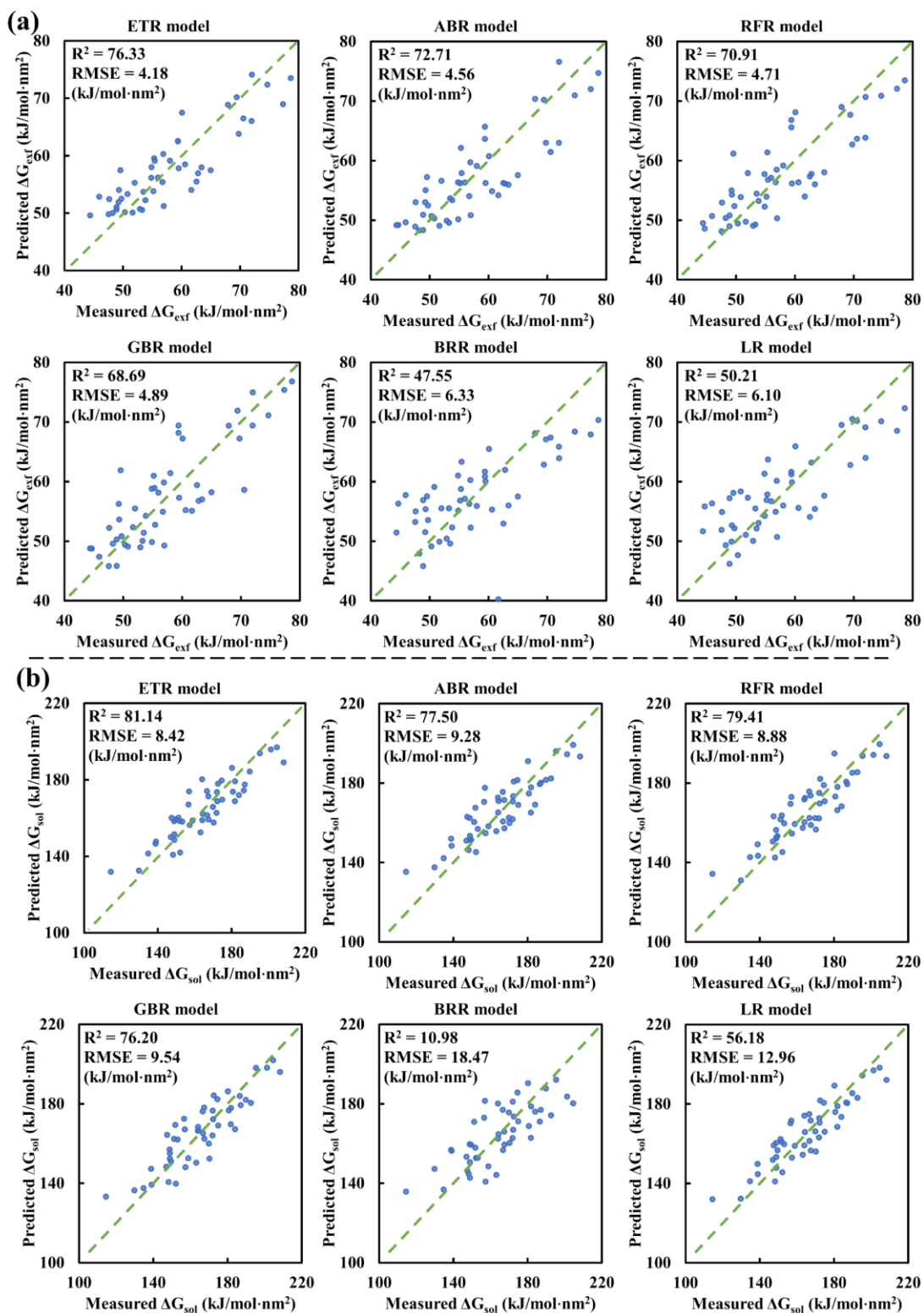


Figure B4. Predicted vs measured values for (a)  $\Delta G_{\text{exf}}$  and (b)  $\Delta G_{\text{sol}}$  with the selected descriptors.  $R^2$  and RMSE of each model are also shown.

Table B3. Optimal parameters obtained from hyper-parameter tuning after descriptor selection.

ML models	Hyper-parameters
ETR	bootstrap: False; cost-complexity pruning (ccp): 0.0, criterion: squared error; maximum depth: None; maximum features: auto; maximum leaf nodes: None; maximum samples: None; minimum impurity decrease: 0.0; minimum samples leaf: 1; minimum samples split: 2; minimum weight fraction leaf: 0.0; number of estimators: 100; out-of-bag (oob) score: False
RFR	bootstrap: True; ccp: 0.0; criterion: squared error; maximum depth: None; maximum features: auto; maximum leaf nodes: None; maximum samples: None; minimum impurity decrease: 0.0; minimum samples leaf: 1; minimum samples split: 2; minimum weight fraction leaf: 0.0; number of estimators: 100; oob score: False
GBR	alpha: 0.9; ccp: 0.0; criterion: Friedman mean squared error; learning rate: 0.1; loss: squared error; maximum depth: 3; maximum features: None; maximum leaf nodes: None; minimum impurity decrease: 0.0; minimum samples leaf: 1; minimum samples split: 2; minimum weight fraction leaf: 0.0; number of estimators: 100; subsample: 1.0; tolerance: 0.0001; validation fraction: 0.1; verbose: 0
ABR	base estimator: None; learning rate: 1.0; loss: linear; number of estimators: 50
GBR	Alpha <sub>1</sub> : 10 <sup>-6</sup> ; Alpha <sub>2</sub> : 10 <sup>-6</sup> ; Alpha <sub>init</sub> : None; compute score: False; fit intercept: True; lambda <sub>1</sub> : 10 <sup>-6</sup> ; lambda <sub>2</sub> : 10 <sup>-6</sup> ; lambda <sub>init</sub> : None; number of iterations: 300; normalize: deprecated; tolerance: 0.001

