# Predicting Metalloprotein Redox Potentials with Machine Learning: A Focus on Iron–Sulfur Systems

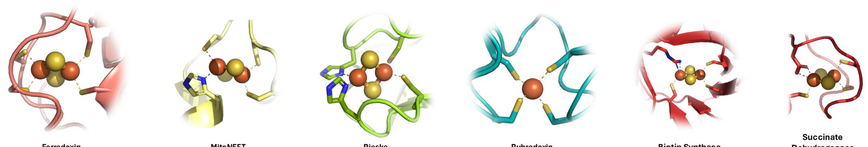

**F. Persico,[1] B. G. Galuzzi,[2] M. Pellegrino,[3] A. Claudel,[4] L. De Gioia,[1] F. Nastri,[3] G. Gilardi,[4] C. Damiani,[1] F. Valetti,[4] M. Chino,[3] and F. Arrigoni[1]***

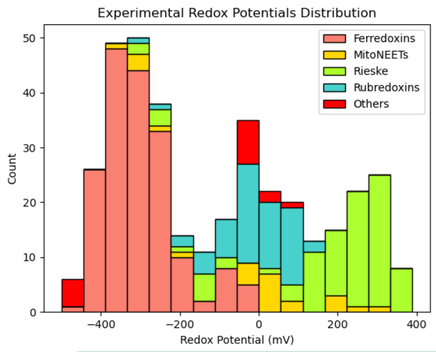
1. Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milano 20126, Italy
2. Institute of Bioimaging and Complex Biological Systems, National Research Council, Segrate 20054, Italy; National Biodiversity Future Center, Palermo 90133, Italy
3. Department of Chemical Sciences, University of Naples "Federico II", Napoli 80126, Italy
4. Department of Life Sciences and Systems Biology, University of Torino, Torino 10124, Italy


## Introduction

Iron-Sulfur (Fe-S) clusters play a critical role as highly efficient electron transfers in many energy-related metalloenzymes such as hydrogenases, CO dehydrogenases, formate dehydrogenases and nitrogenases. Fe–S proteins occupy most of the biological RP spectrum while exhibiting remarkable redox versatility. Understanding Fe-S clusters is, therefore, essential for developing new sustainable strategies for energy storage and transformation. A key property of these clusters is their reduction potential, which significantly influences enzyme reactivity. However, calculating reduction potential poses significant challenges. Current in silico methods are computationally expensive and often yield moderate accuracy.[1] Machine Learning offers a promising approach to correlate a protein's 3D structure with its properties, including the reduction potential of its clusters.

## Training Dataset with [1Fe-0S] and [2Fe-2S] proteins



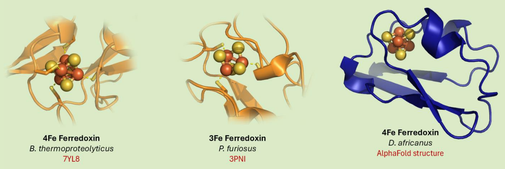



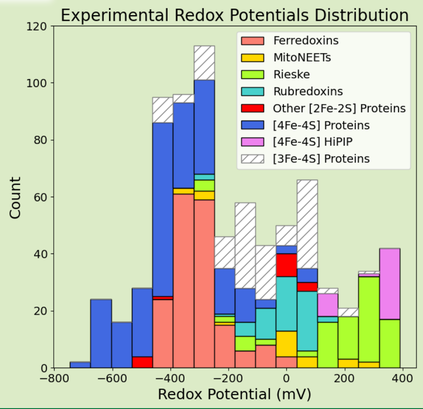

The dataset includes proteins with available 3D structures and known reduction potential. We included 56 different proteins and their mutants in the dataset, resulting in a total of 349 dataset entries. If the PDB structure of the mutants was unavailable, we generated it using the residue and loop mutation tool in Maestro, minimizing the structure within a 5Å radius around the mutation. The proteins in the dataset cover a broad range of RP, from -460 mV to +390 mV.[2]

### Ongoing work

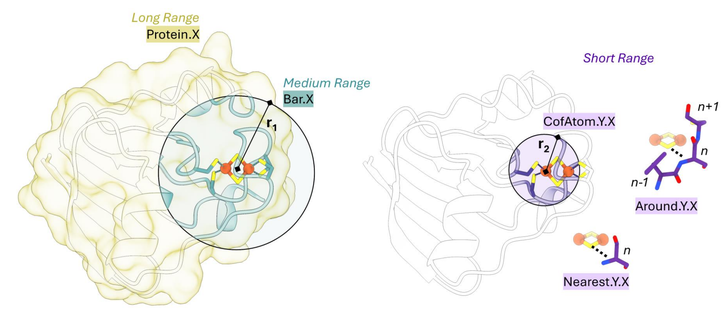
The dataset is being extended with:

- [4Fe-4S] proteins,
- [3Fe-4S] proteins,
- AlphaFold generated structures




## Molecular Descriptors



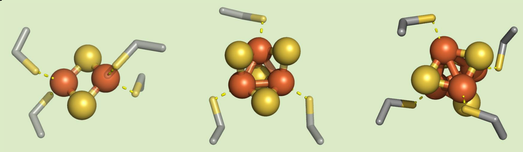

We designed a set of 66 molecular descriptors. These features are automatically extracted from each protein's PDB structure and are recalculated across three spacial ranges, defined as follows:

1. **Long range** - the whole protein;
2. **Medium range** - the portion of the protein within a sphere of radius $r_1$ centered at the cluster's barycenter;
3. **Short range:**
   i. the portion of the protein within a sphere of radius $r_2$ centered on each atom of the cofactor;
   ii. property of the amino acids located in the proximity of each cofactor's atom;

The values of the descriptors depend on the chosen radii values, $r_1 = 8, 9, \cdots, 16$ and $r_2 = 3, 4, 5$. We trained a separate ML model for each combination of radii, as well as a combined model that simultaneously considers all radius combinations. In addition to these spatial features, we also considered the pH at which the redox potential was measured.

### Ongoing work

We are now adapting the molecular descriptors to better represent all the different cofactor in our extended dataset.




## ML Models

Based on previous results,[3] we chose to train Extreme Gradient Boosting (XGB) models. To optimize the computational costs and model's performance, we performed different types of training:

- We trained an ML model for each combination of $r_1$ and $r_2$ using all the descriptors mentioned above → **A**-XGB
- We trained an ML model that simultaneously considers all features across all combi-nations of radii → **B**-XGB

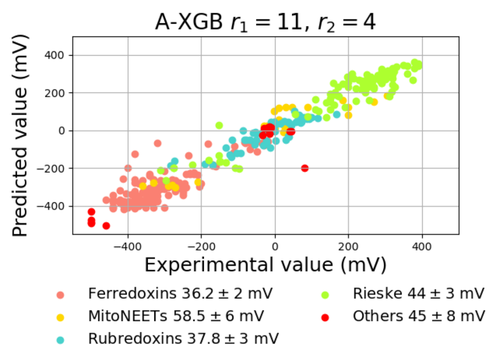
### Ongoing work

Exploring the use of a Tabular Foundation Model (TabPFN), developed specifically for small datasets.[4]

## ML Models performance

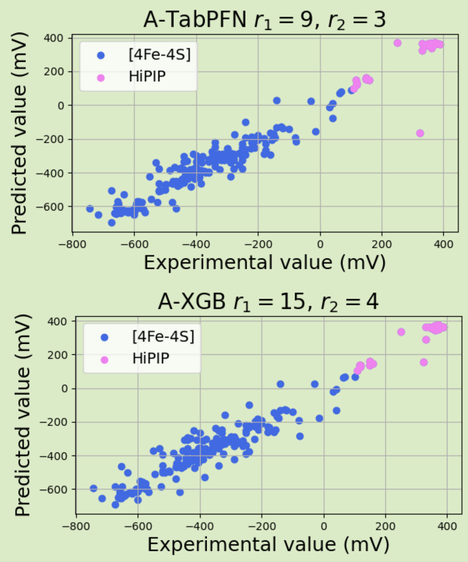
The following table reports the performance of the different models, evaluated on an unseen subset of data (20% of the dataset) during training, calculated over 10 repetitions of the training. For each type of training, we report the performance of the best model.

| Model | MAE (mV) | RMSE (mV) | $R^2$ | SC |
|---|---|---|---|---|
| **A**-XGB $r_1$=11, $r_2$=4 | 39.9 ± 4 | 57.7 ± 8 | 0.94 ± 0.02 | 0.97 ± 0.01 |




