Refining Molecular Characterization to allow machine learning of the effectiveness of corrosion inhibitors

Ivan Cole,[1] Paul White,[1] José María Castillo,[1,2] Mazhar Iqbal,[1,2] Ernane Martins,[2] Pablo Ordejón,[2] Deng Qiushi,[1]
[1]RMIT University, Melbourne Australia,
[2]ICN2 (CSIC-BIST), Barcelona, Spain
Ivan.cole@rmit.edu.au

There is a strong need to develop new and non-toxic corrosion inhibitors to protect industrial metals (steel, galvanized steel, zinc, copper) from corrosion. Such inhibitors would be embedded in the primer of a paint system or used in initial metal passivation. Traditionally, inhibitors have been based on chromate or other toxic compounds that are being banned by legislation around the world. Small hetero-cyclic compounds are a promising alternative, but, as the exact molecular structure is critical to determining the inhibition efficiency, there are literally tens of thousands of possibilities. Various methods, including high throughput experimentation and computational modelling, have been developed to select or design the optimum molecular structure. A very promising approach is the use of inverse design. In this approach, high throughput experiments defining electrochemical performance and computation methods defining inhibitor characteristics and attributes are linked by a machine learning algorithm to define the molecular attributes most critical for inhibition performance. These critical molecular attributes are then used to search molecular databases and select promising candidate inhibitors that are then subject to testing for verification. Prof. Cole's teams have been developing this approach for around 15 years. Early work was able to use quantitative structural activity relationships (QSAR) methods based on a neural network approach to obtain reasonable models of the features controlling inhibition [1]. However, while these models could represent existing data, they were not very effective in predicting performance of new molecules. It was considered that this may have been due to either the relevance of molecular characterization and attribute definitions or the size and coverage of both the computational and experimental database. A large program was undertaken between where the databases were significantly enhanced and great care was undertaken to ensure that both experimental and computational data were accurate and reproducible. Further, the molecular attributes were refined to better represent molecular interactions with solvent and metal surfaces. A range of statistical and QSAR techniques were again used to define the relationships between the molecular attributes and electrochemical performance. These models demonstrated an enhanced ability to represent existing data, but their predictive ability could still be improved. Within this program, substantial experimental and modelling work was undertaken to reach a deeper understanding of inhibitor formation and stability. In some cases, inhibitor films reach a peak performance after a day in inhibited saline solution, with voids appearing in the coating that expanded with time. Molecular dynamics models indicated that the inhibitor film may be subject to electroporation, where charge at the metal surface causes the inhibitors to clump together allowing water to again reach the metal surface. To gain a deeper understanding of inhibitor bonding and inhibitor layer formation, an extensive modelling program was initiated between RMIT and ICN2. Recognizing the limitation of previous models (inadequate or no representation of solvent, no potential effects, and relatively small models) a new methodology was developed based on a combined quantum mechanics/molecular mechanics/non-equilibrium Green's functions or QM/MM/NEGF approach [2,3,4,5], implemented in SIESTA, a DFT method for ab-initio electronic structure calculations [6]. In this approach, larger models that include both solvent and voltage effects can be constructed, as depicted in Figure 1.

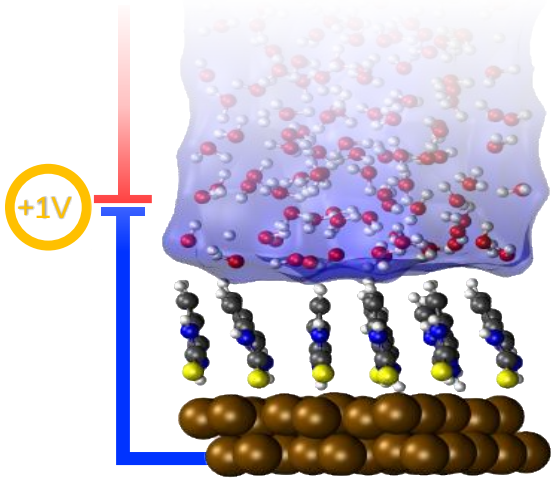

Figure 1. Sketch for the complete model where copper atoms are covered by a full self-assembled monolayer of corrosion inhibitors, including both the solvent and voltage effects. Ochre, yellow, gray, red, and white spheres represent Cu, S, C, O, and H atoms respectively.

The system has been applied to the inhibition of both copper and zinc surfaces by 2-mercapto-benzimidazole (MBI). The first principle of the protection provided by the MBI molecule is forming a physical barrier, protecting copper from interacting with water molecules, as shown in Figure 2. As one can see by comparing the red areas of the plots, the non-protected surface is exposed to the water molecules, and the case with a low coverage (1 MBI molecule) offers some protection as the water peak decreases. Full coverage, i.e., the formation of a packed self-assembled monolayer does not allow water to penetrate the inhibitor film and interact with copper. Another important result of the study is that, when MBI binds to the surface, there is a major electronic re-alignment across the molecular film which itself forms a dipole, as one can see in Figure 3. The traditional theory of inhibitors is that they form a barrier to both water and solutes and charge transfer. These studies indicate that, while MBI can

be an effective barrier against water, it cannot be regarded as a charge barrier and, in fact, the way charges re-align and the formation of the dipole will have a profound influence on the deposition of the second inhibitor layer.

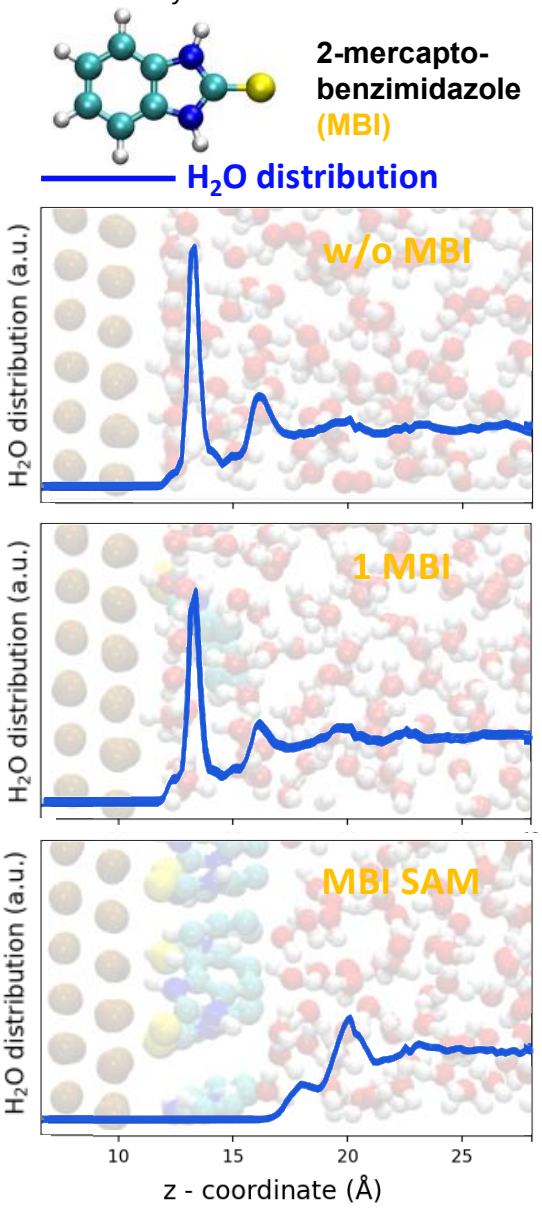

Figure 2. Atomic structure of MBI molecule, a corrosion inhibitor for copper, and the water distribution as a function of the z coordinate for varying MBI coverages. The curves are averaged over all molecular dynamics steps and the background transparent atomic structures are provided to help localize the positions of the atoms.

Current work now looks both at the second inhibitor layer formation and building more complex surfaces into the model. The work also has implications to the definition of molecular attributes for machine learning (ML) approaches. It confirms the importance of inhibitor surface bonding and again shows how a range of attributes are required to ensure this is correctly represented in ML. Furthermore, it highlights that monolayer formation and second layer inhibitor formation are also critical to understand inhibition and inhibitor layer stability. The molecular attributes that can reflect these processes are quite different from those that reflect surface bonding and, thus, our data sets used in ML approaches need significant redesign. This may be one reason why existing models have shown to have limitations in predicting new inhibitors.

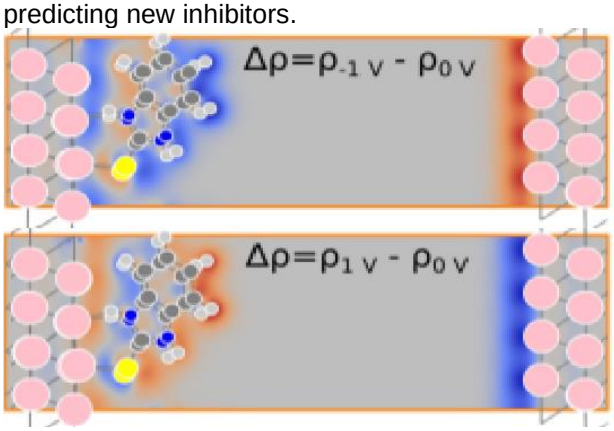

Figure 3. Electronic density redistribution due to the voltage applied in the system composed of two copper electrodes and a self-assembled monolayer of MBI.

Acknowledgments

ICN2 was funded by Grant PID2022-139776NBC62 from Spanish MCIN/AEI/10.13039/501100011033 and ERDF "A way of making Europe", Grant 2021SGR1519 from AGAUR, the CERCA programme (Generalitat de Catalunya), and the Severo Ochoa Centres of Excellence programme, MCIN/AEI/10.13039.501100011033/, Grant CEX2021-001214-S.

References

[1] Chen FF, Breedon M, White P, Chu C, Mallick D, Thomas S, Sapper E, Cole I, Materials & Design, 112 (2016), 410-418.

[2] Ordejón, P., Bulletin of the APS, https://meetings.aps.org/Meeting/MAR20/Session/F45.1

[3] Ordejón, P., de Freitas Martins, E., Febrer, P., Pedron, F., Bulletin of the APS, https://meetings.aps.org/Meeting/MAR23/Session/S60.2

[4] Sanz-Navarro et al., Theor. Chem. Acc. 128 (2011), 825.

[5] Papior et al., Comput. Phys. Commun. 212 (2017) , 8.

[6] Garcia, A. et al., The Journal of Chemical Physics, 152 (2020), 204108.