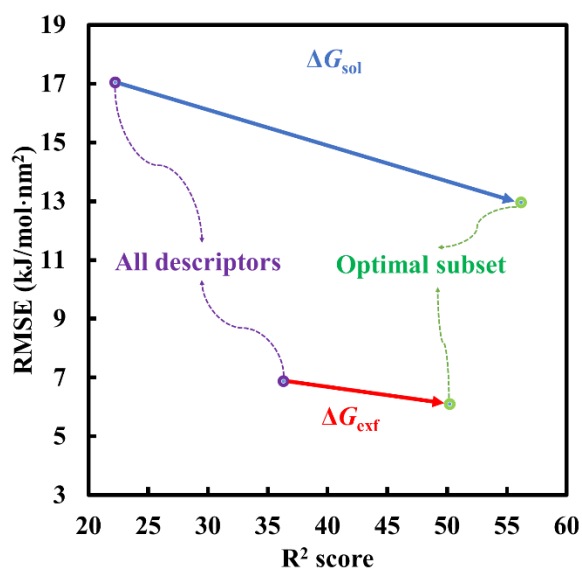


Figure B5. RMSE and  $R^2$  score for the LR model with all descriptors and the optimal subset identified from ASM. The arrows indicate the improvement of each target property after descriptor selection.

## B4. Model deployment

Table B4.  $\Delta G_{\text{exf}}$  (kJ/mol.nm<sup>2</sup>) predicted for 49 solvents from our dataset (along with measured values from MD) and 105 solvents in Coleman et al. [2]. The prediction here is made with ETR and the optimal subset of descriptors.

Solvent	Measured $\Delta G_{\text{exf}}$	Predicted $\Delta G_{\text{exf}}$	Solvent	Measured $\Delta G_{\text{exf}}$	Predicted $\Delta G_{\text{exf}}$
1. Chloroform	59.41	58.86	50. Dichloromethane	-	59.77
2. Dichlorofluoromethane	60.09	60.80	51. Methanal	-	72.87
3. Dibromomethane	49.56	51.74	52. Bromomethane	-	59.09
4. Methanoic acid	55.33	70.58	53. Nitromethane	-	57.58
5. Formamide	74.69	61.15	54. 1,1,2,2-tetrachloroethane	-	52.70
6. Methanol	78.68	63.36	55. 1,1-dichloroethene	-	60.72
7. Pentachloroethane	53.49	54.87	56. 1,1,2-trichloroethane	-	54.17
8. Acetamide	56.76	58.06	57. Acetonitrile	-	67.54
9. N-Methylformamide	50.79	57.69	58. 1,2-dibromoethane	-	54.36
10. Ethanol	67.96	58.76	59. 1,1-dichloroethane	-	59.60
11. Dimethyl sulfoxide	55.19	57.45	60. 1,2-dichloroethane	-	56.50
12. Ethanolamine	71.98	57.16	61. Methyl formate	-	67.18
13. Ethane-1,2-diamine	69.79	55.07	62. Bromoethane	-	58.64
14. gamma-Butyrolactone	47.55	55.27	63. Chloroethane	-	67.75
15. Propanenitrile	70.58	64.96	64. 2-chloroethanol	-	51.04
16. 1,2-Dibromopropane	51.66	54.54	65. Nitroethane	-	56.19
17. Methyloxirane	72.04	71.17	66. Methoxymethane	-	74.71
18. Acetone	69.44	61.52	67. 1,2-ethanedithiol	-	55.41
19. Tetrahydrofuran	59.41	69.02	68. Methyl disulfanylmethane	-	57.72
20. Dimethylformamide	54.85	53.05	69. Methylsulfanylmethane	-	66.52
21. N-Methylacetamide	53.28	52.07	70. Prop-2-enenitrile	-	67.76
22. Glycerol	61.68	61.83	71. 1,3-dioxolan-2-one	-	61.08
23. Isobutane	77.36	73.00	72. 1,3-dichloropropane	-	55.96
24. Butane-1,4-diol	62.83	53.83	73. Methyl acetate	-	59.85
25. 2-Methylpropan-2-amine	55.45	62.09	74. 1,3-dioxolane	-	58.93
26. Pyrimidine	49.34	51.90	75. 2-iodopropane	-	52.93
27. 1,4-dichlorobutane	60.6	52.94	76. 1-bromopropane	-	57.49
28. Dimethylacetamide	45.95	61.60	77. 1-nitropropane	-	55.70
29. 1,2,3,5-Tetrafluorobenzene	52.01	56.04	78. 2-nitropropane	-	56.96
30. 1,3-Difluorobenzene	54.89	58.56	79. Dimethoxymethane	-	61.38
31. Benzenethiol	65.01	54.17	80. Propan-1-amine	-	66.61
32. 4-Methylpyridine	55.99	55.02	81. Propan-2-amine	-	68.58
33. Cyclohexanone	63.42	55.59	82. Ethylsulfanylethane	-	57.70
34. Triethyl phosphate	48.26	50.23	83. Butane-1-thiol	-	57.41
35. N,N-diethylethanamine	47.64	59.95	84. Butan-1-ol	-	52.50
36. Benzonitrile	56.98	48.58	85. 2-methylpropan-2-ol	-	58.48
37. Toluene	56.91	59.92	86. (2-hydroxyethoxy)ethan-2-ol	-	56.82
38. 3-Methylphenol	50.34	54.31	87. N-ethylethanamine	-	60.10
39. 1,2-Dimethoxybenzene	52.9	53.73	88. Butan-1-amine	-	57.54
40. 1-Butoxybutane	48.9	59.10	89. 2-(2-hydroxyethylamino)ethanol	-	55.86
41. Quinoline	44.37	52.18	90. Furan	-	63.90
42. 2,6-Dimethyl-4-heptanone	53.84	51.89	91. Thiophene	-	61.06
43. Phenoxybenzene	48.91	54.68	92. 1H-pyrrole	-	54.97
44. Chlorobenzene	58.04	57.93	93. Ethenyl acetate	-	58.26
45. 1-Cyclohexylpyrrolidin-2-one	62.56	58.22	94. Acetyl acetate	-	52.06
46. Dichlorobenzene	49.26	54.55	95. Ethoxyethene	-	62.28
47. Propan-2-ol	59.52	59.40	96. Ethyl acetate	-	57.75
48. N-Vinylpyrrolidone	49.72	52.27	97. Tetrahydrothiophene 1,1-dioxide	-	58.80
49. Water	119.38	74.35	98. Thiolane	-	58.26
			99. 1-bromobutane	-	56.20