### Ongoing work
### Results on [4Fe-4S] proteins

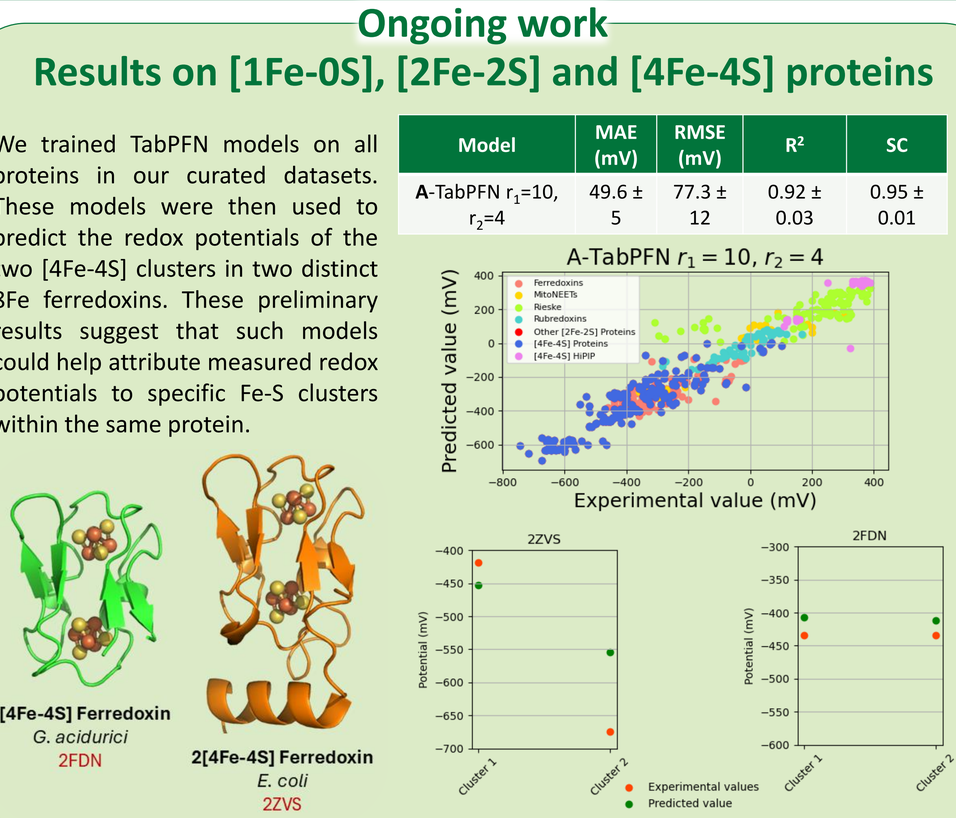
| Model | MAE (mV) | RMSE (mV) | $R^2$ | SC |
|---|---|---|---|---|
| **A**-TabPFN $r_1$=9, $r_2$=3 | 37.6 ± 6 | 60.2 ± 15 | 0.95 ± 0.03 | 0.94 ± 0.02 |
| **A**-XGB $r_1$=15, $r_2$=4 | 39.5 ± 7 | 62.6 ± 10 | 0.95 ± 0.02 | 0.93 ± 0.03 |




### Ongoing work
### Results on [1Fe-0S], [2Fe-2S] and [4Fe-4S] proteins

We trained TabPFN models on all proteins in our curated datasets. These models were then used to predict the redox potentials of the two [4Fe-4S] clusters in two distinct 8Fe ferredoxins. These preliminary results suggest that such models could help attribute measured redox potentials to specific Fe-S clusters within the same protein.

| Model | MAE (mV) | RMSE (mV) | $R^2$ | SC |
|---|---|---|---|---|
| **A**-TabPFN $r_1$=10, $r_2$=4 | 49.6 ± 5 | 77.3 ± 12 | 0.92 ± 0.03 | 0.95 ± 0.01 |

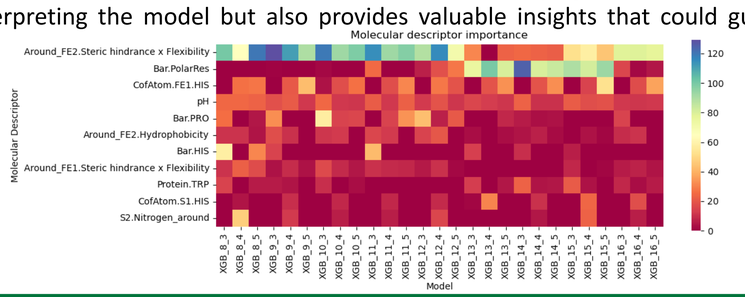






## Molecular descriptors analysis

To gain a deeper understanding of how protein structure influences RP, we assessed the contribution of individual molecular descriptors to the model's predictions. Identifying the most relevant features is not only essential for interpreting the model but also provides valuable insights that could guide protein design and engineering.




## Conclusions

Our ML framework achieved competitive predictive performance, demonstrating that the structure-based molecular descriptors effectively capture the key features of the protein environment around the cofactor. Benchmarked against state-of-the-art computational methods, our method achieves an average error of ~40 mV across cluster classes — without scaling factors or expensive quantum calculations. Training is computationally lightweight, and predictions on new proteins are obtained within seconds, making the approach ideal for large-scale screening.

The model delivers interpretable outputs and demonstrates stable, low-variance performance, suggesting it captures genuine structure–function relationships. Current applications cover mononuclear and [2Fe–2S] clusters; extension to higher-nuclearity systems ([3Fe–4S], [4Fe-4S]) is underway. Future developments include integration of AI-predicted structures (AlphaFold, RoseTTAFold), protein dynamics and conformational ensembles, and expansion to other metalloprotein families.

## References


[1] Jafari et al. *Inorg. Chem.* **2022**, 61, 5991-6007
[2] Persico et al. *Journal of Chemical Information and Modeling* **2025** 65 (21), 11631-11643
[3] Galuzzi et al. *Journal of Chemical Information and Modeling* **2022** *62* (19), 4748-4759
[4] Hollmann et al. *Nature* **2025**, 637, 319-326



Funded by the European Union NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA