Bio-computational evaluation of *Cryptosporidium* inosine monophosphate dehydrogenase as a viable target in the attenuation of cryptosporidiosis.

Kehinde Foluke Omolabi

217077656

A thesis submitted to the School of Health Sciences, College of Health Science, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

Supervisor

Prof. Mahmoud Soliman

2021

Bio-computational evaluation of *Cryptosporidium* inosine monophosphate dehydrogenase as a viable target in the attenuation of cryptosporidiosis.

Kehinde Foluke Omolabi

217077656

2021

A thesis submitted to the School of Health Sciences, College of Health Science, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

This is the thesis in which the chapters are written as a set of discrete research publications with a general introduction and final summary. Typically, these chapters will have been published in internationally recognized, peer-reviewed journals.

This is to certify that the content of this thesis is the original research work of Ms. Kehinde Foluke Omolabi, carried out under my supervision at the Molecular Bio-computation and Drug Design Unit, Westville Campus, University of KwaZulu-Natal, Durban, South Africa.

Supervisor

Signed: Name: Professor Mahmoud Soliman Date: 25/04/2021

PREFACE

This thesis is grouped into eight chapters, with the present chapter included:

Chapter 1

This chapter provided the study's background, rationale, aim and objectives, significance, and the novelty of the findings reported.

Chapter 2

This chapter gave an overview of the research and expounded the background information. It discussed the life-cycle and transmission of cryptosporidiosis, target-based drug design for cryptosporidiosis, and inosine monophosphate dehydrogenase as a viable drug target in *Cryptosporidium parvum*.

Chapter 3

This chapter elaborated on the specific computational tools, methods, and analyses that were utilized in achieving the aim and objectives of this study. It explained how the methodologies were integrated in inferring the conclusions.

Chapter 4: (Published work- this chapter is presented in the required format of the journal and is the final version of the published manuscript)

This chapter employed meta-analysis to report the incidence and prevalence of cryptosporidiosis in humans in the southern Africa region over a twenty-year span (2000-2020). The article has been published in Journal of Parasitic Diseases (IF= 1.21)

Chapter 5: (Published work- this chapter is presented in the required format of the journal and is the final version of the published manuscript)

This section is entitled "Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity – A Structural, Dynamical and Mechanistic Stance". The chapter unraveled the mechanism of inhibition and selectivity of P131- a promising anticryptosporidial. We combined both bioinformatics and computational tool to elucidate the structural mechanism of inhibition and selection by P131. Sequence alignment, total binding free energy, conformational perturbations

were examples of the analysis used. The article has been published in Cell Biochemistry and Biophysics (IF=2.03).

Chapter 6: (Published work- this chapter is presented in the required format of the journal and is the final version of the published manuscript)

This chapter is captioned, "A Probable Means to an End: Exploring P131 Pharmacophoric Scaffold to Identify Potential Inhibitors of *Cryptosporidium parvum* Inosine Monophosphate Dehydrogenase". The study selected the pharmacophoric moieties of P131 that were responsible for its inhibitory potential. A pharmacophore scaffold of P131 was created and used as a screen in ZINC database to mine for other ligands that share similar moieties. The ligands that were identified were docked in the co-factor binding site of *Cp*IMPDH, and the top-three compounds that had the highest molecular docking scores were selected. The complexes were subjected to molecular dynamics simulation and thereafter, conformational and pharmacokinetic analyses were done. The article has been published in the Journal of Molecular Modeling (IF=1.34)

Chapter 7: (Published work -this chapter is presented in the required format of the journal and is the final version of the submitted manuscript).

This chapter is entitled "Finding the Needle in the Haystack"- Will Natural Products Fit for Purpose in the Treatment of Cryptosporidiosis? - A Theoretical Perspective. We employed structure-based virtual screening of 107,000 natural compounds deposited in ZINC database. This was in a bid to find an alternative therapy to cryptosporidiosis. The top three compounds that best fit in the co-factor binding site of *Cp*IMPDH were selected. The complexes were simulated on an extended scale. Their respective total binding free energy was estimated, their drug-likeness was also evaluated. The manuscript has been published in Molecular Simulation (IF=1.71)

Chapter 8

This is the final chapter that has the concluding remarks and proposes future work

ABSTRACT

Cryptosporidiosis caused by *Cryptosporidium* species is an enteric disease infecting vertebrate. The disease is associated with poor living conditions, such as contamination of recreational water facilities with infective oocysts. Transmission can be zoonotic or anthroponotic through direct contact or faeco-oral routes. Cryptosporidiosis is a significant opportunistic parasitic organism of immune-compromised patients. Globally, The World Health Organisation (WHO) reported three million infections annually, mostly in developing countries. Notably, cryptosporidiosis self-resolve in immune-competent individuals after few days of watery diarrhea. However, infection is often treated by nitazoxanide, a thiazolide class of drug. Previous studies have reported nitazoxanide's ineffectiveness in immune-compromised patients (such as pregnant women, children, and HIV/AIDS individuals). Hence, there is a need for continuous research on effective drugs against cryptosporidiosis in individuals with a challenged immune system. This study investigates the pooled prevalence of *Cryptosporidium* infection in the southern Africa region and also the molecular mechanism of action of potential anticryptosporidials.

Meta-analysis was conducted by screening literature database (Google Scholar, PubMed, Ovid Medline, AJOL, and Web of Science) between 2000 and 2020 to estimate the recent pooled prevalence of *Cryptosporidium* infection in southern Africa. This thesis investigated the inhibitory dynamics of the promising anticryptosporidial drug P131 on *Cryptosporidium parvum* inosine monophosphate dehydrogenase (*Cp*IMPDH) compared to the orthologous mouse protein (*m*IMPDH). Crucial moieties of P131 were identified and subsequently adopted to create a pharmacophore model for virtual screening in the ZINC database through the per residue energy decomposition approach. This was done to mine for compounds that could be as effective as or more effective than P131. The potential inhibitory mechanism of these compounds was probed using molecular dynamics simulation and Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA) analyses. In addition, a dedicated library of 107,000 natural compounds available in the ZINC database was virtually screened against *Cp*IMPDH –NAD^+ binding site to determine the compounds that have the best complementarity to the binding site.

Cryptosporidium infection in southern Africa shows the pooled prevalence of 16.8% (95%CI 9.7-25.3), with subgroup analyses revealing the highest pooled prevalence of 25.2% in HIV/AIDS

patients. The high prevalence of *Cryptosporidium* spp. infections among immune-compromised patients in southern Africa showed that the pathogen is of significant importance in this region. The relatively high-affinity interactions occurring at the *Cp*IMPDH-NAD^+ site were majorly mediated by SER22, VAL24, PRO26, SER354, GLY357 and, TYR358 located on chain D of *C*pIMPDH. These residues are unique to the parasite IMPDH and not in their eukaryotic host, thereby explaining the selective action of P131. Three compounds ZINC46542062, ZINC58646829, and ZINC89780094, which contained the pharmacophore of P131, showed a respective, favorable docking score of -8.3kcal/mol, -8.2 kcal/mol, and -7.5kcal/mol in *Cp*IMPDH-NAD^+ site. Results revealed that one of the hits (ZINC46542062) exhibited a higher binding free energy of -39.52kcal/mol than P131, which had -34.6 kcal/mol. Conformational perturbation induced by the binding of the identified hits to *Cp*IMPDH was similar to P131, suggesting a similarity in inhibitory mechanisms. The top three natural compounds identified with the best complementarity to the *Cp*IMPDH–NAD^+ binding site included ZINC5225833, ZINC4258873, and ZINC3841381. The latter (ZINC3841381) had the best binding free energy of -58.43kcal/mol.

The high prevalence of *Cryptosporidium* infections among immune-compromised patients in southern Africa revealed that cryptosporidiosis is of significant importance in this region. The molecular dynamics approaches presented revealed positive prospects toward the identification of novel anticryptosporidials. Identified ZINC compounds from both inorganic and natural sources could serve as the basis for further experimental investigations, optimization for improved selectivity, thereby providing several therapeutic options of treatment.

DECLARATIONS

DECLARATION 1 - PLAGIARISM

I, **Kehinde Foluke Omolabi** declare that

1. The research report in this thesis, except where otherwise indicated, is my original work.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other person's data, pictures, graphs, or other information, unless expressly acknowledged as being sourced from other persons.
4. This thesis does not contain other person's writing, unless expressly acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. Their words have been re-written, but the general information attributed to them has been referenced.
    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet unless expressly acknowledged, and the source being detailed in the thesis and the references sections.

A detailed contribution to publications that form part or/and include research presented in this thesis is stated (include publications submitted, accepted, in the press, and published).

Signed

Kehinde Foluke Omolabi

DECLARATION II – LIST OF PUBLICATIONS

The written manuscripts in chapters 4, 5, 6 and 7 are presented in the required format of the journal and are the final version of the accepted manuscript and are listed below:

1. **Kehinde F. Omolabi**, Paul O. Odeniran and Mahmoud E. S. Soliman (2021). A meta-analysis of *Cryptosporidium* species in humans from southern Africa (2000 - 2020). *Journal of Parasitic Diseases*. doi: 10.1007/s12639-021-01436-4 (Published)

Contribution:

Kehinde F. Omolabi: Extracted, sorted, and analyzed the data. Manuscript preparation and writing.

Paul O. Odeniran: Extraction of data and manuscript writing.

Mahmoud E. S. Soliman: Supervisor.

Appendix A: Pdf version of the publication

2. **Kehinde F. Omolabi**, Clement Agoni, Fisayo A. Olotu, and Mahmoud E. S. Soliman (2020), Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity – A Structural, Dynamical and Mechanistic Stance. *Cell Biochemistry and Biophysics.* doi: 10.1007/s12013-020-00950-1. (Published)

Contribution:

Kehinde F. Omolabi: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Clement Agoni: manuscript editing.

Fisayo A. Olotu: manuscript editing.

Mahmoud E. S. Soliman: Supervisor.

Appendix B: Pdf version of the publication

3. **Kehinde F. Omolabi**, Emmanuel A. Iwuchukwu, Clement Agoni, Fisayo A. Olotu and Mahmoud E. S. Soliman (2020). A Probable Means to an End: Exploring P131 Pharmacophoric Scaffold to Identify Potential Inhibitors of *Cryptosporidium parvum* Inosine Monophosphate Dehydrogenase. *Journal of Molecular Modeling*. doi: 10.1007/s00894-020-04663-3. (Published)

Contribution:

Kehinde F. Omolabi: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Iwuchukwu A. Emmanuel: contributed to the experimental work

Clement Agoni: manuscript editing.

Fisayo A. Olotu: manuscript editing.

Mahmoud E. S. Soliman: Supervisor.

Appendix C: Pdf version of the publication

4. **Kehinde F. Omolabi**, Clement Agoni, Fisayo A. Olotu and Mahmoud E. S. Soliman (2021), entitled "Finding the Needle in the Haystack"- Will Natural Products Fit for Purpose in the Treatment of Cryptosporidiosis? - A Theoretical Perspective. *Molecular Simulation* doi: doi.org/10.1080/08927022.2021.1895435. (Published)

Contribution:

Kehinde F. Omolabi: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Clement Agoni: manuscript editing.

Fisayo A. Olotu: manuscript editing.

Mahmoud E. S. Soliman: Supervisor.

Appendix D: Pdf version of the publication

RESEARCH OUTPUT

A- LIST OF PUBLICATIONS

1. Kehinde F. Omolabi, Paul O. Odeniran and Mahmoud E. S. Soliman (2021). A meta-analysis of *Cryptosporidium* species in humans from southern Africa (2000 - 2020). *Journal of Parasitic Diseases*. doi: 10.1007/s12639-021-01436-4

2. Kehinde F. Omolabi, Clement Agoni, Fisayo A. Olotu, and Mahmoud E. S. Soliman (2020), Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity – A Structural, Dynamical and Mechanistic Stance. *Cell Biochemistry and Biophysics.* doi: 10.1007/s12013-020-00950-1

3. Kehinde F. Omolabi, Emmanuel A. Iwuchukwu, Clement Agoni, Fisayo A. Olotu and Mahmoud E. S. Soliman (2020). A Probable Means to an End: Exploring P131 Pharmacophoric Scaffold to Identify Potential Inhibitors of *Cryptosporidium parvum* Inosine Monophosphate Dehydrogenase. *Journal of Molecular Modeling*. doi: 10.1007/s00894-020-04663-3

4. Kehinde F. Omolabi, Clement Agoni, Fisayo A. Olotu and Mahmoud E. S. Soliman (2021), entitled "Finding the Needle in the Haystack"- Will Natural Products Fit for Purpose in the Treatment of Cryptosporidiosis? - A Theoretical Perspective. *Molecular Simulation* doi: doi.org/10.1080/08927022.2021.1895435

ACKNOWLEDGEMENTS

There is none holy as the Lord, and there is none beside Him. All the glory and honor of this journey belong to Jesus Christ alone. Thank you, Lord, for this that you have done again. It was You all the way.

My profound gratitude goes to

- My supervisor: Professor Mahmoud Soliman for his invaluable inputs, initiatives and guidance. I have learned hard work, tenacity, and humility from him.
- University of KwaZulu-Natal (College of Health Sciences), for financial support.
- The Molecular Bio-computational and Drug Design group.
- My parents Deacon and Mrs. J.A Omolabi and Pastor and Mrs. A.O Odeniran for their long-suffering and for relentlessly praying all my siblings and me into a brighter future.
- My twin brother; Mr. Taiwo Omolabi, for constantly checking up on me every day
- Many thanks to Olaoluwa Omolabi, Mr. and Mrs. Razak, my ever-supportive siblings.
- My handsome nephews: Emmanuel, Ephraim, and Erastus.
- Bukola Omotosho- My great friend
- Members and pastorate of Deeper Life Campus Fellowship (Westville) for their spiritual support.
- Lesley Greeves and family, for always leaning in.

My most profound appreciation goes to my husband: Dr. Paul Odeniran, for being again so understanding, patient, and highly resourceful. Thank you yet again for not stifling my potentials; you instead fanned it to flames. It was expensive for you, but you paid the price nevertheless. I will not forget it.

DEDICATION

To Dr. Paul Odeniran – A husbandman indeed!

What exactly can I say?

Thank you.

To my parents- Deacon & Mrs. J.A Omolabi, this is for you both. Thank you for laying a solid foundation for your children.

LIST OF ABREVIATIONS

3D	Three-Dimensional
Å	Amperes
ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicology
AIDS	Acquired Immuno-Deficiency Syndrome
AMP	Adenine Monophosphate
ANOVA	Analysis of Variance
ART	Antiretrovirals
BFE	Binding Free Energy
CDPK-1	Calcium-Dependent Protein Kinase 1
*Cp*IMPDH	*Cryptosporidium parvum* inosine monophosphate dehydrogenase
*Cp*LC-FACS	*Cryptosporidium parvum* Long Chain Fatty Acid Co-Enzyme A Synthetase
DCCM	Dynamic Cross Correlation Matrix
df	Degree of Freedom
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked Immunosorbent Assay
FAO	Food and Agricultural Organization
FDA	Food and Drug Administration
GAFF	General amber Force Field
GMP	Guanine Monophosphate
GPU	Graphical Processing Unit
GUI	Graphical User Interface
HIV	Human Immuno-Deficiency Virus
IFAT	Indirect Fluorescent Antibody Test
IMP	Inosine Monophosphate
IMPDH	Inosine Monophosphate Dehydrogenase
K	Kelvin

LE	Ligand Efficiency
LELP	Ligand efficiency-dependent lipophilicity
LFK	Luis Furuya Kanamori
LLE	Lipophilic Ligand Efficiency
MC	Monte Carlo
MD	Molecular Dynamics
*m*IMPDH	Mouse inosine monophosphate dehydrogenase
MM	Molecular Mechanics
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics/Poisson-Boltzmann Surface Area
NAD	Nicotinamide Adenine Dinucleotide
NMR	Nuclear Magnetic Resonance
ns	nanoseconds
PBVS	Pharmacophore-based virtual screening
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PMEMD	Particle-mesh Ewald Molecular Dynamics
PRED	Per Residue Decomposition
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
qGIS	Quantum Geographic Information System
QM	Quantum Mechanics
QSAR	Qualitative Structure-Activity Relationship
QSPR	Qualitative Structure-Property Relationship
RDD	Rational Drug design
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic acid

RoG	Radius of Gyration
RSCB	Research Collaboration for Structural Bioinformatics
SASA	Solvent accessible surface area
SDF	Standard Data Files
UCSF	University of California at San Francisco
VS	Virtual screening
WHO	World Health Organization
XMP	Xanthine Monophosphate

LIST OF AMINO ACIDS

Three Letter Code	Amino acid
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1

1.1 Background and Rationale

Cryptosporidiosis remains a significant parasitic disease, most especially in sub-Saharan Africa. It is one of the leading causes of mortality and morbidity in infants (< 5 years) [1–6]. Nitazoxanide is the only FDA-approved for treating cryptosporidiosis, though other drugs have also been adapted for its treatment, e.g., azithromycin, paromomycin, rifamycin, etc. [7–10]. These drugs can resolve cryptosporidiosis in immune-competent patients. However, in immune-compromised patients, e.g., HIV/AIDS patients and infants, the drugs are not very effective [11,12].

Target-based drug design, which is one of the approaches to rational drug design, has been explored to design and develop drugs that have the propensity to inhibit one or multiple targets in the disease-causing organism. These targets must be crucial to the survival of the organism. This approach is faster and economically efficient than the traditional method of drug design.

The limitations of existing anticryptosporidials have led to the synthesis of other compounds that can overcome the disadvantage of mainstream drugs. P131 emerged as the most promising ligand from this effort [13,14,23,15–22]. The compound specifically targets inosine monophosphate dehydrogenase in *Cryptosporidium parvum*. IMPDH is crucial to the differentiation and replication of the parasite as it controls the guanine and the adenine pool in the cell. The inhibition of *Cp*IMPDH will lead to the blockade of guanine and adenine nucleotide synthesis thereby, causing the death of the parasite [24].

It has also been established that *Cp*IMPDH, despite a eukaryote, has a prokaryotic IMPDH. This property has allowed for the selective targeting of the prokaryotic IMPDH in a eukaryotic host. P131, therefore, only inhibits the parasite IMPDH and leaves the host IMPDH unaffected [25]. However, the structural mechanism of inhibition and selection of P131 has not been investigated.

The evolution and progression of computational tools have made possible the elucidation of ligands' molecular mechanism of action. It has also facilitated the processes of design and optimization of drugs *in silico* before they are subjected to further test in the laboratory. One of the computational approaches known as pharmacophore-based virtual screening [26] was used to

build a scaffold containing the active moieties of P131 to obtain from the ZINC database other compounds that might be as active as or more active than P131. These compounds can serve as leads for potential anticryptosporidials. Through structure-based virtual screening, another computational tool, natural compounds were tested *in silico* for their potential to mitigate cryptosporidiosis.

1.2 Aims and Objectives

This study is aimed at investigating the significance and burden of cryptosporidiosis in the southern African region. The structural mechanism of inhibition and selectivity of anticryptosporidials targeting *Cp*IMPDH with P131 as the test ligand was investigated. We wish to determine why P131 will inhibit the prokaryotic CpIMPDH and leave the eukaryotic host IMPDH unaffected from a dynamical perspective. Secondly, to mine for compounds sharing similar pharmacophoric moieties with P131 that can have comparable or superior efficacy to P131 when bound to the co-factor site of *Cp*IMPDH. Lastly, in addition to the synthetic compounds discovered as potential anticryptosporidials, do natural compounds also have a place in mitigating cryptosporidiosis?

To achieve the aim above, the following objectives were set

1. The pooled prevalence of cryptosporidiosis and associated risk factors in southern African countries among immune-compromised patients. This will be achieved by:

    1.1 Meta-analysis of cryptosporidiosis among children in southern Africa

    1.2 Meta-analysis of cryptosporidiosis among pregnant women in southern Africa

    1.3 Meta-analysis of cryptosporidiosis among HIV/AIDS in southern Africa

2. Understand the observed selectivity in action of the compound and the structural inhibitory dynamics of P131 when bound to *Cp*IMPDH and mouse IMPDH (*m*IMPDH). This will be accomplished by:

    2.1 Comparing the amino acid sequence of *Cp*IMPDH (prokaryotic model) and *m*IMPDH-eukaryotic model to know the percentage similarity and disparity.

2.2 Subjecting the complexes (*Cp*IMPDH-P131 and *m*IMPDH-P131) to a long scale MD simulation (280ns) and subsequently analyzed to generate data that can be meaningfully extrapolated.

2.3 Estimating of the binding free energies of P131 in *Cp*IMPDH and *m*IMPDH using the MMPB/SA method

2.4 Evaluating of structural computation matrices to explain the motion, dynamics, mechanics, and structural variations in understanding the mechanism of selection of P131 in *Cp*IMPDH and *m*IMPDH.

3. As P131 has been shown to be more efficacious than paromomycin, we screened for other potential drug candidates that have a similar scaffold with P131 and, for this reason, may be more effective than P131

3.1 Pharmacophore modeling will be employed, which incorporate both structure-based and ligand-based virtual screening. The per-residue energy decomposition method will be employed in modeling the pharmacophore.

3.2 The pharmacophore of P131 will be used as a screen in the ZINC database to output ZINC compounds with similar pharmacophore

3.3 The ZINC compounds and P131 bound to the co-factor site of *Cp*IMPDH will be subjected to 200ns molecular dynamics simulation, and after that, the total binding free energy will be quantified and also the resulting structural perturbation patterns will be compared to P131

3.4 The pharmacokinetic parameters (ADMET) of the ZINC compounds will be evaluated and compared to P131 to determine the compound that has superior ADMET properties and is more drug-likely

4. To investigate, using computational tools, the potential of natural compounds in cryptosporidiosis treatment. This will be made possible by

4.1 Virtually screening a dedicated library of 107,000 natural compounds available in the ZINC database against the co-factor site of *Cp*IMPDH

4.2 Computing the molecular docking scores of the natural compounds

4.3 The top three compounds with the highest docking scores in complex with *Cp*IMPDH will be subjected to molecular dynamics simulation (300ns).

4.4 The total binding free energy, conformational dynamics, and druggability scores of the natural compound will be evaluated.

1.3 Novelty and Significance of the study

The current treatment of cryptosporidiosis has its challenges, most especially its inefficacy in immune-compromised patients. This has led to the synthesis of potential anticryptosporidials, which have the propensity of overcoming the drawbacks of mainstream drugs. The progression in the integration of bioinformatics, computer science, physics, mathematics, etc., has made the possible target-based drug design and the *in silico* investigation of the mechanism of action of promising drugs. P131, an alternative yet promising drug, binds to the co-factor site of *Cp*IMPDH and inhibits it while sparing the IMPDH of the eukaryotic host it infects. To the best of our knowledge, this study is the first to highlight from a structural and mechanistic perspective why anticryptosporidials are selective in their therapeutic action using P131 as a model. We also projected through the pharmacophore modeling of P131, other compounds that have greater potential than P131 in mitigating cryptosporidiosis infection. To the best of our knowledge, this study is the first to report the virtual screening of a library of natural compounds in the ZINC database that displayed good complementarity with the NAD^+ binding site of *Cp*IMPDH. Insight from this study could serve as a starting point in optimizing these new inhibitors for *Cp*IMPDH (from pharmacophore modeling and the virtual screening with natural compounds) with improved therapeutic properties.

1.4 References

[1] Bouzid M, Kintz E, Hunter PR. Risk factors for Cryptosporidium infection in low and middle income countries: A systematic review and meta-analysis. PLoS Negl Trop Dis. 2018;12.

[2] Fischer Walker CL, Aryee MJ, Boschi-Pinto C, et al. Estimating diarrhea mortality among

young children in low and middle income countries. PLoS One. 2012.

[3] Valentiner-Branth P, Steinsland H, Fischer TK, et al. Cohort study of Guinean children: Incidence, pathogenicity, conferred protection, and attributable risk for enteropathogens during the first 2 years of life. J Clin Microbiol. 2003;41:4238–4245.

[4] Naghavi M, Abajobir AA, Abbafati C, et al. Global, regional, and national age-sex specifc mortality for 264 causes of death, 1980-2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet [Internet]. 2017 [cited 2020 Oct 8];390:1151–1210. Available from: https://pubmed.ncbi.nlm.nih.gov/28919116/.

[5] Kotloff K, Nataro J, Blackwelder W. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. 2013.

[6] Sow S, Muhsen K, Nasrin D. The Burden of Cryptosporidium Diarrheal Disease among Children &lt. 2016.

[7] Lee S, Harwood M, Girouard D, et al. The therapeutic efficacy of azithromycin and nitazoxanide in the acute pig model of Cryptosporidium hominis. Yeruva L, editor. PLoS One [Internet]. 2017 [cited 2020 Apr 20];12:e0185906. Available from: https://dx.plos.org/10.1371/journal.pone.0185906.

[8] Checkley W, Jr. AW, Jaganath D. A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. 2015.

[9] Hewitt RG, Yiannoutsos CT, Higgs ES, et al. Paromomycin: no more effective than placebo for treatment of cryptosporidiosis in patients with advanced human immunodeficiency virus infection. AIDS Clinical Trial Group. Clin Infect Dis [Internet]. 2000 [cited 2019 Oct 7];31:1084–1092. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11049793.

[10] Hussien SMM, Abdella OH, Abu-Hashim AH, et al. Comparative study between the effect of nitazoxanide and paromomycine in treatment of cryptosporidiosis in hospitalized children. J Egypt Soc Parasitol [Internet]. 2013 [cited 2019 Oct 7];43:463–470. Available

from: http://www.ncbi.nlm.nih.gov/pubmed/24260825.

[11] Cabada MM, White AC. Treatment of cryptosporidiosis: do we know what we think we know? Curr Opin Infect Dis [Internet]. 2010 [cited 2019 Oct 7];23:494–499. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20689422.

[12] Sparks H, Nair G, Castellanos-Gonzalez A, et al. Treatment of Cryptosporidium: What We Know, Gaps, and the Way Forward. Curr Trop Med Reports. 2015;2:181–187.

[13] Gorla SK, Kavitha M, Zhang M, et al. Selective and potent urea inhibitors of cryptosporidium parvum inosine 5′-monophosphate dehydrogenase. J Med Chem. 2012;55:7759–7771.

[14] Gorla SK, McNair NN, Yang G, et al. Validation of IMP dehydrogenase inhibitors in a mouse model of cryptosporidiosis. Antimicrob Agents Chemother. 2014;58:1603–1614.

[15] Johnson CR, Gorla SK, Kavitha M, et al. Phthalazinone inhibitors of inosine-5'-monophosphate dehydrogenase from Cryptosporidium parvum. Bioorg Med Chem Lett [Internet]. 2013 [cited 2019 Oct 19];23:1004–1007. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23324406.

[16] Kirubakaran S, Gorla SK, Sharling L, et al. Structure-activity relationship study of selective benzimidazole-based inhibitors of Cryptosporidium parvum IMPDH. Bioorg Med Chem Lett [Internet]. 2012 [cited 2019 Oct 19];22:1985–1988. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22310229.

[17] Maurya SK, Gollapalli DR, Kirubakaran S, et al. Triazole inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. J Med Chem [Internet]. 2009 [cited 2019 Oct 19];52:4623–4630. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19624136.

[18] Sharling L, Liu X, Gollapalli DR, et al. A screening pipeline for antiparasitic agents targeting cryptosporidium inosine monophosphate dehydrogenase. PLoS Negl Trop Dis [Internet]. 2010 [cited 2019 Oct 19];4:e794. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/20706578.

[19] Umejiego NN, Gollapalli D, Sharling L, et al. Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against cryptosporidiosis. Chem Biol [Internet]. 2008 [cited 2019 Oct 19];15:70–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18215774.

[20] Sun Z, Khan J, Makowska-Grzyska M, et al. Synthesis, in vitro evaluation and cocrystal structure of 4-oxo-[1]benzopyrano[4,3-c]pyrazole Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase (CpIMPDH) inhibitors. J Med Chem [Internet]. 2014 [cited 2019 Oct 19];57:10544–10550. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25474504.

[21] Gorla SK, Kavitha M, Zhang M, et al. Optimization of benzoxazole-based inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. J Med Chem [Internet]. 2013 [cited 2019 Oct 19];56:4028–4043. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23668331.

[22] Macpherson IS, Kirubakaran S, Gorla SK, et al. The structural basis of Cryptosporidium -specific IMP dehydrogenase inhibitor selectivity. J Am Chem Soc [Internet]. 2010 [cited 2019 Oct 19];132:1230–1231. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20052976.

[23] Umejiego NN, Li C, Riera T, et al. Cryptosporidium parvum IMP dehydrogenase: Identification of functional, structural, and dynamic properties that can be exploited for drug design. J Biol Chem. 2004;279:40320–40327.

[24] Wang W, Hedstrom L. Kinetic mechanism of human inosine 5'-monophosphate dehydrogenase type II: random addition of substrates and ordered release of products. Biochemistry [Internet]. 1997 [cited 2019 Oct 26];36:8479–8483. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9214292.