100. 1-chlorobutane	-	58.84
101. Pyrrolidine	-	57.25
102. Morpholine	-	55.84
103. Pyridine	-	57.20
104. Cyclopentanone	-	56.24
105. 1-cyclopropylethanone	-	58.11
106. Pentane-2,4-dione	-	56.24
107. Methyl 2-methylprop-2-enoate	-	59.22
108. Pentanenitrile	-	56.72
109. Ethyl propanoate	-	58.46
110. Diethyl carbonate	-	61.89
111. Pentan-1-ol	-	51.27
112. Pentan-3-ol	-	51.59
113. 2-methylbutan-2-ol	-	54.45
114. Pentane-1,5-diol	-	53.90
115. Pentan-3-amine	-	58.14
116. 1,2,3,4-tetrafluorobenzene	-	55.37
117. 1,2-difluorobenzene	-	57.62
118. Fluorobenzene	-	59.21
119. Nitrobenzene	-	52.16
120. 2-chloroaniline	-	50.12
121. Phenol	-	51.89
122. 2-methylpyridine	-	56.28
123. 3-methylpyridine	-	54.86
124. (E)-hex-2-ene	-	63.13
125. Hexan-2-one	-	57.57
126. 2,4,6-trimethyl-1,3,5-trioxane	-	56.99
127. Cyclohexanamine	-	57.03
128. 2-propan-2-yloxypropane	-	59.14
129. 1-methoxy-2-(2-methoxyethoxy)ethane	-	56.69
130. N-propan-2-ylpropan-2-amine	-	59.62
131. Trifluoromethylbenzene	-	56.69
132. Benzaldehyde	-	58.71
133. Methoxybenzene	-	57.52
134. Phenylmethanol	-	49.11
135. 2-methylphenol	-	53.91
136. 4-methylphenol	-	50.60
137. Diethyl propanedioate	-	52.35
138. 2,4-dimethylpentan-3-one	-	57.41
139. Heptan-2-one	-	55.66
140. Ethenylbenzene	-	58.95
141. 1-phenylethanone	-	50.28
142. Methyl benzoate	-	53.26
143. Methyl 2-hydroxybenzoate	-	51.14
144. Ethylbenzene	-	60.49
145. 1,2-dimethylbenzene	-	59.62
146. 2,4,6-trimethylpyridine	-	51.96
147. Octan-1-ol	-	51.00
148. N-butylbutan-1-amine	-	55.67
149. Isoquinoline	-	53.79
150. (1-methylethyl)benzene	-	57.45
151. 1,2,4-trimethylbenzene	-	57.68
152. 1-chloronaphthalene	-	53.04
153. N-Methyl-2-pyrrolidone		49.43
154. Methanesulfonic acid		49.89

## B5. Experimental details

1-Methyl-2-pyrrolidinone (NMP), Benzyl alcohol (Bn), Acetonitrile (ACN) were purchased from Fisher Scientific. Melamine (99%) and Methanesulfonic acid (MSA) (99%) were purchased from Acros. All solvents were of HPLC grade and used as received.

Graphitic carbon nitride was synthesized by heating melamine in a ceramic crucible from room temperature to 550 C at the rate of 5 C/min, and maintaining the temperature for 2 hours within the muffle furnace. The crucible was then allowed to settle naturally to room temperature, and the obtained light-yellow product was ground into a fine powder for additional experimental work. The samples were examined by various methods to confirm the synthesis of traditional carbon nitride. The X-ray diffraction (XRD) was collected with a Rigaku Ultimate IV equipped with a Cu-K $\alpha$  radiation source (40 kV,  $\lambda=0.15418$  nm). Raman active vibrational modes of the material were investigated with a Raman spectrometer (Renishaw InVia Raman Microscope) operating at the 532 nm laser excitation line. The powder sample was deposited on a glass slides and spectra were accumulated using 50um confocal pinhole aperture slit. The absorption spectrum of the bulk carbon nitride in the UV/VIS region was determined using diffuse reflectance spectroscopy UV/VIS spectrophotometer (Hitachi U-3900H). For the dispersibility tests, the aquasonicator model 75T from VWR Scientific Products, operating at 117/120 V, 2 A, and 50/60 Hz, was used to disperse bulk carbon nitride in the four different solvents. The sonication was conducted for 15 mins under the temperature of 26°C. The dispersibility of carbon nitride is photographically recorded immediately after sonication and over one hour keeping it in stable condition.

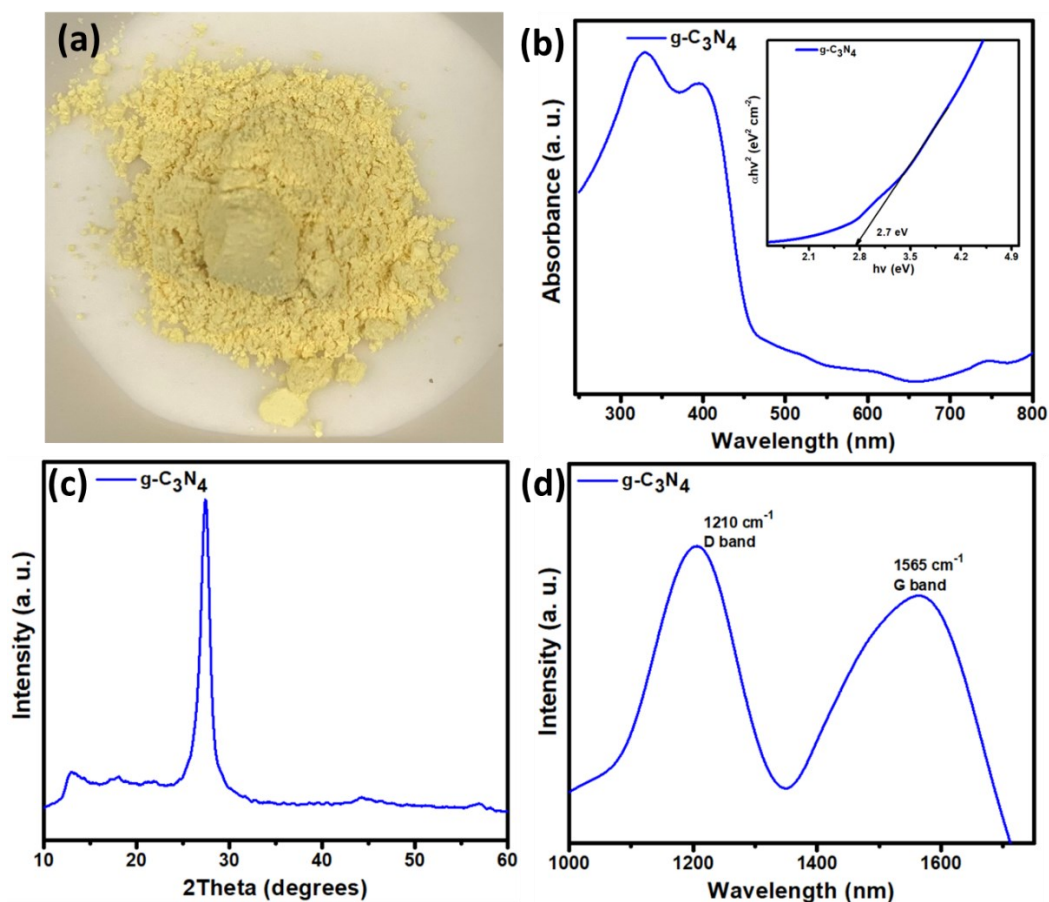


Figure B6. (a) Macroscopic image, (b) DRS UV/Vis spectra, (c) XRD pattern, and (d) Raman spectra of the carbon nitride powder.

The pale-yellow colored powder (Figure B6a) suggests the successful synthesis of carbon nitride using melamine as the precursor. Figure B6b shows the DRS UV/VIS absorbance spectrum of g-C<sub>3</sub>N<sub>4</sub> thin film. The spectrum demonstrates the intrinsic absorption edge at 420 nm [3]. The Kubelka-Munk plot shows that the bandgap of the g-C<sub>3</sub>N<sub>4</sub> is 2.7 eV [4]. Figure B6c displays the XRD pattern, and it clearly demonstrates the characteristic diffraction peaks at 13.05° and 27.5° [5]. The modest diffraction peak at 13.05 ° corresponds to the (100) plane of heptazine units. The strong peak at 27.5 ° is assigned to the stacking of conjugated aromatic rings and designated as (002) plane. The extremely intense diffraction pattern developed from interlayer stacking of carbon nitride along the hexagonal c-axis was observed for melamine derived g-C<sub>3</sub>N<sub>4</sub> material [6]. Figure B6d shows the Raman spectra for the carbon nitride where the D band (around 1210 cm<sup>-1</sup>) exists

due to the  $sp^3$  defects, while the G band (around  $1650\text{ cm}^{-1}$ ) holds the signature of the in-plane vibration of  $sp^2$  carbon in the graphitic carbon nitride network [7,8].

## B6. Additional plots

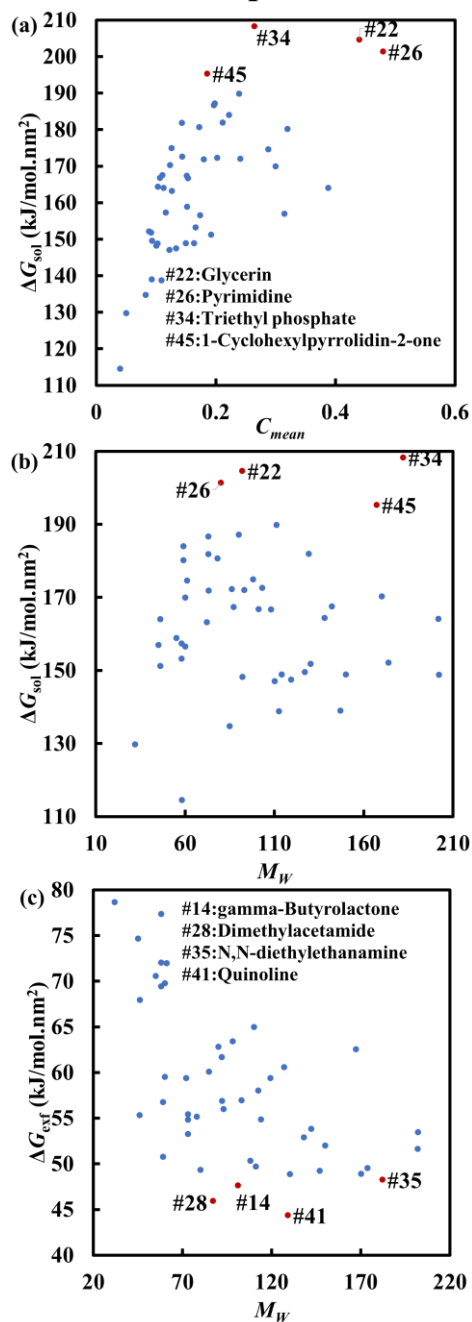


Figure B7. Pair plots of (a)  $\Delta G_{\text{sol}}$  vs.  $C_{\text{mean}}$ , (b)  $\Delta G_{\text{sol}}$  vs.  $M_W$ , and (c)  $\Delta G_{\text{exf}}$  vs.  $M_W$ .  $\Delta G_{\text{sol}}$  and  $\Delta G_{\text{exf}}$  are measured (from MD) values for the 48 organic solvents in our dataset. Four solvents with the lowest  $\Delta G_{\text{exf}}$  and highest  $\Delta G_{\text{sol}}$  are highlighted with red color.

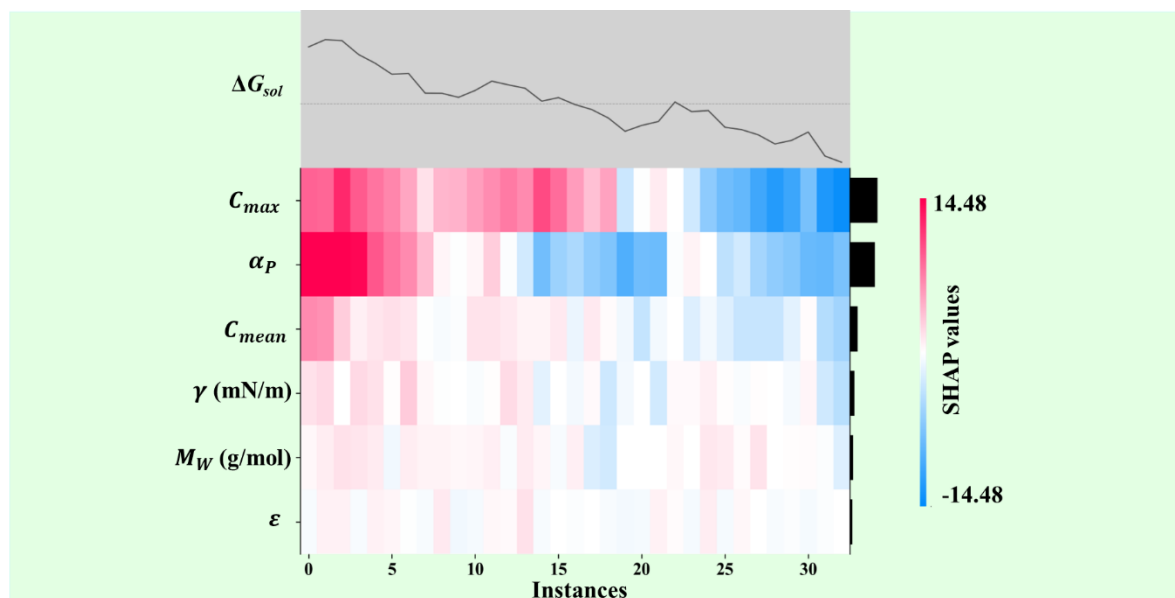


Figure B8. Heatmap of SHAP values for each descriptor in the ETR model. The color intensity represents the magnitude of the SHAP value, with red indicating a positive impact and blue indicating a negative impact on the predicted  $\Delta G_{sol}$ .

## References

- [1] A. Gelman, Exploratory Data Analysis for Complex Models, [Http://Dx.Doi.Org/10.1198/106186004X11435](http://dx.doi.org/10.1198/106186004X11435) 13 (2012) 755–779. <https://doi.org/10.1198/106186004X11435>.
- [2] C. Caleman, P.J. Van Maaren, M. Hong, J.S. Hub, L.T. Costa, D. Van Der Spoel, Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant, *J Chem Theory Comput* 8 (2012) 61–74. <https://doi.org/10.1021/ct200731v>.
- [3] P. Niu, M. Qiao, Y. Li, L. Huang, T. Zhai, Distinctive defects engineering in graphitic carbon nitride for greatly extended visible light photocatalytic hydrogen evolution, *Nano Energy* 44 (2018) 73–81. <https://doi.org/10.1016/J.NANOEN.2017.11.059>.
- [4] K. Ogata, J. Fernando, D. Chubaci, F. Fujimoto, J. Fernando, F. Fujinioto, Properties of carbon nitride films with composition ratio C/N=0.5–3.0 prepared by the ion and vapor deposition method, *J Appl Phys* 76 (1994) 3791–3796. <https://doi.org/10.1063/1.358497>.
- [5] P. Kumar, E. Vahidzadeh, U.K. Thakur, P. Kar, K.M. Alam, A. Goswami, N. Mahdi, K. Cui, G.M. Bernard, V.K. Michaelis, K. Shankar, C<sub>3</sub>N<sub>5</sub>: A Low Bandgap Semiconductor Containing an Azo-Linked Carbon Nitride Framework for Photocatalytic, Photovoltaic and Adsorbent Applications, *J Am Chem Soc* 141 (2019) 5415–5436. [https://doi.org/10.1021/JACS.9B00144/SUPPL\\_FILE/JA9B00144\\_SI\\_001.PDF](https://doi.org/10.1021/JACS.9B00144/SUPPL_FILE/JA9B00144_SI_001.PDF).
- [6] T. Mahvelati-Shamsabadi, H. Fattahimoghaddam, B.K. Lee, S. Bae, J. Ryu, Synthesis of hexagonal rosettes of g-C<sub>3</sub>N<sub>4</sub> with boosted charge transfer for the enhanced visible-light photocatalytic hydrogen evolution and hydrogen peroxide production, *J Colloid Interface Sci* 597 (2021) 345–360. <https://doi.org/10.1016/J.JCIS.2021.04.019>.
- [7] N. Chaulagain, K.M. Alam, S. Kadian, N. Kumar, J. Garcia, G. Manik, K. Shankar, Synergistic Enhancement of the Photoelectrochemical Performance of TiO<sub>2</sub>Nanorod Arrays through Embedded Plasmon and Surface Carbon Nitride Co-sensitization, *ACS Appl Mater Interfaces* 14 (2022) 24309–24320. [https://doi.org/10.1021/ACSAMI.2C02649/SUPPL\\_FILE/AM2C02649\\_SI\\_001.PDF](https://doi.org/10.1021/ACSAMI.2C02649/SUPPL_FILE/AM2C02649_SI_001.PDF).

- [8] K.M. Alam, N. Chaulagain, E. Shahini, M. Masud Rana, J. Garcia, N. Kumar, A.E. Kobryn, S. Gusarov, T. Tang, K. Shankar, Low bandgap carbon nitride nanoparticles incorporated in titania nanotube arrays by in situ electrophoretic anodization for photocatalytic CO<sub>2</sub> reduction, *Chemical Engineering Journal* 456 (2023) 141067. <https://doi.org/10.1016/J.CEJ.2022.141067>.



# Appendix C

## Supporting Information for Chapter 5

### C1. Organic solvents used for machine learning modeling

- |   |                                  |                                   |
|---|----------------------------------|-----------------------------------|
| 1. (1-methylethyl)benzene               | 58. Benzaldehyde                 | 115. Methoxybenzene               |
| 2. (2-hydroxyethoxy)ethan-2-ol          | 59. Benzenethiol                 | 116. Methoxymethane               |
| 3. (2R)-2-methyloxirane                 | 60. Benzonitrile                 | 117. Methyl 2-hydroxybenzoate     |
| 4. (E)-hex-2-ene                        | 61. Benzyl-benzoate              | 118. Methyl 2-methylprop-2-enoate |
| 5. 1,1,1,2,2-pentachloroethane          | 62. Benzylether                  | 119. Methyl acetate               |
| 6. 1,1,2,2-tetrachloroethane            | 63. Bromobenzene                 | 120. Methyl benzoate              |
| 7. 1,1,2-trichloroethane                | 64. Bromoethane                  | 121. Methyl formate               |
| 8. 1,1-dichloroethane                   | 65. Bromomethane                 | 122. Methyl disulfanylmethane     |
| 9. 1,1-dichloroethene                   | 66. Butan-1-amine                | 123. Methyl sulfanylmethane       |
| 10. 1,2,3,4-tetrafluorobenzene          | 67. Butan-1-ol                   | 124. Methyl sulfinylmethane       |
| 11. 1,2,3,5-tetrafluorobenzene          | 68. Butane-1,4-diol              | 125. 4-Methylpentan-2-one         |
| 12. 1,2,4-trimethylbenzene              | 69. Butane-1-thiol               | 126. Morpholine                   |
| 13. 1,2-dibromoethane                   | 70. Chlorobenzene                | 127. Methanesulfonic acid         |
| 14. 1,2-dibromopropane                  | 71. Chloroethane                 | 128. 2-Methoxy-2-methylpropane    |
| 15. 1,2-dichloroethane                  | 72. Chloroform                   | 129. N,N-diethylethanamine        |
| 16. 1,2-difluorobenzene                 | 73. 1-Cyclohexylpyrrolidin-2-one | 130. N,N-dimethylacetamide        |
| 17. 1,2-dimethoxybenzene                | 74. Cyclohexanamine              | 131. N,N-dimethylformamide        |
| 18. 1,2-dimethylbenzene                 | 75. Cyclohexane                  | 132. N-dodecyl-pyrrolidone        |
| 19. 1,2-ethanedithiol                   | 76. Cyclohexanol                 | 133. N-butylbutan-1-amine         |
| 20. 1,3-dichloropropane                 | 77. Cyclohexanone                | 134. N-ethylethanamine            |
| 21. 1,3-difluorobenzene                 | 78. Cyclopentanone               | 135. 1-formylpiperidine           |
| 22. 1,3-dioxolan-2-one                  | 79. Dichlorobenzene              | 136. Nitrobenzene                 |
| 23. 1,3-dioxolane                       | 80. Diethyl phthalate            | 137. Nitroethane                  |
| 24. 1,4-dichlorobutane                  | 81. Diaminodiethylamine          | 138. Nitromethane                 |
| 25. 1-bromobutane                       | 82. Dibromomethane               | 139. N-methylacetamide            |
| 26. 1-bromopropane                      | 83. Dichloro(fluoro)methane      | 140. N-methylformamide            |
| 27. 1-butoxybutane                      | 84. Dichloromethane              | 141. N-Methylpyrrolidone          |
| 28. 1-chlorobutane                      | 85. Diethyl carbonate            | 142. N-propan-2-ylpropan-2-amine  |
| 29. 1-chloronaphthalene                 | 86. Diethyl propanedioate        | 143. N-Vinyl-2-pyrrolidone        |
| 30. 1-cyclopropylethanone               | 87. Dimethoxymethane             | 144. Octan-1-ol                   |
| 31. 1H-pyrrole                          | 88. Dimethylethyleneurea         | 145. Oxolan-2-one                 |
| 32. 1-methoxy-2-(2-methoxyethoxy)ethane | 89. Pyrimidinone                 | 146. Oxolane                      |
| 33. 1-nitropropane                      | 90. Epichlorohydrin              | 147. Pentan-1-ol                  |
| 34. 1-phenylethanone                    | 91. Ethanamide                   | 148. Pentan-3-amine               |

- |                                    |                          |                                      |
|------------------------------------|--------------------------|--------------------------------------|
| 35. 2-(2-hydroxyethylamino)ethanol | 92. Ethane-1,2-diamine   | 149. Pentan-3-ol                     |
| 36. 2,4,6-trimethyl-1,3,5-trioxane | 93. Ethanol              | 150. Pentane-1,5-diol                |
| 37. 2,4,6-trimethylpyridine        | 94. Ethenyl acetate      | 151. Pentane-2,4-dione               |
| 38. 2,4-dimethylpentan-3-one       | 95. Ethenylbenzene       | 152. Pentanenitrile                  |
| 39. 2,6-dimethylheptan-4-one       | 96. Ethoxyethene         | 153. Phenol                          |
| 40. 2-aminoethanol                 | 97. Ethyl acetate        | 154. Phenoxybenzene                  |
| 41. 2-chloroaniline                | 98. Ethyl propanoate     | 155. Phenylmethanol                  |
| 42. 2-chloroethanol                | 99. Ethylbenzene         | 156. Prop-2-enenitrile               |
| 43. 2-iodopropane                  | 100. Ethylene glycol     | 157. Propan-1-amine                  |
| 44. 2-methylbutan-2-ol             | 101. Ethylene-glycol     | 158. Propan-2-amine                  |
| 45. 2-methylphenol                 | 102. Ethylsulfanylethane | 159. Propan-2-one                    |
| 46. 2-methylpropan-2-amine         | 103. Fluorobenzene       | 160. Propane-1,2,3-triol             |
| 47. 2-methylpropan-2-ol            | 104. Furan               | 161. Propanenitrile                  |
| 48. 2-methylpropane                | 105. Heptan-2-one        | 162. Pyridine                        |
| 49. 2-methylpyridine               | 106. Hexan-2-one         | 163. Pyrimidine                      |
| 50. 2-nitropropane                 | 107. Hexane              | 164. Pyrrolidine                     |
| 51. 2-propan-2-yloxypropane        | 108. Isopropanol         | 165. Quinoline                       |
| 52. 3-methylphenol                 | 109. Isoquinoline        | 166. Tetrahydrothiophene 1,1-dioxide |
| 53. 3-methylpyridine               | 110. Methylacetone       | 167. Thiolane                        |
| 54. 4-methylphenol                 | 111. Methanal            | 168. Thiophene                       |
| 55. 4-methylpyridine               | 112. Methanamide         | 169. Toluene                         |
| 56. Acetonitrile                   | 113. Methanoic acid      | 170. Triethyl phosphate              |
| 57. Acetyl acetate                 | 114. Methanol            | 171. Trifluoromethylbenzene          |

