I'm gonna need like 10 Irish Car Bombs, or margaritas when this is all over
UAB

Pat the Great and Powerful

29 June 2021

Abstract

Metalloproteins compose approximately 40 percent (look up how to do percents in latex) of all known proteins, and use some metallic group to accomplish their chemistry. One such metallic group is heme. Heme is a member of the porphyrin family, which are able to catalyze a broad range of reactions. Heme in particular catalyzes many different reactions and is present in many proteins. However, the underlying structural requirements to host heme in a protein are not well studied.

In this study, all heme or heme-c containing proteins as of xx were downloaded and processed in order to determine underlying structural characteristics these proteins may have in common. Parameters that were examined include: xx. Overall, we found: xx. These results may have implications for protein engineering; or if I fucked up this illustrates the difficulty of the field and demonstrate the wide range of acceptable environments of heme; it may therefore be more appropriate to take a more hands-on approach until perhaps other computational methods evolve to better examine structure-function relationships.

See? Not so bad of a worst-case scenario. Just, an unusual sentiment to see in modern science.

Lay Summary

Proteins are biological molecules that perform essential functions in our bodies and in all living things. They are responsible for everything from transporting oxygen in our blood to photosynthesis in plants. They can be extracted from living things and cultured in laboratories and factories. They can then be used as drugs, or be used to perform functions in industrial processes.

Many proteins require additional molecules to perform their function, and these molecules must be bound, docked to the molecule like a boat in a harbor. The molecules can be bound to the protein inside a specialized space, or pocket, within or on the surface of the protein.

One class of proteins is hemoproteins - proteins that use the molecule heme to perform their function. The heme is critical to the proper function of these proteins. Heme enables many specific chemical reactions to be performed, or assisted by the protein. In hemoglobin, in blood, the heme molecule allows the overall protein of hemoglobin to carry and transport oxygen around the body.

But the specifics of the pocket that binds heme in these hemoproteins is not well understood. What are the conditions inside the pocket? Is it a very specific size or can it vary? Is there anything in common among the pockets of many different hemoproteins? These are the questions this research hoped to answer; or at the very least, provide some data and lay the groundwork for others to continue the research later.

This investigation was not conducted in a lab. Rather, using various software packages and the 3D structures of hemoproteins published in a database, various properties of these hemoproteins were calculated. The data produced were analyzed with various statistical methods to extract potentially useful information.

Overall, the following was found: Grad school sucks but bravas are delicious.

Acknowledgments

In case anyone reads this in the future, some context may be appreciated: I attended and completed this Master's during the COVID-19 global pandemic from September 2020 to September 2021.

Thanks professors

Thanks lab

Thanks UAB

Thanks Spain, and Catalonia, allowing me in and then also having public health measures unlike Donny's America

Thanks classmates

Thanks fam, friends

Thanks to the media and the creators of media that facilitated the survival of my sanity through the pandemic.

Finally, I'd like to quote a well-known artist from California. He was referencing his own work, but I wholly identify with his appreciation for the subject of his esteem:

"Last but not least, I wanna thank me. I wanna thank me for believing in me. I wanna thank me for doing all this hard work. I wanna thank me for having no days off. I wanna thank me for, for never quitting. I wanna thank me for always being a giver, and trying to give more than I receive. I wanna thank me for trying to do more right than wrong. I wanna thank me for just being me at all times."

– Calvin Cordozar Broadus Jr.

Contents

List of Figures

List of Tables

Introduction

Proteins may catalyze reactions, and many require ligands to enable their chemistry. A significant portion of all proteins, approximately 40%, require a metallic group as a ligand in order to function correctly - these proteins are known as metalloproteins.

One of these metallic groups is heme. Heme is a member of the porphyrin family, a group of molecules capable of catalyzing a broad range of reactions. Heme can catalyze many different reactions and is present in many proteins. However, the underlying structural requirements to host heme in a protein are not well understood. [MAY ADD CITATIONS]

There have only been a handful of studies dedicated to understanding the structure-chemical relationship between heme and the proteins that use heme for their chemistry (these proteins are known as hemoproteins).

In the most significant previous work, approximately 125 hemoproteins were studied[2]. Although pdbs were thoroughly examined and the datasets were culled, the sample size of this study is very small compared to the amount of hemoproteins available in the pdb a decade later (10,000 HEM-containing proteins and xx). The dataset is also limited in that there is a somewhat homogenous group of proteins examined (?). The characteristics examined were limited to: xx.

It is hypothesized that the following characteristics all have an impact on the binding of heme and function of the hemoprotein: XXXXXXXXX.

In this study, some of these characteristics were examined. They include: XX. The remainder are thus far not feasible to calculate.

All of these characteristics have implications in the field of protein engineering or basic research into hemoproteins. Examples of the uses of these results include [SUPER BLOOD STUDY] and [OTHER PROTEIN ENGINEERING STUFF]. Not sure how much we can reference those other papers besides doing that besides in the conclusion.

Notable results from some of the prior studies include: xx and xx. These characteristics are also examined in this dataset, while some are not due to different study approaches.

Methods

Perhaps I need something right here first?

- Download from PDB using the script they've provided at RCSB for many, many files
- Use UCSF Chimera to determine:
  - Volume
  - SA
  - Nearby AA
- R to process raw data and produce tables
- Whatever other software we use to achieve the other results. E.g. E, or availability to solvent etc. likely will stick w Chimera I suspect. Or somehow implement the Python script to open both chimera for the first part of what we've done or for something else later. The script we've written is a python script, not a chimera script. We're initializing it with chimera and excluding the necessary code... to initialize chimera and specify chimera to receive the commands

Equipment

Results

Discussion

Limitations of this brilliant work: Limited sample size Limited experimental data to reference to verify NO experimental data in this study to verify, all theoretical Only one software package/few algorithms used to calculate all these properties. Others were evaluated but none are compared w. Algorithms may introduce bias based on how they work e.g. all the bubbles Arbitrary selection of parameters; some based on rule of thumb or visual evaluation but all or almost all arbitrary Unknown if the qualities measured are truly the most critical for the heme binding. Some papers suggest other properties may also be important but cannot be calculated, at least right now Visual examination itself to OK the parameters/algorithms can introduce bias Precision of algorithms needs to be evaluated** or at least IDK how PRECISE they are

Conclusion

this is just as master's and in basic research don't feel the need to replicate what took some god forsaken, sad, overworked, impoverished PhD students years + with help of their PIs and with generous word fluff to hide fuck ups. 0 -¿ thesis in approx. 3-4 months during global catastrophe is nifty

Bibliography

[1] Chris E Cooper et al. "Engineering tyrosine residues into hemoglobin enhances heme reduction, decreases oxidative stress and increases vascular retention of a hemoglobin based blood substitute". In: *Free Radical Biology and Medicine* 134 (June 2019), pp. 106–118. ISSN: 18734596. DOI: 10.1016/j.freeradbiomed.2018.12.030.

[2] Ting Li, Herbert L Bonkovsky, and Jun Tao Guo. "Structural analysis of heme proteins: Implications for design and prediction". In: *BMC Structural Biology* 11 (2011). ISSN: 14726807. DOI: 10.1186/1472-6807-11-13.