[25] Kim Y, Makowska-Grzyska M, Gorla SK, et al. Structure of Cryptosporidium IMP

dehydrogenase bound to an inhibitor with in vivo antiparasitic activity. Acta Crystallogr Sect FStructural Biol Commun. 2015;71:531–538.

[26] Sliwoski G, Kothiwale S, Meiler J, et al. Computational methods in drug discovery [Internet]. Pharmacol. Rev. American Society for Pharmacology and Experimental Therapeutics; 2014 [cited 2021 Mar 25]. p. 334–395. Available from: /pmc/articles/PMC3880464/.

[27] Rossle NF, Latif B. Cryptosporidiosis as threatening health problem: A review [Internet]. Asian Pac. J. Trop. Biomed. China Humanity Technology Publishing House; 2013 [cited 2021 Mar 25]. p. 916–924. Available from: /pmc/articles/PMC3793167/.

[28] Holubová N, Zikmundová V, Limpouchová Z, et al. Cryptosporidium proventriculi sp. n. (Apicomplexa: Cryptosporidiidae) in Psittaciformes birds. Eur J Protistol [Internet]. 2019 [cited 2021 Feb 26];69:70–87. Available from: https://pubmed.ncbi.nlm.nih.gov/30981203/.

[29] Koehler A V., Whipp MJ, Haydon SR, et al. Cryptosporidium cuniculus - New records in human and kangaroo in Australia. Parasites and Vectors [Internet]. 2014 [cited 2021 Feb 26];7:492. Available from: http://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-014-0492-8.

[30] Li N, Xiao L, Alderisio K, et al. Subtyping Cryptosporidium ubiquitum, a zoonotic pathogen emerging in humans. Emerg Infect Dis [Internet]. 2014 [cited 2021 Feb 26];20:217–224. Available from: https://pubmed.ncbi.nlm.nih.gov/24447504/.

[31] Wang SN, Sun Y, Zhou HH, et al. Prevalence and genotypic identification of Cryptosporidium spp. And Enterocytozoon bieneusi in captive Asiatic black bears (Ursus thibetanus) in Heilongjiang and Fujian provinces of China. BMC Vet Res [Internet]. 2020 [cited 2021 Feb 26];16:84. Available from: https://bmcvetres.biomedcentral.com/articles/10.1186/s12917-020-02292-9.

[32] Diarrheal diseases - Our World in Data [Internet]. [cited 2021 Feb 26]. Available from:

https://ourworldindata.org/diarrheal-diseases.

[33] Rosado-García FM, Guerrero-Flórez M, Karanis G, et al. Water-borne protozoa parasites: The Latin American perspective. Int. J. Hyg. Environ. Health. Elsevier GmbH; 2017. p. 783–798.

[34] Ahmed SA, Karanis P. An overview of methods/techniques for the detection of Cryptosporidium in food samples [Internet]. Parasitol. Res. Springer Verlag; 2018 [cited 2021 Feb 26]. p. 629–653. Available from: https://doi.org/10.1007/s00436-017-5735-0.

[35] Ryan U, Hijjawi N, Xiao L. Foodborne cryptosporidiosis. Int. J. Parasitol. Elsevier Ltd; 2018. p. 1–12.

[36] Odeniran PO, Ademola IO. Epidemiology of Cryptosporidium infection in different hosts in Nigeria: A meta-analysis. Parasitol Int [Internet]. 2019;71:194–206. Available from: https://doi.org/10.1016/j.parint.2019.04.007.

[37] Squire SA, Ryan U. Cryptosporidium and Giardia in Africa: current and future challenges. Parasites and Vectors. 2017;10:1–32.

[38] Mor SM, Tzipori S. Cryptosporidiosis in children in sub-Saharan Africa: A lingering challenge [Internet]. Clin. Infect. Dis. Clin Infect Dis; 2008 [cited 2020 Oct 12]. p. 915–921. Available from: https://pubmed.ncbi.nlm.nih.gov/18715159/.

[39] Farthing MJ. Clinical aspects of human cryptosporidiosis. [Internet]. Contrib. Microbiol. Contrib Microbiol; 2000 [cited 2021 Mar 25]. p. 50–74. Available from: https://pubmed.ncbi.nlm.nih.gov/10943507/.

[40] O'Hara SP, Chen XM. The cell biology of cryptosporidium infection [Internet]. Microbes Infect. 2011 [cited 2021 Mar 26]. p. 721–730. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1286457911000918.

[41] Thompson A. Review of "Cryptosporidium and cryptosporidiosis" by Ronald Fayer and Lihua Xiao (eds.). Parasit Vectors [Internet]. 2008 [cited 2021 Mar 25];1:47. Available

from: http://parasitesandvectors.biomedcentral.com/articles/10.1186/1756-3305-1-47.

[42] Chappell C, Okhuysen P, Sterling C. Infectivity of Cryptosporidium parvum in healthy adults with pre-existing anti-C. Parvum serum immunoglobulin G. 1999.

[43] FAYER R, LEEK RG. The Effects of Reducing Conditions, Medium, pH, Temperature, and Time on in Vitro Excystation of Cryptosporidium. J Protozool [Internet]. 1984 [cited 2021 Mar 25];31:567–569. Available from: https://pubmed.ncbi.nlm.nih.gov/6512726/.

[44] REDUKER DW, SPEER CA, BLIXT JA. Ultrastructure of Cryptosporidium parvum Oocysts and Excysting Sporozoites as Revealed by High Resolution Scanning Electron Microscopy. J Protozool [Internet]. 1985 [cited 2021 Mar 25];32:708–711. Available from: https://pubmed.ncbi.nlm.nih.gov/4067883/.

[45] Robertson LJ, Campbell T, Smith H V. In vitro exeystation of Cryptosporidium parvum. Parasitology [Internet]. 1993 [cited 2021 Mar 25];106:13–19. Available from: https://www.cambridge.org/core/journals/parasitology/article/abs/in-vitro-excystation-of-cryptosporidium-parvum/A134917B120C7C0C68E84391901DA4C3.

[46] Hijjawi NS, Meloni BP, Ng'anzo M, et al. Complete development of Cryptosporidium parvum in host cell-free culture. Int J Parasitol [Internet]. 2004 [cited 2021 Mar 25];34:769–777. Available from: https://pubmed.ncbi.nlm.nih.gov/15157759/.

[47] Okhuysen PC, DuPont HL, Sterling CR, et al. Arginine aminopeptidase, an integral membrane protein of the Cryptosporidium parvum sporozoite [Internet]. Infect. Immun. Infect Immun; 1994 [cited 2021 Mar 25]. p. 4667–4670. Available from: https://pubmed.ncbi.nlm.nih.gov/7927738/.

[48] Snelling WJ, Lin Q, Moore JE, et al. Proteomics analysis and protein expression during sporozoite excystation of Cryptosporidium parvum (coccidia, apicomplexa). Mol Cell Proteomics [Internet]. 2007 [cited 2021 Mar 25];6:346–355. Available from: https://pubmed.ncbi.nlm.nih.gov/17124246/.

[49] Forney JR, Yang S, Healey MC. Protease activity associated with excystation of

Cryptosporidium parvum oocysts. J Parasitol. 1996;82:889–892.

[50] Pollok RCG, McDonald V, Kelly P, et al. The role of Cryptosporidium parvum-derived phospholipase in intestinal epithelial cell invasion. Parasitol Res [Internet]. 2003 [cited 2021 Mar 25];90:181–186. Available from: https://pubmed.ncbi.nlm.nih.gov/12783305/.

[51] White AC. Nitazoxanide: A new broad spectrum antiparasitic agent. Expert Rev Anti Infect Ther [Internet]. 2004 [cited 2021 Mar 25];2:43–49. Available from: https://pubmed.ncbi.nlm.nih.gov/15482170/.

[52] Fox LM, Saravolatz LD. Nitazoxanide: A new thiazolide antiparasitic agent [Internet]. Clin. Infect. Dis. Clin Infect Dis; 2005 [cited 2021 Mar 25]. p. 1173–1180. Available from: https://pubmed.ncbi.nlm.nih.gov/15791519/.

[53] Doumbo O, Rossignol JF, Pichard E, et al. Nitazoxanide in the treatment of cryptosporidial diarrhea and other intestinal parasitic infections associated with acquired immunodeficiency syndrome in tropical Africa. Am J Trop Med Hyg [Internet]. 1997 [cited 2021 Mar 25];56:637–639. Available from: https://pubmed.ncbi.nlm.nih.gov/9230795/.

[54] Favennec L, Jave Ortiz J, Gargala G, et al. Double-blind, randomized, placebo-controlled study of nitazoxanide in the treatment of fascioliasis in adults and children from northern Peru. Aliment Pharmacol Ther [Internet]. 2003 [cited 2021 Mar 25];17:265–270. Available from: https://pubmed.ncbi.nlm.nih.gov/12534412/.

[55] Rossignol JF, Hidalgo H, Feregrino M, et al. A double-'blind' placebo-controlled study of nitazoxanide in the treatment of cryptosporidial diarrhoea in AIDS patients in Mexico. Trans R Soc Trop Med Hyg [Internet]. 1998 [cited 2021 Mar 25];92:663–666. Available from: https://pubmed.ncbi.nlm.nih.gov/10326116/.

[56] Mandal S, Moudgil nal, Mandal SK. Rational drug design. Eur J Pharmacol. 2009;625:90–100.

[57] Greer J, Erickson JW, Baldwin JJ, et al. Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design. J Med Chem [Internet]. 1994

[cited 2021 Mar 25];37:1035–1054. Available from: https://pubs.acs.org/doi/abs/10.1021/jm00034a001.

[58] Muller B. Imatinib and Its Successors – How Modern Chemistry has Changed Drug Development. Curr Pharm Des [Internet]. 2008 [cited 2021 Mar 25];15:120–133. Available from: https://pubmed.ncbi.nlm.nih.gov/19149608/.

[59] Love MS, Beasley FC, Jumani RS, et al. A high-throughput phenotypic screen identifies clofazimine as a potential treatment for cryptosporidiosis. Geary TG, editor. PLoS Negl Trop Dis [Internet]. 2017 [cited 2021 Mar 25];11:e0005373. Available from: https://dx.plos.org/10.1371/journal.pntd.0005373.

[60] Irwin JJ, Shoichet BK. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. J Chem Inf Model [Internet]. 2005 [cited 2020 Oct 31];45:177. Available from: /pmc/articles/PMC1360656/?report=abstract.

[61] Fuller JC, Burgoyne NJ, Jackson RM. Predicting druggable binding sites at the protein-protein interface [Internet]. Drug Discov. Today. Drug Discov Today; 2009 [cited 2021 Mar 25]. p. 155–161. Available from: https://pubmed.ncbi.nlm.nih.gov/19041415/.

[62] Lionta E, Spyrou G, Vassilatis D, et al. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. Curr Top Med Chem [Internet]. 2014 [cited 2021 Mar 25];14:1923–1938. Available from: /pmc/articles/PMC4443793/.

[63] Hughes JP, Rees SS, Kalindjian SB, et al. Principles of early drug discovery [Internet]. Br. J. Pharmacol. Wiley-Blackwell; 2011 [cited 2021 Mar 25]. p. 1239–1249. Available from: /pmc/articles/PMC3058157/.

[64] Hait M. Lead Optimization : Research Pathways in Medicinal Chemistry Lead Optimization : Research Pathways in Medicinal Chemistry . Asian J. 2016;2–4.

[65] Bambini S, Rappuoli R. The use of genomics in microbial vaccine development [Internet]. Drug Discov. Today. Drug Discov Today; 2009 [cited 2021 Mar 25]. p. 252–260. Available from: https://pubmed.ncbi.nlm.nih.gov/19150507/.

[66] Lundstrom K. Genomics and drug discovery [Internet]. Future Med. Chem. Future Med Chem; 2011 [cited 2021 Mar 25]. p. 1855–1858. Available from: https://pubmed.ncbi.nlm.nih.gov/22023028/.

[67] Wang R, Fang X, Lu Y, et al. The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. J Med Chem [Internet]. 2004 [cited 2021 Mar 25];47:2977–2980. Available from: https://pubmed.ncbi.nlm.nih.gov/15163179/.

[68] Castellanos-Gonzalez A, White AC, Ojo KK, et al. A novel calcium-dependent protein kinase inhibitor as a lead compound for treating cryptosporidiosis. J Infect Dis [Internet]. 2013 [cited 2021 Mar 25];208:1342–1348. Available from: /pmc/articles/PMC3778970/.

[69] Vinayak S, Jumani RS, Miller P, et al. Bicyclic azetidines kill the diarrheal pathogen Cryptosporidium in mice by inhibiting parasite phenylalanyl-tRNA synthetase. Sci Transl Med [Internet]. 2020 [cited 2021 Mar 25];12. Available from: https://stm.sciencemag.org/content/12/563/eaba8412.

[70] Love MS, McNamara CW. Phenotypic screening techniques for *Cryptosporidium* drug discovery. Expert Opin Drug Discov [Internet]. 2021 [cited 2021 Mar 25];16:59–74. Available from: https://www.tandfonline.com/doi/full/10.1080/17460441.2020.1812577.

[71] Ji Y, Gu J, Makhov AM, et al. Regulation of the Interaction of Inosine Monophosphate Dehydrogenase with Mycophenolic Acid by GTP *. J Biol Chem. 2005;281:206–212.

[72] Zimmermann A, Gu JJ, Spychala J, et al. Inosine monophosphate dehydrogenase expression: Transcriptional regulation of the type I and type II genes. Adv Enzyme Regul [Internet]. Elsevier Ltd; 1996 [cited 2020 May 10]. p. 75–84. Available from: https://linkinghub.elsevier.com/retrieve/pii/0065257195000127.

[73] Pawlowic MC, Somepalli M, Sateriale A, et al. Genetic ablation of purine salvage in Cryptosporidium parvum reveals nucleotide uptake from the host cell. Proc Natl Acad Sci U S A [Internet]. 2019 [cited 2021 Mar 25];116:21160–21165. Available from:

/pmc/articles/PMC6800313/.

[74] Sintchak MD, Fleming MA, Futer O, et al. Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. Cell. 1996;85:921–930.

[75] Zhang RG, Evans G, Rotella FJ, et al. Characteristics and crystal structure of bacterial inosine-5’- monophosphate dehydrogenase. Biochemistry. 1999;38:4691–4700.

CHAPTER 2

2.1 Introduction

For both individuals and animals worldwide, cryptosporidiosis is a serious diarrheal disease caused by the many parasitic protozoa- *Cryptosporidium,* which oocysts are widespread in the ecosystem [27]. *Cryptosporidium* belongs to the Phylum- Protozoa, Subphylum- Apicomplexa, Class- Conoidasida, Subclass- Coccidia, Order- Eucoccidiorida, Suborder- Eimeriorina, Family- Cryptosporidiidae, and Genus- *Cryptosporidium*. At present, 41 species of *Cryptosporidium* have been recorded with greater than 60 genotypes. These species include C. *meleagridis,* C. *ubiquitum,* C. *muris,* C. *andersoni,* C. *parvum,* C. *hominis*. The commonest pathogenic species affecting both immune-competent and immune-compromised human species are C. *parvum* and C. *hominis,* which account for close to 90 percent of human infection [28–31]. Diarrhoeal diseases killed 1.6 million people worldwide in 2017. Children under five years of age accounted for one-third of these deaths, with the highest mortality rate from South Asia and sub-Saharan Africa, which are low-resource countries [32]. *Cryptosporidium* contamination has been reported as the most prominent cause of 905 waterborne outbreaks globally [33]. More than 8 million foodborne disease cases per year and 25 confirmed foodborne outbreaks have been reported to be responsible by *Cryptosporidium* [34,35].

2.2 Transmission and lifecycle of *Cryptosporidium*.

Cryptosporidium is the parasite responsible for cryptosporidiosis, affecting humans and animals worldwide. Cryptosporidiosis has been reported to be one of the leading causes of diarrhea. Hence it is an enteric pathogen parasitizing the epithelia of host intestinal cells [36]. There are several reported Cryptosporidium species, of which the first recognized species in humans was *Cryptosporidium parvum*. However, in recent years, both zoonotic and anthroponotic species have been reported, with *Cryptosporidium parvum* and *C. hominis* standing out as opportunistic pathogen of significance in immune-compromised individuals. Significant infections are more likely in children, pregnant women, aged, HIV/AIDS patients, and very ill individuals. Other reported species in humans include, *C. meleagridis, C. viatorum, C. bovis, C. cuniculus, C. ubiquitum, C. muris, C. xiaoi, C. felis, C. canis* and *C. suis,* most of which have their origins from

animals [37]. Clinical signs include acute diarrhea, weakness, dehydration, amongst others. It has also been established to be a disease of poverty, being related to contaminated drinking water and poor hygiene across the world [38]. In healthy individuals, cryptosporidiosis is self-limiting with 9-15 days [39]. Hence, infected individuals' immune system status could determine the severity and duration of disease symptoms, especially in intestinal cryptosporidiosis. Extraintestinal cryptosporidiosis has been reported in HIV patients, such as biliary cryptosporidiosis, though less common in immunocompetent patients [40]. Transmission of *Cryptosporidium* occurs through ingestion of water contaminated with feces (recreational, drinking or occupational water) or food (vegetables, raw milk, and fruits) or those associated with direct contact with infected animals (farm, domestic or wild) or humans. Excystation occurs after ingestion (and possibly inhalation) by the susceptible vertebrate host.

Cryptosporidium sporulated oocysts are shed in feces containing four sporozoites. The oocysts are resistant to tensed environmental factors such as temperature, rainfall, and humidity. The oocyst wall is composed of inner and outer protective walls containing protein-lipid-carbohydrate matrix [41]. Vertebrates ingest the oocysts from the environment, and the infectious cycle commences, of which all the developmental stages are completed in a single host [42] (Figure 2.1). The oocysts containing the four infective sporozoites excyst in the intestinal tract with the help of intrinsic (pancreatic enzymes, carbon dioxide and bile salts) and extrinsic (temperature and pH) factors [43–46]. Other molecules associated with excystation include arginine aminopeptidase [47], protein-associated molecules such as ribosomal-associated and heat shock proteins [48], sporozoite-associated serine and cysteine proteases [48,49] and secretory phospholipase A2 [50]. The sporozoite becomes encapsulated by the host membrane in form of a parasitophorous vacuole. The vacuole presents an extra-cytoplasmic form, yet largely considered intracellular because of its position within the host-derived parasitophorous vacuole membrane on top of epithelial cells. The trophozoite is formed and undergoes asexual reproduction by merogony forming type-I meronts [40]. Autoinfection by thin-walled oocysts and the recycling of type-I meronts is the major cause of persistent chronic infection, as observed in immunocompromised patients. Type II meronts are formed after series of developmental stages. Each of these can infect another enterocyte, with cells differentiating into gametocytes to form the sexual stage of cell development. Male or female gametocytes are formed, which are also known as microgametocytes and macrogametocytes,

respectively. Meanwhile the non-fertile microgametocytes fertilize the motile macrogametocytes to form the diploid zygote, which undergoes sporogony to produce the four haploid sporozoites within the oocyst, which is sporulated [40].

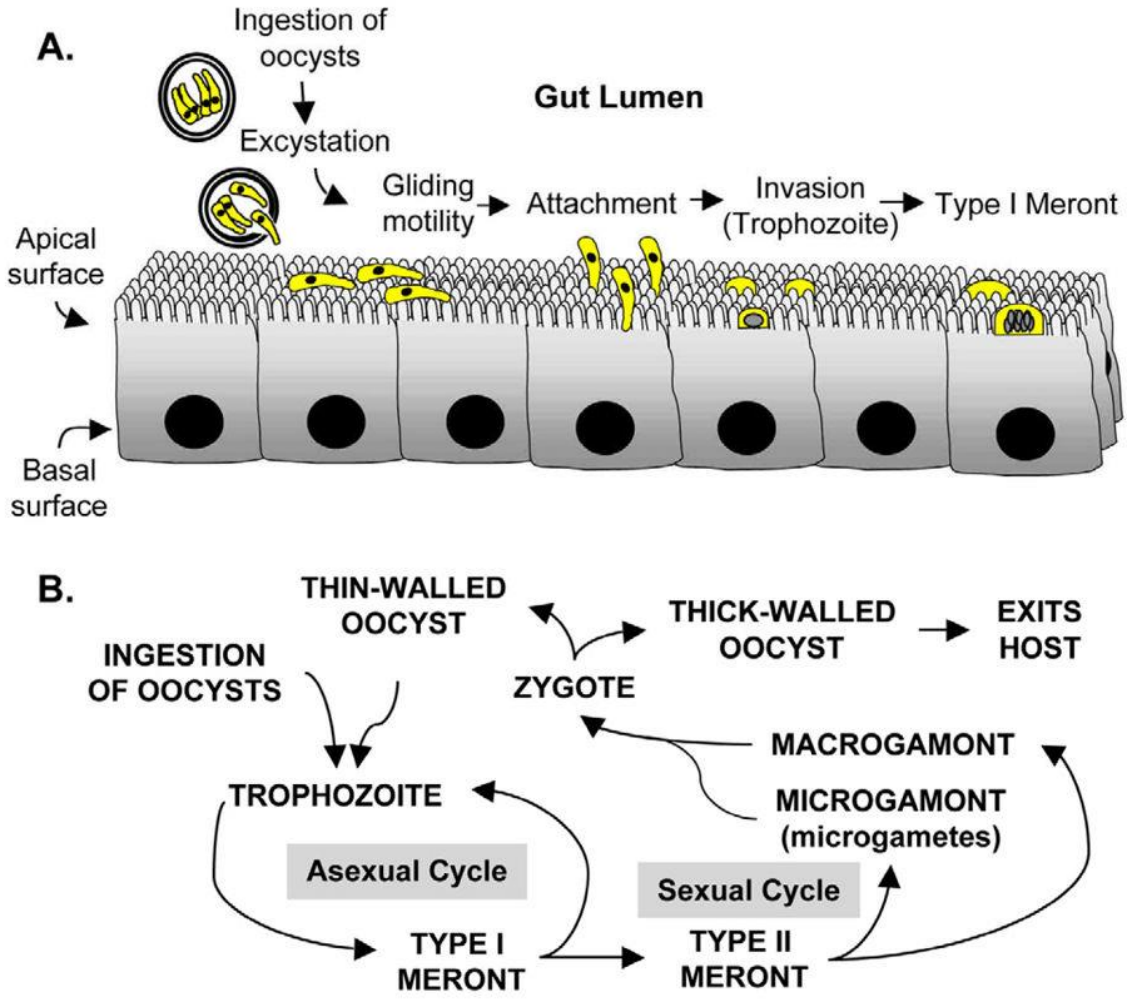


Figure 2.1 *Cryptosporidium* species lifecycle stages. A. Oocysts ingestion during *Cryptosporidium* infection. B. Sexual and asexual stages of *Cryptosporidium* in a single host [40]. (Public domain)

2.3 Treatment options of cryptosporidiosis

Immune-competent individuals will recover from cryptosporidiosis without treatment. Cases involving fatal diarrhea can be managed by drinking plenty of fluids to prevent dehydration. However, immune-compromised individuals are at higher risk for more severe and prolonged illness. Fluid replacement therapy options for children, pregnant women, and other immune-compromised patients are essential in fatal diarrhea cases to avoid sudden deaths associated with dehydration. Anti-diarrhoeic drugs could assist in slowing down diarrhea, but not reduce oocyst burden. Nitazoxanide has been FDA-approved for the treatment of diarrhea caused by *Cryptosporidium* species in immune-competent individuals [51]. Meanwhile, the effectiveness of nitazoxanide in immune-compromised individuals is still unclear. There are reports of individuals with HIV/AIDS but are placed on anti-retroviral therapy to improve their immune status, hence decreasing the burden associated with cryptosporidiosis. There is often cryptosporidiosis relapse when the immune status worsens. Nitazoxanide is presently not approved to treat immune-compromised individuals because nitazoxanide tablets and nitazoxanide oral suspension have not been shown to be superior to placebo for the treatment of fatal diarrhea caused by *Cryptosporidium* in HIV-infected or immune-deficient individuals [52]. Immune-competent individuals with cryptosporidiosis treated with multiple three-day courses of nitazoxanide have been effective. During early studies, seven-day treatment courses have been used for cryptosporidiosis and other parasitic infections with positive responses [53,54]. HIV/AIDS patients with *Cryptosporidium*-associated diarrhea received the drug for 28 days [55].

2.4 Rational drug targeting in the treatment of cryptosporidiosis

Rational drug design (RDD) harnesses the advances in the current bioinformatics tool. RDD identifies a biomolecular target in the parasite, which is crucial to the survival of the organism and the inhibition of which will lead to the impairment in the parasite's pathogenesis [56–58]. Once the targets have been recognized, small-molecule inhibitors are screened against them via high-throughput screening of chemical libraries or virtual screening using the available drug database [59,60]. The small inhibitors can perturb the functions of the identified targets in two ways.

1. It can stop the cross-talks between the identified targets and other biomolecular targets necessary for the progression of the disease [61].

2. By activating and up-regulating the biomolecules that are important for the normal functioning of the living system but has been hitherto suppressed in diseased condition [56].

Generally, in practice, the inhibitors must have a more competitive and higher binding affinity for the identified targets' active sites more than the natural ligands of the targets [62]. The small molecules which inhibit the target effectively are tagged as the lead compounds [63]. The compounds are then subjected to further optimization processes. In the optimization process, the properties below are evaluated and optimized initially by *in silico* methods [56,64]

- The hydrophilicity/lipophilicity balance
- The ADMET properties (Absorption, Distribution, Metabolism, Excretion, and Toxicity)
- Quantitative structure-activity relationship (QSAR)
- Quantitative structure-property relationship (QSPR)
- Biodegradation and biotransformation products, etc.

RDD in drug discovery has grown exponentially due to the advances in computer-aided methodologies, molecular docking tools, and availability of the 3D X-ray and NMR structures of the biological targets[65–67]. Using the RDD approach, some advances have been made in the discovery of drugs against cryptosporidiosis. *Cryptosporidim parvum* calcium-dependent protein kinase 1 (CDPK1) is a validated target, and the bumped kinase inhibitor-1294 (BK1-1294) has been recognized as a putative and effective inhibitor of CDPK-1 [68]. Other targets in *Cryptosporidium parvum spp* are phenylalanl-tRNA synthetase with BRD7929 as its inhibitor [69], long chain fatty acid co-enzyme A (*Cp*LC-FACS) is inhibited by compound R134 [70]. Lysyl tRNA-synthetase has compound 5 as its inhibitor [70]. Another critical target in *Cryptosporidium* species is inosine monophosphate dehydrogenase (IMPDH) which will be discussed in the next section.

2.5 IMPDH as a potential therapeutic target in *Cryptosporidium*

Inosine monophosphate dehydrogenase (IMPDH) is the first and rate-limiting enzyme in the de novo synthesis of guanine nucleotide. It does this by converting inosine monophosphate to xanthosine monophosphate (XMP) while reducing NAD^+ to NADH at the same time [71]. This reaction bridges the synthesis of adenine and guanine nucleotide biosynthesis (Figure 2.2). This

implies that IMPDH regulates the pool of nucleotides in the living system. As it controls the RNA and DNA biosynthesis, IMPDH is vital in signal transduction and other mechanisms that are important to cell proliferation [72]. Therefore the inhibition of *Cryptosporidium parvum* IMPDH (*Cp*IMPDH) will lead to the diminution of both guanine and the adenylate nucleotide pool, eventually resulting in cell death [73]. Thus, targeting IMPDH has been validated, and it is repeatedly being targeted in the treatment of cryptosporidiosis [14,15,21,25]. It is noteworthy that *Cryptosporidium parvum*, despite being an eukaryote, has a prokaryotic IMPDH made possible by lateral gene transfer from ε-proteobacterium [19]. This makes selective targeting of a prokaryotic IMPDH in a eukaryotic host possible.

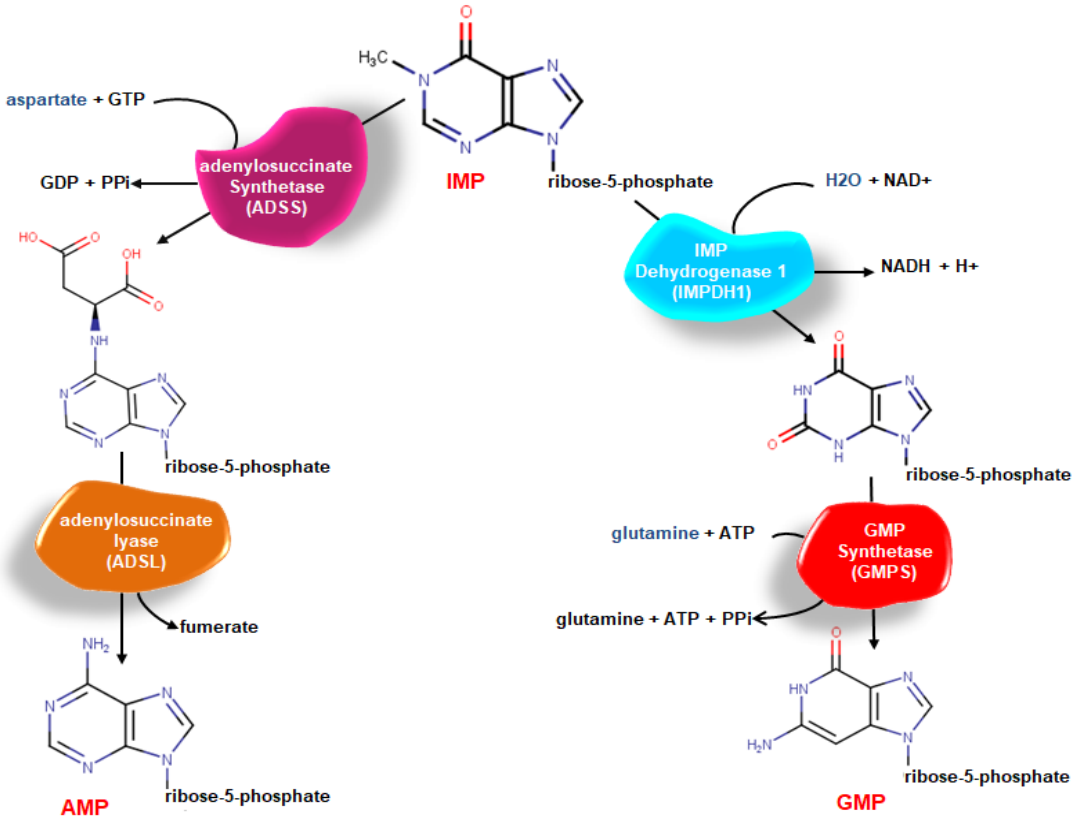


Figure 2.2 Reaction Pathway showing the conversion of inosine monophosphate (IMP) to guanine monophosphate (GMP) and adenosine monophosphate (AMP). IMP dehydrogenase (IMPDH) catalyzes the rate-limiting step in the conversion to GMP (prepared by author)

The x-ray crystal structures of IMPDH have been determined by various research groups either as an apo, in complex with ligands, cofactors, inhibitors, etc. [25,74,75]. IMPDH is a 4-fold

symmetric and homotetrameric structure (Figure 2.3). The tetramer has four active sites, which are situated close to the subunit interfaces [25]. The monomer has two binding pockets, which are the active site and the co-factor site. The active site naturally binds IMP, it is confined in a monomer, and its binding residues are strictly conserved in all IMPDHs. The co-factor site which binds NAD^+ has adenosine and nicotinamide subsites. The nicotinamide subsite lies within the same monomer as the IMP binding site. However, the adenosine subsite is contained in an adjacent monomer (peculiar to prokaryotic IMPDH alone). Therefore any inhibitor targeting the NAD^+ site (cofactor site) of prokaryotic IMPDH, their binding will traverse an IMPDH dimer [25]. In eukaryotes, however, the adenosine subsite is located in the same monomer as the active site. The adenosine site is highly divergent among organisms, and this accounts for the mechanism of selection of anticryptosporidials [25].

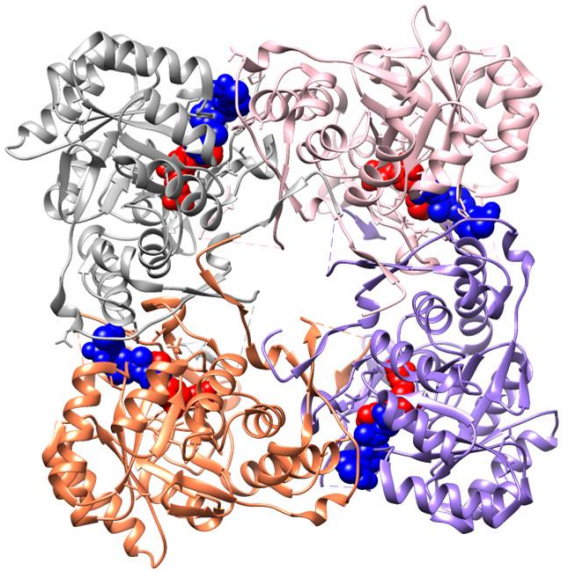


Figure 2.3 X-ray crystal structure of tetrameric IMPDH. Chain A (dark grey), chain B (coral), chain C (medium purple), and chain D (pink) The NAD^+ site (co-factor site) is colored blue while the active site is in red (prepared by author)

2.6 P131 as a potential drug candidate

The research group of Hedstrom *et al.* was involved in a program to develop potential anticryptosporidials which specifically target IMPDH. This led to the synthesis of numerous compounds that were structurally distinct. These compounds were tested in a mouse model and

displayed a 1000-fold in potency and selectivity over the human IMPDH counterpart [13,15–17,19,22]. Of all these compounds, eight were most promising when tested in an interleukin-12 knockout mouse model. Compound P131 had the most remarkable antiparasitic effect. P131 when administered at a single dose of 250mg/kg body weight, showed an equivalent activity compared to the control group, which was treated with paromomycin at 2000mg/kg body weight. However, when both drugs were administered thrice-daily at the same concentration (200mg/kg), P131 elicited a superior anticryptosporidial activity when compared to paromomycin[14]. X-ray crystal structure of *Cp*IMPDH showed that P131 binds in the co-factor site of *Cp*IMPDH. P131 has two aromatic rings (Figure 2.4). In *Cp*IMPDH, one of these rings interacts with the hypoxanthine ring of IMP while the other with Tyr358 residue in the adjacent subunit [25].

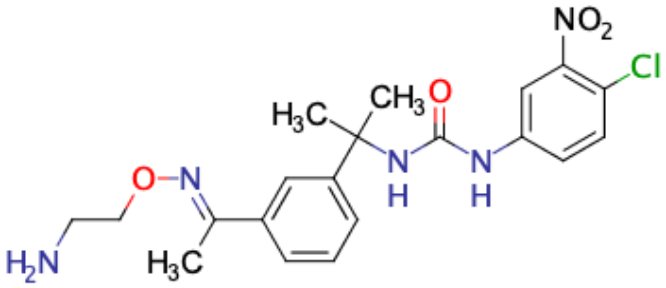


Figure 2.4 2D structure of P131, a novel *Cryptosporidium* inhibitor that elicited a superior parasiticidal activity compared to paromomycin. (prepared by author)

Computational chemistry is vital to drug discovery processes as it reduces the time and capital investment involved in the pure traditional method of drug discovery [26]. Newer and more efficacious drug candidates are needed for the treatment of cryptosporidiosis. The faster means of achieving this is the integration of computational techniques. To achieve the study's aim and objectives stated earlier, we employed various computational methods such as molecular docking and modeling, homology modeling, virtual screening, pharmacophore modeling, molecular dynamics simulation, etc. All these are expounded in subsequent chapters.

2.7 References

[1] Bouzid M, Kintz E, Hunter PR. Risk factors for Cryptosporidium infection in low and middle income countries: A systematic review and meta-analysis. PLoS Negl Trop Dis. 2018;12.

[2] Fischer Walker CL, Aryee MJ, Boschi-Pinto C, et al. Estimating diarrhea mortality among young children in low and middle income countries. PLoS One. 2012.

[3] Valentiner-Branth P, Steinsland H, Fischer TK, et al. Cohort study of Guinean children: Incidence, pathogenicity, conferred protection, and attributable risk for enteropathogens during the first 2 years of life. J Clin Microbiol. 2003;41:4238–4245.

[4] Naghavi M, Abajobir AA, Abbafati C, et al. Global, regional, and national age-sex specifc mortality for 264 causes of death, 1980-2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet [Internet]. 2017 [cited 2020 Oct 8];390:1151–1210. Available from: https://pubmed.ncbi.nlm.nih.gov/28919116/.

[5] Kotloff K, Nataro J, Blackwelder W. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. 2013.

[6] Sow S, Muhsen K, Nasrin D. The Burden of Cryptosporidium Diarrheal Disease among Children &lt. 2016.

[7] Lee S, Harwood M, Girouard D, et al. The therapeutic efficacy of azithromycin and nitazoxanide in the acute pig model of Cryptosporidium hominis. Yeruva L, editor. PLoS One [Internet]. 2017 [cited 2020 Apr 20];12:e0185906. Available from: https://dx.plos.org/10.1371/journal.pone.0185906.

[8] Checkley W, Jr. AW, Jaganath D. A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. 2015.

[9] Hewitt RG, Yiannoutsos CT, Higgs ES, et al. Paromomycin: no more effective than placebo

for treatment of cryptosporidiosis in patients with advanced human immunodeficiency virus infection. AIDS Clinical Trial Group. Clin Infect Dis [Internet]. 2000 [cited 2019 Oct 7];31:1084–1092. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11049793.

[10] Hussien SMM, Abdella OH, Abu-Hashim AH, et al. Comparative study between the effect of nitazoxanide and paromomycine in treatment of cryptosporidiosis in hospitalized children. J Egypt Soc Parasitol [Internet]. 2013 [cited 2019 Oct 7];43:463–470. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24260825.

[11] Cabada MM, White AC. Treatment of cryptosporidiosis: do we know what we think we know? Curr Opin Infect Dis [Internet]. 2010 [cited 2019 Oct 7];23:494–499. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20689422.

[12] Sparks H, Nair G, Castellanos-Gonzalez A, et al. Treatment of Cryptosporidium: What We Know, Gaps, and the Way Forward. Curr Trop Med Reports. 2015;2:181–187.

[13] Gorla SK, Kavitha M, Zhang M, et al. Selective and potent urea inhibitors of cryptosporidium parvum inosine 5′-monophosphate dehydrogenase. J Med Chem. 2012;55:7759–7771.

[14] Gorla SK, McNair NN, Yang G, et al. Validation of IMP dehydrogenase inhibitors in a mouse model of cryptosporidiosis. Antimicrob Agents Chemother. 2014;58:1603–1614.

[15] Johnson CR, Gorla SK, Kavitha M, et al. Phthalazinone inhibitors of inosine-5'-monophosphate dehydrogenase from Cryptosporidium parvum. Bioorg Med Chem Lett [Internet]. 2013 [cited 2019 Oct 19];23:1004–1007. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23324406.

[16] Kirubakaran S, Gorla SK, Sharling L, et al. Structure-activity relationship study of selective benzimidazole-based inhibitors of Cryptosporidium parvum IMPDH. Bioorg Med Chem Lett [Internet]. 2012 [cited 2019 Oct 19];22:1985–1988. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22310229.

[17] Maurya SK, Gollapalli DR, Kirubakaran S, et al. Triazole inhibitors of Cryptosporidium

parvum inosine 5’-monophosphate dehydrogenase. J Med Chem [Internet]. 2009 [cited 2019 Oct 19];52:4623–4630. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19624136.

[18] Sharling L, Liu X, Gollapalli DR, et al. A screening pipeline for antiparasitic agents targeting cryptosporidium inosine monophosphate dehydrogenase. PLoS Negl Trop Dis [Internet]. 2010 [cited 2019 Oct 19];4:e794. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20706578.

[19] Umejiego NN, Gollapalli D, Sharling L, et al. Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against cryptosporidiosis. Chem Biol [Internet]. 2008 [cited 2019 Oct 19];15:70–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18215774.

[20] Sun Z, Khan J, Makowska-Grzyska M, et al. Synthesis, in vitro evaluation and cocrystal structure of 4-oxo-[1]benzopyrano[4,3-c]pyrazole Cryptosporidium parvum inosine 5’-monophosphate dehydrogenase (CpIMPDH) inhibitors. J Med Chem [Internet]. 2014 [cited 2019 Oct 19];57:10544–10550. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25474504.

[21] Gorla SK, Kavitha M, Zhang M, et al. Optimization of benzoxazole-based inhibitors of Cryptosporidium parvum inosine 5’-monophosphate dehydrogenase. J Med Chem [Internet]. 2013 [cited 2019 Oct 19];56:4028–4043. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23668331.

[22] Macpherson IS, Kirubakaran S, Gorla SK, et al. The structural basis of Cryptosporidium -specific IMP dehydrogenase inhibitor selectivity. J Am Chem Soc [Internet]. 2010 [cited 2019 Oct 19];132:1230–1231. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20052976.

[23] Umejiego NN, Li C, Riera T, et al. Cryptosporidium parvum IMP dehydrogenase: Identification of functional, structural, and dynamic properties that can be exploited for drug design. J Biol Chem. 2004;279:40320–40327.

[24] Wang W, Hedstrom L. Kinetic mechanism of human inosine 5'-monophosphate dehydrogenase type II: random addition of substrates and ordered release of products. Biochemistry [Internet]. 1997 [cited 2019 Oct 26];36:8479–8483. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9214292.

[25] Kim Y, Makowska-Grzyska M, Gorla SK, et al. Structure of Cryptosporidium IMP dehydrogenase bound to an inhibitor with in vivo antiparasitic activity. Acta Crystallogr Sect FStructural Biol Commun. 2015;71:531–538.

[26] Sliwoski G, Kothiwale S, Meiler J, et al. Computational methods in drug discovery [Internet]. Pharmacol. Rev. American Society for Pharmacology and Experimental Therapeutics; 2014 [cited 2021 Mar 25]. p. 334–395. Available from: /pmc/articles/PMC3880464/.

[27] Rossle NF, Latif B. Cryptosporidiosis as threatening health problem: A review [Internet]. Asian Pac. J. Trop. Biomed. China Humanity Technology Publishing House; 2013 [cited 2021 Mar 25]. p. 916–924. Available from: /pmc/articles/PMC3793167/.

[28] Holubová N, Zikmundová V, Limpouchová Z, et al. Cryptosporidium proventriculi sp. n. (Apicomplexa: Cryptosporidiidae) in Psittaciformes birds. Eur J Protistol [Internet]. 2019 [cited 2021 Feb 26];69:70–87. Available from: https://pubmed.ncbi.nlm.nih.gov/30981203/.

[29] Koehler A V., Whipp MJ, Haydon SR, et al. Cryptosporidium cuniculus - New records in human and kangaroo in Australia. Parasites and Vectors [Internet]. 2014 [cited 2021 Feb 26];7:492. Available from: http://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-014-0492-8.

[30] Li N, Xiao L, Alderisio K, et al. Subtyping Cryptosporidium ubiquitum, a zoonotic pathogen emerging in humans. Emerg Infect Dis [Internet]. 2014 [cited 2021 Feb 26];20:217–224. Available from: https://pubmed.ncbi.nlm.nih.gov/24447504/.

[31] Wang SN, Sun Y, Zhou HH, et al. Prevalence and genotypic identification of

Cryptosporidium spp. And Enterocytozoon bieneusi in captive Asiatic black bears (Ursus thibetanus) in Heilongjiang and Fujian provinces of China. BMC Vet Res [Internet]. 2020 [cited 2021 Feb 26];16:84. Available from: https://bmcvetres.biomedcentral.com/articles/10.1186/s12917-020-02292-9.

[32] Diarrheal diseases - Our World in Data [Internet]. [cited 2021 Feb 26]. Available from: https://ourworldindata.org/diarrheal-diseases.

[33] Rosado-García FM, Guerrero-Flórez M, Karanis G, et al. Water-borne protozoa parasites: The Latin American perspective. Int. J. Hyg. Environ. Health. Elsevier GmbH; 2017. p. 783–798.

[34] Ahmed SA, Karanis P. An overview of methods/techniques for the detection of Cryptosporidium in food samples [Internet]. Parasitol. Res. Springer Verlag; 2018 [cited 2021 Feb 26]. p. 629–653. Available from: https://doi.org/10.1007/s00436-017-5735-0.

[35] Ryan U, Hijjawi N, Xiao L. Foodborne cryptosporidiosis. Int. J. Parasitol. Elsevier Ltd; 2018. p. 1–12.

[36] Odeniran PO, Ademola IO. Epidemiology of Cryptosporidium infection in different hosts in Nigeria: A meta-analysis. Parasitol Int [Internet]. 2019;71:194–206. Available from: https://doi.org/10.1016/j.parint.2019.04.007.

[37] Squire SA, Ryan U. Cryptosporidium and Giardia in Africa: current and future challenges. Parasites and Vectors. 2017;10:1–32.

[38] Mor SM, Tzipori S. Cryptosporidiosis in children in sub-Saharan Africa: A lingering challenge [Internet]. Clin. Infect. Dis. Clin Infect Dis; 2008 [cited 2020 Oct 12]. p. 915–921. Available from: https://pubmed.ncbi.nlm.nih.gov/18715159/.

[39] Farthing MJ. Clinical aspects of human cryptosporidiosis. [Internet]. Contrib. Microbiol. Contrib Microbiol; 2000 [cited 2021 Mar 25]. p. 50–74. Available from: https://pubmed.ncbi.nlm.nih.gov/10943507/.

[40] O'Hara SP, Chen XM. The cell biology of cryptosporidium infection [Internet]. Microbes Infect. 2011 [cited 2021 Mar 26]. p. 721–730. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1286457911000918.

[41] Thompson A. Review of "Cryptosporidium and cryptosporidiosis" by Ronald Fayer and Lihua Xiao (eds.). Parasit Vectors [Internet]. 2008 [cited 2021 Mar 25];1:47. Available from: http://parasitesandvectors.biomedcentral.com/articles/10.1186/1756-3305-1-47.

[42] Chappell C, Okhuysen P, Sterling C. Infectivity of Cryptosporidium parvum in healthy adults with pre-existing anti-C. Parvum serum immunoglobulin G. 1999.

[43] FAYER R, LEEK RG. The Effects of Reducing Conditions, Medium, pH, Temperature, and Time on in Vitro Excystation of Cryptosporidium. J Protozool [Internet]. 1984 [cited 2021 Mar 25];31:567–569. Available from: https://pubmed.ncbi.nlm.nih.gov/6512726/.

[44] REDUKER DW, SPEER CA, BLIXT JA. Ultrastructure of Cryptosporidium parvum Oocysts and Excysting Sporozoites as Revealed by High Resolution Scanning Electron Microscopy. J Protozool [Internet]. 1985 [cited 2021 Mar 25];32:708–711. Available from: https://pubmed.ncbi.nlm.nih.gov/4067883/.

[45] Robertson LJ, Campbell T, Smith H V. In vitro exeystation of Cryptosporidium parvum. Parasitology [Internet]. 1993 [cited 2021 Mar 25];106:13–19. Available from: https://www.cambridge.org/core/journals/parasitology/article/abs/in-vitro-excystation-of-cryptosporidium-parvum/A134917B120C7C0C68E84391901DA4C3.

[46] Hijjawi NS, Meloni BP, Ng'anzo M, et al. Complete development of Cryptosporidium parvum in host cell-free culture. Int J Parasitol [Internet]. 2004 [cited 2021 Mar 25];34:769–777. Available from: https://pubmed.ncbi.nlm.nih.gov/15157759/.

[47] Okhuysen PC, DuPont HL, Sterling CR, et al. Arginine aminopeptidase, an integral membrane protein of the Cryptosporidium parvum sporozoite [Internet]. Infect. Immun. Infect Immun; 1994 [cited 2021 Mar 25]. p. 4667–4670. Available from: https://pubmed.ncbi.nlm.nih.gov/7927738/.

[48] Snelling WJ, Lin Q, Moore JE, et al. Proteomics analysis and protein expression during sporozoite excystation of Cryptosporidium parvum (coccidia, apicomplexa). Mol Cell Proteomics [Internet]. 2007 [cited 2021 Mar 25];6:346–355. Available from: https://pubmed.ncbi.nlm.nih.gov/17124246/.

[49] Forney JR, Yang S, Healey MC. Protease activity associated with excystation of Cryptosporidium parvum oocysts. J Parasitol. 1996;82:889–892.

[50] Pollok RCG, McDonald V, Kelly P, et al. The role of Cryptosporidium parvum-derived phospholipase in intestinal epithelial cell invasion. Parasitol Res [Internet]. 2003 [cited 2021 Mar 25];90:181–186. Available from: https://pubmed.ncbi.nlm.nih.gov/12783305/.

[51] White AC. Nitazoxanide: A new broad spectrum antiparasitic agent. Expert Rev Anti Infect Ther [Internet]. 2004 [cited 2021 Mar 25];2:43–49. Available from: https://pubmed.ncbi.nlm.nih.gov/15482170/.

[52] Fox LM, Saravolatz LD. Nitazoxanide: A new thiazolide antiparasitic agent [Internet]. Clin. Infect. Dis. Clin Infect Dis; 2005 [cited 2021 Mar 25]. p. 1173–1180. Available from: https://pubmed.ncbi.nlm.nih.gov/15791519/.

[53] Doumbo O, Rossignol JF, Pichard E, et al. Nitazoxanide in the treatment of cryptosporidial diarrhea and other intestinal parasitic infections associated with acquired immunodeficiency syndrome in tropical Africa. Am J Trop Med Hyg [Internet]. 1997 [cited 2021 Mar 25];56:637–639. Available from: https://pubmed.ncbi.nlm.nih.gov/9230795/.

[54] Favennec L, Jave Ortiz J, Gargala G, et al. Double-blind, randomized, placebo-controlled study of nitazoxanide in the treatment of fascioliasis in adults and children from northern Peru. Aliment Pharmacol Ther [Internet]. 2003 [cited 2021 Mar 25];17:265–270. Available from: https://pubmed.ncbi.nlm.nih.gov/12534412/.

[55] Rossignol JF, Hidalgo H, Feregrino M, et al. A double-'blind' placebo-controlled study of nitazoxanide in the treatment of cryptosporidial diarrhoea in AIDS patients in Mexico. Trans R Soc Trop Med Hyg [Internet]. 1998 [cited 2021 Mar 25];92:663–666. Available

from: https://pubmed.ncbi.nlm.nih.gov/10326116/.

[56] Mandal S, Moudgil nal, Mandal SK. Rational drug design. Eur J Pharmacol. 2009;625:90–100.

[57] Greer J, Erickson JW, Baldwin JJ, et al. Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design. J Med Chem [Internet]. 1994 [cited 2021 Mar 25];37:1035–1054. Available from: https://pubs.acs.org/doi/abs/10.1021/jm00034a001.

[58] Muller B. Imatinib and Its Successors – How Modern Chemistry has Changed Drug Development. Curr Pharm Des [Internet]. 2008 [cited 2021 Mar 25];15:120–133. Available from: https://pubmed.ncbi.nlm.nih.gov/19149608/.

[59] Love MS, Beasley FC, Jumani RS, et al. A high-throughput phenotypic screen identifies clofazimine as a potential treatment for cryptosporidiosis. Geary TG, editor. PLoS Negl Trop Dis [Internet]. 2017 [cited 2021 Mar 25];11:e0005373. Available from: https://dx.plos.org/10.1371/journal.pntd.0005373.

[60] Irwin JJ, Shoichet BK. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. J Chem Inf Model [Internet]. 2005 [cited 2020 Oct 31];45:177. Available from: /pmc/articles/PMC1360656/?report=abstract.

[61] Fuller JC, Burgoyne NJ, Jackson RM. Predicting druggable binding sites at the protein-protein interface [Internet]. Drug Discov. Today. Drug Discov Today; 2009 [cited 2021 Mar 25]. p. 155–161. Available from: https://pubmed.ncbi.nlm.nih.gov/19041415/.

[62] Lionta E, Spyrou G, Vassilatis D, et al. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. Curr Top Med Chem [Internet]. 2014 [cited 2021 Mar 25];14:1923–1938. Available from: /pmc/articles/PMC4443793/.

[63] Hughes JP, Rees SS, Kalindjian SB, et al. Principles of early drug discovery [Internet]. Br. J. Pharmacol. Wiley-Blackwell; 2011 [cited 2021 Mar 25]. p. 1239–1249. Available from: /pmc/articles/PMC3058157/.

[64] Hait M. Lead Optimization : Research Pathways in Medicinal Chemistry Lead Optimization : Research Pathways in Medicinal Chemistry . Asian J. 2016;2–4.

[65] Bambini S, Rappuoli R. The use of genomics in microbial vaccine development [Internet]. Drug Discov. Today. Drug Discov Today; 2009 [cited 2021 Mar 25]. p. 252–260. Available from: https://pubmed.ncbi.nlm.nih.gov/19150507/.

[66] Lundstrom K. Genomics and drug discovery [Internet]. Future Med. Chem. Future Med Chem; 2011 [cited 2021 Mar 25]. p. 1855–1858. Available from: https://pubmed.ncbi.nlm.nih.gov/22023028/.

[67] Wang R, Fang X, Lu Y, et al. The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. J Med Chem [Internet]. 2004 [cited 2021 Mar 25];47:2977–2980. Available from: https://pubmed.ncbi.nlm.nih.gov/15163179/.

[68] Castellanos-Gonzalez A, White AC, Ojo KK, et al. A novel calcium-dependent protein kinase inhibitor as a lead compound for treating cryptosporidiosis. J Infect Dis [Internet]. 2013 [cited 2021 Mar 25];208:1342–1348. Available from: /pmc/articles/PMC3778970/.

[69] Vinayak S, Jumani RS, Miller P, et al. Bicyclic azetidines kill the diarrheal pathogen Cryptosporidium in mice by inhibiting parasite phenylalanyl-tRNA synthetase. Sci Transl Med [Internet]. 2020 [cited 2021 Mar 25];12. Available from: https://stm.sciencemag.org/content/12/563/eaba8412.

[70] Love MS, McNamara CW. Phenotypic screening techniques for *Cryptosporidium* drug discovery. Expert Opin Drug Discov [Internet]. 2021 [cited 2021 Mar 25];16:59–74. Available from: https://www.tandfonline.com/doi/full/10.1080/17460441.2020.1812577.

[71] Ji Y, Gu J, Makhov AM, et al. Regulation of the Interaction of Inosine Monophosphate Dehydrogenase with Mycophenolic Acid by GTP *. J Biol Chem. 2005;281:206–212.

[72] Zimmermann A, Gu JJ, Spychala J, et al. Inosine monophosphate dehydrogenase expression: Transcriptional regulation of the type I and type II genes. Adv Enzyme Regul

[Internet]. Elsevier Ltd; 1996 [cited 2020 May 10]. p. 75–84. Available from: https://linkinghub.elsevier.com/retrieve/pii/0065257195000127.

[73] Pawlowic MC, Somepalli M, Sateriale A, et al. Genetic ablation of purine salvage in Cryptosporidium parvum reveals nucleotide uptake from the host cell. Proc Natl Acad Sci U S A [Internet]. 2019 [cited 2021 Mar 25];116:21160–21165. Available from: /pmc/articles/PMC6800313/.

[74] Sintchak MD, Fleming MA, Futer O, et al. Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. Cell. 1996;85:921–930.

[75] Zhang RG, Evans G, Rotella FJ, et al. Characteristics and crystal structure of bacterial inosine-5'- monophosphate dehydrogenase. Biochemistry. 1999;38:4691–4700.

CHAPTER THREE

3.1 Introduction to methods in Computational Chemistry

Computational chemistry, which is also known as molecular modeling, is defined as the consortium of techniques carried out principally on a computer to model, probe, and understand the atomistic behavior of molecular systems, whether large or small [1]. In theoretical chemistry, drug design driven by molecular modeling protocols is rapidly gaining ground as it allows the integration of chemical theories and models with experimental inferences [2]. Computational chemistry also utilizes artificial intelligence, algorithms, and statistics to coin out large database, meaningful, understandable, and effective solutions to chemical and chemical-related challenges. Though a scientifically exciting field, computational chemistry is laden with its challenges, two of which are the choice of theory to adapt in interpreting a chemical problem and assessing the quality of and synthesizing the results obtained to draw out a meaningful extrapolation [3]. In the description of computational chemistry, three standard methods can be used. They are quantum mechanics, molecular mechanics, and molecular dynamics simulation with the latter mainly focusing on the Newtonian laws of physics. In this chapter, the computational methods utilized in this study are elucidated in a step-wise fashion.

3.2 Quantum Mechanics.

Quantum mechanics, which involves studying the chemistry of molecules at the electron levels was initiated, invented, and developed by Werner Heisenberg, Max Born, Pascal Jordan, Max Planck, Wolfang Pauli, Neils Bohr, Albert Einstein, and Paul Dirac [4]. In this approach, the electrons making up an atomic particle are wave-like properties and are regarded as having their energies (quantum particles) [5]. When analyzing a biomolecular system using the quantum mechanics method, the electrons are arranged in a 3D space and subsequently mapped using the continuous electronic density method [6]. The Schrodinger equation gives insight into the bioenergetics and movement of the biomolecular systems and how these wave-like properties behave is presented in the next section.

3.3 The Schrödinger equation

Erwin Schrodinger, born in 1926, put forward the time-dependent equation of quantum mechanics [7, 8]. This is described as

$$\mathbf{H}\psi = \mathbf{E}\psi \qquad \textbf{Eq. 3.1}$$

Where,

H – Hamiltonian operator,

E - Total energy of the system

Ψ - Wave function of the chemical system. It also represents the Eigenfunction which gives the probable but not exact location of the nuclei and electrons. The molecular Hamiltonian is an expression of the sum of its operators:

$$H=T+V \qquad \textbf{Eq. 3.2}$$

Where,

H denotes Hamiltonian operator, T is the system's kinetic energy operator, and V denotes potential energy operator. The Hamiltonian has the propensity to be composed of all the kinetic and potential energy operators for the entire electrons and nuclei if particles are treated as point masses ad other effects are neglected. The Hamiltonian equation can also be denoted as:

$$H = -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 - \frac{\hbar^2}{2}\sum_A \frac{1}{M_A}\nabla_A^2 - \sum_A \sum_i \frac{Z_A e^2}{r_{Ai}} + \sum_i \sum_{j>i} \frac{e^2}{r_{ij}} + \sum_A \sum_{B>A} \frac{Z_A Z_B e^2}{R_{AB}} \qquad \textbf{Eq. 3.3}$$

Here, A and B are the nuclei, i and j are the electrons, M_A is the nucleus mass, m_e is electron mass, R_{AB} is a nucleus A and B stand for inter-distance, r_{ij} is the distance between i and j electrons, Z_A is a nucleus A charge, r_{Ai} a nucleus A and electron i inter-distance. In the equation above, the first notation represents the kinetic energy operator of the electrons, and second is the nuclei kinetic energy operator, third represents the electron-nuclei attractions kinetic energy operator, fourth notation is the potential energy operator for electron-electron repulsions, and finally, the fifth

notation is the nuclei-nuclei repulsions, potential energy operator. The main disadvantage of the Schrodinger equation is its inability to be used in solving molecular system problems. In this instance, the Born-Oppenheimer approximation is used [9].

3.4 Born-Oppenheimer approximation

Born and Oppenheimer in 1927 posited that nuclear coordinates are the parameters that define molecular geometry and also give the relationship between energy and nuclear coordinates. It is anchored on the fact that the nucleus mass is more than the electron mass, which leads to the faster movement of the latter compared with the former [10, 11]. The opposite charges possessed by both nucleus and electrons also allow for the generation of an attractive force, leading to the rate of movement of both molecules. This approximation equation excludes the nucleus's motion assumes that the movement of electrons and nucleus can be independently treated [10, 11]. Therefore, statically distributed electron contributions within a molecule can be solved as

$$T^{elec} = \left[-\frac{h^2}{8\pi^2 m} \sum_i^{electrons} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \quad \textbf{Eq. 3.4}$$

While Schrodinger equation for electrons in a fixed nucleus is

$$H^{elec} \varphi^{elec}(r, R) = E^{eff}(R) \varphi^{elec}(r, R) \quad \textbf{Eq. 3.5}$$

3.5 Molecular Mechanics

Owing to the greater computational speed and efficiency of molecular mechanics, it is used in calculating thermodynamics properties of biomolecular systems (through molecular dynamics), molecular docking etc. It is also used to analyze molecules that contain thousands and millions of atoms, deducing the structural conformations given by the MD ensemble and all other applications that call for numerous energy valuations [12–15]. The principle guiding molecular mechanics include the following-

- Unlike quantum mechanics, the mathematical model analyses the nucleus' position while neglecting that of electrons' contribution.
- Electrons and nuclei atom-like particles are regarded as spheres
- The bonds between each particle are considered harmonic oscillators
- The spatial distribution of particles and their individual energies determine the interaction
- Springs and classical potential govern interaction

Molecular mechanics is also known as force field calculation. A force field is a mathematical function that correlates the coordinates of specific particles to the molecular systems' overall energy and analyzes the potential energy of the interatomic interactions [16]. Force-field parameters are either determined semi-empirically through QM calculations or experimental data such as infrared, X-ray diffraction, NMR, etc. There are different force fields with their differing complexities used in treating biomolecular systems. Examples are AMBER [17], CHARMM [18], NAMD [19], GROMOS [20], GROMACS [21], DL_POLY [22]. According to the specific type of molecules, each force field must be chosen carefully for its parameterization. Due to the force field variation within a molecule, there must be a corresponding adjustment of the force field parameters [23]. Typically, a force-field expression is like this:

$$U = \sum_{\text{bonds}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\phi - \delta)]$$
$$+ \sum_{\text{improper}} V_{imp} + \sum_{\text{LJ}} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) + \sum_{\text{elec}} \frac{q_i q_j}{r_{ij}},$$

Eq. 3.6

The first four notations are referred to as the contributions of the intramolecular forces to the binding energy. They are bond stretching, angle bending, dihedral, and improper torsions, while the last two notations are Lennard-Jones potential and the Coulombic force, which is also the electrostatic interactions between the non-bonded atoms. Force fields are only limited to analyzing bond interactions. They cannot be used in forming and breaking bonds [24]. In this study, we made use of the AMBER force field in parameterizing the biomolecular systems.

3.6 Molecular Dynamics

Computer simulation serves as a compliment and a bridge to experimental investigations in the sense that we may test a theory by simulation and correlate it with the same model's experimental findings. Simulation allows for the understanding of the interaction between atoms [25]. Simulations can be divided into two distinct classes, and they are molecular dynamics (MD) [26, 27] and Monte Carlo (MC). Other types consist of the hybrid of the two classes mentioned above [28]. MD is preferred to MC because it allows for dynamical analysis and subsequent time-dependent structural perturbations etc. MD also gives room for the evaluation of the thermodynamics and kinetic properties of a biomolecular system (whether bound to a ligand or not). It allows for the manipulation of the factors like temperature, flexibility, and velocity to determine its effect on the time-evolution of interacting particles. Finally, it provides a real-time view of the chemical reaction and biological systems

MD simulation comprises the Newtonian classical equation of motion, which is defined for the atomic system as

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \qquad \textbf{Eq.3.7}$$

In the equation above $r_i(t)$ is the particle position vector, t is the time-evolution, m= mass of the particle, and F_1 is the force acting on the atom.

However, to apply this equation in MD simulation, the individual particle position and velocity making up a biomolecule must be known as this is required, a force-field must be applied for the parameterization of the molecular system, and lastly, there must be specificity in boundary conditions. To perform both quantitative and qualitative analysis on MD simulated systems, trajectories must be generated. For trajectories to be generated, then parameters such as the coordinates, bond-characteristics, acceleration, and potential energy of each atom must be defined. Also, potential energy must be incorporated into the Newtonian equation above. The generated trajectory is then analyzed.

3.7 Molecular Dynamics Post-Analysis

To understand the macromolecules' behavior presented in this thesis, they were MD simulated, and the resulting trajectories generated from the production run were subjected to post dynamical analysis. The MD trajectories are a series of sequential and time-progressive snapshots that give the individual atom's coordinates and velocities. The equilibrium of the systems, binding free energies, conformations, and configuration changes in the 3D structure of the protein are significant results determined by post-MD analysis. Software to be adopted for use in the post-MD analysis should not just be chosen randomly. They must have embedded in the various analytical tools and efficiently process large datasets, and their image resolution must be top-notch.

3.8 Molecular Dynamics Solvent Parameters

Particle-Mesh Ewald (PME) method, which is an efficient computational tool, is best used in parameterizing solvated systems. It is imperative to solvate a system in an explicit solvent as it gives relatively stable trajectories. In the PME method, the potential energy is distinctively separated into Ewald's standard direct and procal sums. The former is determined by cut-offs while the latter by Fast Fourier Transform (FFT). FFT transforms the 3D coloumbic charges to 2D, leading to the reduction of the time and memory for the estimation of PME [29, 30].

3.9 Stability of systems

3.9.1 System convergence

The quality of the post-MD analysis depends on ensuring simulation convergence. Convergence in MD simulation is defined as when conformational ensembles peak after a considerable length of time [31]. It describes the structural stability imbued by bond torsions, angles, and types during protein motion. It is the point at which optimum distance from the reference structure is reached, exhibiting the most energetically stable conformations of our protein-ligand system [32, 33].

3.10 Conformational Analysis

3.10.1 Root Mean Square of Deviation

The root mean square of deviation of atomic position is being used to compare biomolecular structures [34–36]. It has found its use as an index for observing the folding and conformation dynamics of a protein after MD simulation [37–40]. It validates the quality of predicted structures and also experimentally determined structures [41–44]. In relation to its reference structure, it assesses the extent of structural (dis)similarity of a biomolecule brought about during the process of refinement or innate flexibility [45, 46]. RMSD also points out the segment in a protein with more significant fluctuation than its reference structure. Mathematically, RMSD can be represented as

$$\mathbf{RMSD} = \left(\frac{\sum_{\mathbf{N}}(\mathbf{R_i} - \mathbf{R_1^0})^2}{\mathbf{N}}\right)^{\frac{1}{2}} \qquad \textbf{Eq.3.8}$$

N is the total number of atoms in a biomolecular system, i is calculated via the fitting of the least square protocol to the initial conformation and the one obtained after simulation, refinement etc. R_i is the position of the c-alpha atoms. The mean RMSD applies to the receptor, ligand or the ligand complexed to a receptor. It is almost always calculated by averaging the structural deviation observed in the frames in each trajectory.

3.10.2 Radius of Gyration

The radius of gyration (RoG) is defined as the root mean square distance from each atom of the protein to its centroid. It gives insight into how compact or loosely packed the atoms in a biomolecule is [47]. Compactness in this sense evaluates the degree of the accessible surface area of a protein to that of an ideal sphere of the same volume. RoG does not factor in the size of the protein; it is also inversely proportional to the rate of protein folding [48]. Mathematically, it is represented as:

$$\mathbf{r^2gyr} = \frac{(\sum_{i-1}^{n} \mathbf{w_i}\,(\mathbf{r_i} - \mathbf{r^-})^2)}{\sum_{i-1}^{n} \mathbf{w_i}}$$

Eq. 3.9

Here, r_i is the position of i^{th} atom and r is the weight of the center of atom i

3.10.3 Root Mean Square of Fluctuation

RMSF is much more specific than RMSD. RMSD gives the flexibility of the global structure of the protein, while RMSF gives information about the flexibility of each residue making up the protein [49–51]. Furthermore, it provides insight into the heterogeneity and thermal stability of macromolecules [52–54]. RMSF is often obtained through Debye-Waller factors (B factors). B factor quantifies the fluctuation of individual atoms about their mean position. B-factors has various application, few of which are diagnosing errors in the structure of a protein, analysis of active sites and pockets, relating structural mobility with conformation and the extent of disorder in protein structures [55–59]. RMSF is given by the equation below

$$\mathbf{sRMSF} = \frac{(\mathbf{RMSF_i} - \overline{\mathbf{RMSF}})}{\boldsymbol{\sigma}\,(\mathbf{RMSF})}$$

Eq. 3.10

3.10.4 Principal Component Analysis

PCA is a technique used in determining the structural fluctuation within a biomolecular system and the overall motion of the protein backbone [60]. A decomposition process reduces the dataset needed to explain protein dynamics by ordering the motions observed from the largest to the most miniature spatial scale. It linearly transforms the dataset to filter important information via a correlation matrix built from Cartesian coordinates, which is the atomic displacement in a trajectory [61–63]. Essential dynamics is when PCA is applied to protein trajectory to extract important motions from the conformation. In PCA, protein movement is called an eigenvector while the motion's energy contribution is called eigenvalue [60].

3.10.5 Dynamics Cross Correlation Matrix

DCCM depicts graphically time-correlated motions of the residues making up a protein [64]. Visual inspection of the residue time-correlated pattern is principally used in evaluating DCCM [49, 65]. Positively correlated residues tend towards the same direction while that of anticorrelated residues is in the opposite direction. Completely correlated or anti-correlated motion, C(i, j) = 1 or C(i, j) = −1, represent residues which are in the similar phase and period(66). The equation below gives the dynamic cross-correlation matrix:

$$C_{ij} = \frac{< \Delta ri . \Delta rj >}{\left(< \Delta r_i^2 >< \Delta r_j^2 >\right)^{\frac{1}{2}}} \qquad \textbf{Eq. 3.11}$$

C_{ij} is the cross-correlation coefficient, i and j are the i^{th} and j^{th} residue, respectively, and Δri or Δrj stands for displacement vectors of ith and jth residue, respectively.

3.11 Binding Free Energy Calculations

Estimation of binding free energy calculations is crucial in understanding the interaction and the binding mechanism and affinity of ligand-protein complexes. *in silico* analysis of free binding energies of macromolecular complexes can be calculated by two mainstream methods, which are the Molecular Mechanics/Poisson-Boltzmann Surface Area (MMPB-SA) or Molecular Mechanics/Generalized Born Surface Area (MMGB-SA) techniques which combines MM with the continuum solvation of a molecular system at low computational cost [66–68]. In the studies reported in subsequent chapters, the MM-PBSA approach was used. Apart from MM/PBSA and MM/GBSA, other algorithmic approaches in calculating BFE are free energy perturbation estimation, thermodynamic integration, linear interaction energy calculation and, molecular docking calculation [69].

The MM/GBSA and MM/PBSA free energy calculations aids the determination of the inhibitory properties of ligands, has wide application in virtual screening and molecular docking of potential drug candidates [66–68]. The two methods employ the principles of molecular mechanics by using

an implicit solvent model to solve the binding free energy computation. The subset that contribute to the total binding free energy (ΔG_{bind}) is the gas-phase energy (Egas), the solvation free energy (Gsol), and entropy (S).

Mathematically, ΔG_{bind} is given as

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (1)$$

$$\Delta G_{sol} = \Delta G_{pol\ (PB)} + \Delta G_{np} \quad (2)$$

Therefore,

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{pol\ (PB)} + \Delta G_{np} \quad (3)$$

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (4)$$

ΔE_{int} is given by the summation of $\Delta E_{angle,}$ $\Delta E_{bond,}$ and $\Delta E_{torsion.}$ Molecular dynamics simulation was run on the complex only, and a single trajectory was used. This approach minimizes error and noise; therefore, ΔE_{int} was canceled between receptor, ligand, and complex.[70]

$$\Delta E_{MM} = \Delta E_{ele} + \Delta E_{vdW} \quad (5)$$

In the equation presented above, ΔE_{ele} and ΔE_{vdw} represent electrostatic and Vander Waal energy contributions, respectively. ΔE_{MM} is gas-phase energy. The solvation energy contribution is $\Delta G_{sol,}$ which is constituted by polar solvation energy contribution ($\Delta G_{pol\ (PB)}$) and non-polar solvation energy contribution (ΔG_{np}). ΔE_{int} represents the internal energy contribution, and $T\Delta S$, the conformational entropy change.

3.12 Additional Computational Tools Used in this Study

3.12.1 Homology modeling

In target-based drug design, the 3D structure of the protein of interest is essential, and this can be readily downloaded from online database such as protein data bank (PDB). The advances in X-ray crystallography, NMR spectroscopy, and electron microscopy had made the determination of protein structures possible. However, there are several scenarios where some biologically important targets may not have been structurally determined. The structures of these proteins can then be homology modeled. Homology modeled predicts the 3D structure of a protein by using its

protein sequence and the 3D X-ray crystal structures of proteins related to it [71]. Homology modeling is hinged on the fact that the structure of a protein is stringently conserved than its amino acid sequence, and a minor change in its amino acid sequence does not significantly affect its 3D structure [72]. In this study, the homology modeling concept was employed in chapter 5, where the IMPDH of the mouse (UNIPROT ID: P24547) was modeled using *Cricetulus griseus* IMPDH (PDB code IJR1) as a template as it shares 99.22% similarity with IMPDH. The homology modeling was carried out SwissModel algorithm.