## C2. Definition of solvation layers

In order to define the solvent layers for each solvent, the RDF of solvent molecules around nanosheet atoms were calculated from the simulations where a single nanosheet was solvated and equilibrated. Figure C1 shows these RDF curves. The first solvation layer is defined based on where the first prominent peak is observed. For each system, all solvent molecules whose COM is within this distance are considered part of the first solvation layer. A similar approach is used to define the second and third solvation layers, based on the second and third prominent peaks in the RDF curves. If the peaks are not evident in the RDF curves, then the second and third solvation layers are defined as double and triple the distance of the first peak, respectively. It should be noted that in some systems, such as pure NMP and pure MET, there is a peak around 0.2 nm which corresponds to hydrogen bonding and is neglected in the solvation layer definition. The cutoff distances for each system are listed in Table C1.

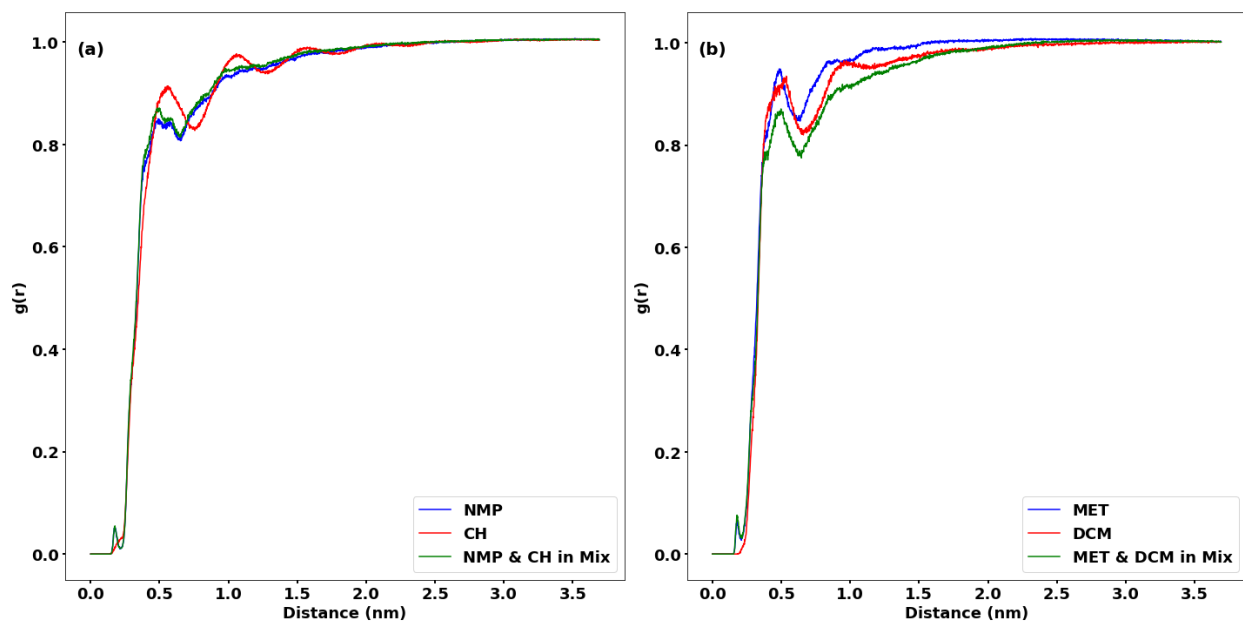


Figure C1. The RDF curves of (a) NMP (in pure NMP), CH (in pure CH), NMP & CH (in NMP:CH mixture), (b) MET (in pure MET), DCM (in pure DCM), and MET & DCM (in MET:DCM mixture) molecules around all atoms of a single nanosheet.

Table C1. Cutoff distances for defining the first, second and third solvation layers.

System	1st solvation layer distance (nm)	2nd solvation layer distance (nm)	3rd solvation layer distance (nm)
NMP	0.50	1.00	1.50
CH	0.56	1.10	1.60
NMP:CH Mixture	0.50	1.00	1.50
MET	0.50	1.00	1.50
DCM	0.53	1.06	1.59
MET:DCM Mixture	0.50	1.00	1.50

### C3. H-bond lifetime calculation

To calculate the lifetime of hydrogen bonds (H-bonds), the following autocorrelation function is used [1]:

$$C(\tau) = \langle s_i(t) s_i(t + \tau) \rangle \quad (C1)$$

Here,  $s_i(t)$  is an indicator function that records whether hydrogen bond  $i$  is present ( $s_i = 1$ ) or absent ( $s_i = 0$ ) at time  $t$ . The angular brackets  $\langle \bullet \rangle$  denote an average over all time  $t$  and all H-bonds  $i$  between the solvent and the nanosheets. By integrating this autocorrelation function over time, we obtain an estimate of the average hydrogen bond lifetime [1]:

$$\text{Average H-bond lifetime} \approx \int_0^\infty C(\tau) d\tau \quad (C2)$$

This method provides a quantitative measure of the duration for which a hydrogen bond persists, giving insights into the dynamic interactions between the solvent molecules and the nanosheets.

#### C4. Organic solvents used for studying their mixture with water

- |                           |                          |                            |
|---------------------------|--------------------------|----------------------------|
| 1. 2-Aminoethanol         | 10. Ethane-1,2-Diamine   | 19. N-Methyl-2-Pyrrolidone |
| 2. 2-Methylpropan-2-Amine | 11. Formaldehyde         | 20. N-Vinylpyrrolidone     |
| 3. Acetamide              | 12. Gamma-Butyrolactone  | 21. Propane-1,2,3-Triol    |
| 4. Acetone                | 13. Isopropyl Alcohol    | 22. Propanenitrile         |
| 5. 1,4-Butanediol         | 14. Methanoic Acid       | 23. Pyrimidine             |
| 6. Dimethylacetamide      | 15. Methanol             | 24. Tetrahydrofuran        |
| 7. Dimethylformamide      | 16. Methanesulfonic Acid | 25. Triethyl Phosphate     |
| 8. Dimethyl Sulfoxide     | 17. N-Methylacetamide    |                            |
| 9. Ethanol                | 18. N-Methylformamide    |                            |

#### References

- [1] A. Luzar, D. Chandler, Hydrogen-bond kinetics in liquid water, *Nature* 1996 379:6560 379 (1996) 55–57. <https://doi.org/10.1038/379055a0>.

# Appendix D

## Supporting Information for Chapter 6

### D1. Calculation of perimeter and area for determining the periphery-to-interior ratio ( $\alpha$ )

To determine  $\alpha$  for a pristine g-C<sub>3</sub>N<sub>4</sub> nanosheet, the pristine nanosheet first undergoes energy minimization in vacuum using the steepest descent method to ensure structural stability. Then, the nitrogen atoms located in the periphery region of the nanosheet are considered boundary atoms (for the definition of the periphery region, refer to Figure 6.1). These atoms are sorted in sequence to form a closed loop around the nanosheet. To calculate the perimeter, the Euclidean distance between each pair of consecutive boundary atoms is determined. Denoting the plane of the nanosheet as the  $x$ - $y$  plane, for atoms  $i$  and  $j$  with coordinates  $(x_i, y_i)$  and  $(x_j, y_j)$ , respectively, the distance  $d_{i,j}$  is computed as follows:

$$d_{i,j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (\text{D1})$$

The perimeter  $P$  is then obtained by summing these distances:

$$P = \sum_{i=1}^{N-1} d_{i,i+1} + d_{N,1} \quad (\text{D2})$$

where  $N$  is the total number of boundary atoms, and  $d_{N,1}$  is the distance between the last and the first boundary atoms to close the loop.

The area of the pristine nanosheet is calculated by discretizing the nanosheet into smaller triangles. Each of these smaller triangles is formed by connecting the boundary atoms to the center of the nanosheet or by using a Delaunay triangulation method. For each small triangle  $i$  with vertices located at  $(x_{i1}, y_{i1})$ ,  $(x_{i2}, y_{i2})$ , and  $(x_{i3}, y_{i3})$ , the area  $A_i$  is calculated using the determinant method:

$$A_i = \frac{1}{2} |x_{i1}(y_{i2} - y_{i3}) + x_{i2}(y_{i3} - y_{i1}) + x_{i3}(y_{i1} - y_{i2})| \quad (\text{D3})$$

The total area  $A_{total}$  of the nanosheet is then determined by summing the areas of all the small triangles:

$$A_{total} = \sum_{i=1}^M A_i \quad (D4)$$

where  $M$  is the total number of small triangles formed. With the perimeter  $P$  and the total area  $A_{total}$  calculated,  $\alpha$  is defined as:

$$\alpha = \frac{P}{A_{total}} \quad (D5)$$

The calculations for the functionalized nanosheets are conducted in the same manner, involving energy minimization followed by evaluations using Eq. (D1)-(D5). The results for  $\alpha$  show less than 1% difference across different nanosheets, allowing a single value ( $0.32 \text{ nm}^{-1}$  for 1L and  $0.64 \text{ nm}^{-1}$  for 2L) to be used for different functionalizations.

## D2. Organic solvents used for studying different functionalizations

- |                                 |                          |                            |
|---------------------------------|--------------------------|----------------------------|
| 1. 1,2-Dibromopropane           | 18. Benzonitrile         | 35. Methanoic Acid         |
| 2. 1,2-Diaminoethane            | 19. Benzenethiol         | 36. Methanol               |
| 3. 1,2-Dimethoxybenzene         | 20. Chloroform           | 37. Methyloxirane          |
| 4. 1,3-Difluorobenzene          | 21. Chlorobenzene        | 38. N,N-Diethylethanamine  |
| 5. 1,4-Butanediol               | 22. Cyclohexanone        | 39. N-Methyl-2-Pyrrolidone |
| 6. 1,4-Dichlorobutane           | 23. Dibromomethane       | 40. N-Methylacetamide      |
| 7. 1-Butoxybutane               | 24. Dichlorobenzene      | 41. N-Methylformamide      |
| 8. 1-Cyclohexylpyrrolidin-2-one | 25. Dichloromethane      | 42. N-Vinylpyrrolidone     |
| 9. 1,2,3,5-Tetrafluorobenzene   | 26. Dimethyl Sulfoxide   | 43. Phenoxybenzene         |
| 10. 2-Aminoethanol              | 27. Dimethylacetamide    | 44. Propane-1,2,3-Triol    |
| 11. 2,6-Dimethylheptan-4-One    | 28. Dimethylformamide    | 45. Propanenitrile         |
| 12. 2-Methylpropane             | 29. Ethane-1,2-Diamine   | 46. Pyrimidine             |
| 13. 2-Methylpropan-2-Amine      | 30. Ethanol              | 47. Quinoline              |
| 14. 3-Methylphenol              | 31. Formaldehyde         | 48. Tetrahydrofuran        |
| 15. 4-Methylpyridine            | 32. Gamma-Butyrolactone  | 49. Toluene                |
| 16. Acetamide                   | 33. Isopropyl Alcohol    | 50. Triethyl Phosphate     |
| 17. Acetone                     | 34. Methanesulfonic Acid |                            |