3.12.2 Virtual Screening (VS)

Virtual screening is indispensable in the field of drug discovery and development. Virtual screening allows for thousands and millions of small molecules to be screened against a biological target's binding site [73–75] . These molecules can then be filtered using different drug-likeness criteria to a sizeable number that has the potential to be a lead compound [76]. Pharmacophore-based virtual screening was employed in chapter 6, where the active moieties of P131 were selected in building a pharmacophore scaffold. The scaffold was then deposited in the ZINC database to screen for ligands that share similar pharmacophoric moieties as P131. The outputted compounds were thereafter docked in the co-factor site of CpIMPDH. Structure-based virtual screening was also introduced in chapter 7. Therein, the 107,000 natural compounds deposited in the ZINC database were sampled for their complementarity in the co-factor binding site of CpIMPDH, and the top three compounds which had the highest docking scores were selected.

3.13 References

1. Genheden S, Reymer A, Saenz-Méndez P, Eriksson LA (2017) Chapter 1. Computational Chemistry and Molecular Modelling Basics. Royal Society of Chemistry, pp 1–38

2. Aminpour M, Montemagno C, Tuszynski JA (2019) An overview of molecular modeling for drug discovery with specific illustrative examples of applications. Molecules 24:. https://doi.org/10.3390/molecules24091693

3. Grant BJ, Gorfe AA, McCammon JA (2009) Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics.

PLoS Comput Biol 5:e1000325. https://doi.org/10.1371/journal.pcbi.1000325

4. Kleppner D, Jackiw R (2000) One hundred years of quantum physics. Science (80-.). 289:893–898

5. Atkins P, Friedman R MOLECULAR QUANTUM MECHANICS, FOURTH EDITION

6. Boström KJ (2015) Quantum mechanics as a deterministic theory of a continuum of worlds. Quantum Stud Math Found 2:315–347. https://doi.org/10.1007/s40509-015-0046-6

7. Schrödinger E (1926) An undulatory theory of the mechanics of atoms and molecules. Phys Rev 28:1049–1070. https://doi.org/10.1103/PhysRev.28.1049

8. Müller-Kirsten HJW (2012) Introduction to Quantum Mechanics. Introd to Quantum Mech 1–20. https://doi.org/10.1142/8428

9. Introduction to Computational Chemistry, 3rd Edition | Wiley. https://www.wiley.com/en-us/Introduction+to+Computational+Chemistry%2C+3rd+Edition-p-9781118825990. Accessed 26 Mar 2021

10. Born M, Oppenheimer R (1927) Zur Quantentheorie der Molekeln. Ann Phys 389:457–484. https://doi.org/10.1002/andp.19273892002

11. Jecko T (2014) On the mathematical treatment of the born-Oppenheimer approximation. J Math Phys 55:053504. https://doi.org/10.1063/1.4870855

12. Maseras F, Morokuma K (1995) IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. J Comput Chem 16:1170–1179. https://doi.org/10.1002/jcc.540160911

13. Vanommeslaeghe K, Guvench O, MacKerell AD (2014) Molecular Mechanics. Curr Pharm Des 20:3281–3292. https://doi.org/10.2174/13816128113199990600

14. Lin H, Truhlar DG (2007) QM/MM: What have we learned, where are we, and where do we go from here? In: Theoretical Chemistry Accounts. pp 185–199

15. Burke K (2012) Perspective on density functional theory. J Chem Phys 136:150901. https://doi.org/10.1063/1.4704546

16. González MA (2011) Force fields and molecular dynamics simulations. École thématique la Société Française la Neutron 12:169–200. https://doi.org/10.1051/sfn/201112009

17. Case DA (2018) Amber 18. Univ California, San Fr

18. Brooks BR, Brooks CL, Mackerell AD, et al (2009) CHARMM: The biomolecular simulation program. J Comput Chem 30:1545–1614. https://doi.org/10.1002/jcc.21287

19. Phillips JC, Braun R, Wang W, et al (2005) Scalable molecular dynamics with NAMD. J. Comput. Chem. 26:1781–1802

20. Christen M, Hünenberger PH, Bakowies D, et al (2005) The GROMOS software for biomolecular simulation: GROMOS05. J. Comput. Chem. 26:1719–1751

21. Van Der Spoel D, Lindahl E, Hess B, et al (2005) GROMACS: Fast, flexible, and free. J. Comput. Chem. 26:1701–1718

22. Smith W, Yong CW, Rodger PM (2002) DL_POLY: Application to molecular simulation. Mol Simul 28:385–471. https://doi.org/10.1080/08927020290018769

23. Lopes PEM, Guvench O, Mackerell AD (2015) Current status of protein force fields for molecular dynamics simulations. Methods Mol Biol 1215:47–71. https://doi.org/10.1007/978-1-4939-1465-4_3

24. Honarparvar B, Govender T, Maguire GEM, et al (2014) Integrated approach to structure-based enzymatic drug design: Molecular modeling, spectroscopy, and experimental bioactivity. Chem. Rev. 114:493–537

25. (2008) Introduction. In: Computational Chemistry and Molecular Modeling. Springer Berlin Heidelberg, pp 1–15

26. Petruk AA, Defelipe LA, Rodríguez Limardo RG, et al (2013) Molecular dynamics

simulations provide atomistic insight into hydrogen exchange mass spectrometry experiments. J Chem Theory Comput 9:658–669. https://doi.org/10.1021/ct300519v

27. Šponer J, Špačková N (2007) Molecular dynamics simulations and their application to four-stranded DNA. Methods 43:278–290. https://doi.org/10.1016/j.ymeth.2007.02.004

28. Neyts EC, Bogaerts A (2013) Combining molecular dynamics with monte carlo simulations: Implementations and applications. Theor Chem Acc 132:1–12. https://doi.org/10.1007/s00214-012-1320-x

29. Case D, Ben-Shalom I, Brozell S, et al (2018) AMBER 2018. Univ. California, San Fr.

30. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R et al (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688

31. Sawle L, Ghosh K (2016) Convergence of Molecular Dynamics Simulation of Protein Native States: Feasibility vs Self-Consistency Dilemma. J Chem Theory Comput 12:861–869. https://doi.org/10.1021/acs.jctc.5b00999

32. Amadei A, Ceruso MA, Di Nola A (1999) On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations. Proteins Struct Funct Genet 36:419–424. https://doi.org/10.1002/(SICI)1097-0134(19990901)36:4<419::AID-PROT5>3.0.CO;2-U

33. Knapp B, Frantal S, Cibena M, et al (2011) Is an intuitive convergence definition of molecular dynamics simulations solely based on the root mean square deviation possible? J Comput Biol 18:997–1005. https://doi.org/10.1089/cmb.2010.0237

34. McLachlan AD (1972) A mathematical procedure for superimposing atomic coordinates of proteins. Acta Crystallogr Sect A 28:656–657. https://doi.org/10.1107/S0567739472001627

35. Kneller GR (1991) Superposition Of Molecular Structures Using Quaternions. Mol Simul 7:113–119. https://doi.org/10.1080/08927029108022453

36. Duan Y, Kollman PA (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. Science (80-) 282:740–744. https://doi.org/10.1126/science.282.5389.740

37. Daura X, Jaun B, Seebach D, et al (1998) Reversible peptide folding in solution by molecular dynamics simulation. J Mol Biol 280:925–932. https://doi.org/10.1006/jmbi.1998.1885

38. Daura X, Van Gunsteren WF, Mark AE (1999) Folding-unfolding thermodynamics of a β-heptapeptide from equilibrium simulations. Proteins Struct Funct Genet 34:269–280. https://doi.org/10.1002/(SICI)1097-0134(19990215)34:3<269::AID-PROT1>3.0.CO;2-3

39. Zagrovic B, Snow CD, Shirts MR, Pande VS (2002) Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. J Mol Biol 323:927–937. https://doi.org/10.1016/S0022-2836(02)00997-X

40. Zagrovic B, Sorin EJ, Pande V (2001) β-Hairpin folding simulations in atomistic detail using an implicit solvent model. J Mol Biol 313:151–169. https://doi.org/10.1006/jmbi.2001.5033

41. Schueler-Furman O, Wang C, Bradley P, et al (2005) Progress in modeling of protein structures and interactions. Science (80-.). 310:638–642

42. Bowman GR, Pande VS (2009) The roles of entropy and kinetics in structure prediction. PLoS One 4:. https://doi.org/10.1371/journal.pone.0005840

43. Rangwala H, Karypis G (2008) fRMSDPred: Predicting local RMSD between structural fragments using sequence information. Proteins Struct Funct Genet 72:1005–1018. https://doi.org/10.1002/prot.21998

44. Zhang Y (2008) Progress and challenges in protein structure prediction. Curr. Opin. Struct. Biol. 18:342–348

45. Brüschweiler R (2002) Efficient RMSD measures for the comparison of two molecular

ensembles. Proteins Struct Funct Bioinforma 50:26–34. https://doi.org/10.1002/prot.10250

46. Zagrovic B, Van Gunsteren WF (2007) Computational analysis of the mechanism and thermodynamics of inhibition of phosphodiesterase 5A by synthetic ligands. J Chem Theory Comput 3:301–311. https://doi.org/10.1021/ct600322d

47. Lobanov MY, Bogatyreva NS, Galzitskaya O V. (2008) Radius of gyration as an indicator of protein structure compactness. Mol Biol 42:623–628. https://doi.org/10.1134/S0026893308040195

48. Lobanov MY, Bogatyreva NS, Galzitskaya O V. (2008) Radius of gyration as an indicator of protein structure compactness. Mol Biol 42:623–628. https://doi.org/10.1134/S0026893308040195

49. Król M, Roterman I, Piekarska B, et al (2005) Analysis of correlated domain motions in IgG light chain reveals possible mechanisms of immunological signal transduction. Proteins Struct Funct Genet 59:545–554. https://doi.org/10.1002/prot.20434

50. Yin J, Bowen D, Southerland WM (2006) Barnase thermal titration via molecular dynamics simulations: Detection of early denaturation sites. J Mol Graph Model 24:233–243. https://doi.org/10.1016/j.jmgm.2005.08.011

51. Sousa SF, Fernandes PA, Ramos MJ (2009) Molecular dynamics simulations on the critical states of the farnesyltransferase enzyme. Bioorganic Med Chem 17:3369–3378. https://doi.org/10.1016/j.bmc.2009.03.055

52. Reetz MT, Soni P, Fernández L (2009) Knowledge-guided laboratory evolution of protein thermolability. Biotechnol Bioeng 102:1712–1717. https://doi.org/10.1002/bit.22202

53. Parthasarathy S, Murthy MRN (2000) Protein thermal stability: Insights from atomic displacement parameters (B values). Protein Eng 13:9–13. https://doi.org/10.1093/protein/13.1.9

54. Vihinen M (1987) Relationship of protein flexibility to thermostability. Protein Eng 1:477–

480. https://doi.org/10.1093/protein/1.6.477

55. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins - A tool for the selection of peptide antigens. Naturwissenschaften 72:212–213. https://doi.org/10.1007/BF01195768

56. Stroud RM, Fauman EB (1995) Significance of structural changes in proteins: Expected errors in refined protein structures. Protein Sci 4:2392–2404. https://doi.org/10.1002/pro.5560041118

57. Vihinen M, Torkkila E, Riikonen P (1994) Accuracy of protein flexibility predictions. Proteins Struct Funct Bioinforma 19:141–149. https://doi.org/10.1002/prot.340190207

58. Mohan S, Sinha N, Smith-Gill SJ (2003) Modeling the Binding Sites of Anti-Hen Egg White Lysozyme Antibodies HyHEL-8 and HyHEL-26: An Insight into the Molecular Basis of Antibody Cross-Reactivity and Specificity. Biophys J 85:3221–3236. https://doi.org/10.1016/S0006-3495(03)74740-7

59. Yuan Z, Zhao J, Wang ZX (2003) Flexibility analysis of enzyme active sites by crystallographic temperature factors. Protein Eng 16:109–114. https://doi.org/10.1093/proeng/gzg014

60. David CC, Jacobs DJ (2014) Principal component analysis: a method for determining the essential dynamics of proteins. Methods Mol Biol 1084:193–226. https://doi.org/10.1007/978-1-62703-658-0_11

61. Pearson K (1901) LIII. On lines and planes of closest fit to systems of points in space . London, Edinburgh, Dublin Philos Mag J Sci 2:559–572. https://doi.org/10.1080/14786440109462720

62. Abdi H, Williams LJ (2010) Principal component analysis. Wiley Interdiscip Rev Comput Stat 2:433–459. https://doi.org/10.1002/wics.101

63. (2006) Introduction. In: Principal Component Analysis. Springer-Verlag, pp 1–9

64. Arnold GE, Ornstein RL (1997) Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3. Biophys J 73:1147–1159. https://doi.org/10.1016/S0006-3495(97)78147-5

65. Swaminathan S, Beveridge DL, Harte WE (1991) Investigation of Domain Structure in Proteins via Molecular Dynamics Simulation: Application to HIV-1 Protease Dimer. J Am Chem Soc 113:2717–2721. https://doi.org/10.1021/ja00007a054

66. Wang C, Greene D, Xiao L, et al (2018) Recent Developments and Applications of the MMPBSA Method. Front Mol Biosci 4:. https://doi.org/10.3389/fmolb.2017.00087

67. Miller BR, Mcgee TD, Swails JM, et al (2012) MMPBSA . py : An E ffi cient Program for End-State Free Energy Calculations. https://doi.org/10.1021/ct300418h

68. Ylilauri M, Pentikäinen OT (2013) MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. J Chem Inf Model 53:2626–2633. https://doi.org/10.1021/ci4002475

69. Huang K, Luo S, Cong Y, et al (2020) An accurate free energy estimator: Based on MM/PBSA combined with interaction entropy for protein-ligand binding affinity. Nanoscale 12:10737–10750. https://doi.org/10.1039/c9nr10638c

70. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82. https://doi.org/10.1021/ci100275a

71. Rai DK, Rieder E (2012) Homology modeling and analysis of structure predictions of the bovine rhinitis B virus RNA dependent RNA polymerase (RdRp). Int J Mol Sci 13:8998–9013. https://doi.org/10.3390/ijms13078998

72. Fiser A (2010) Template-based protein structure modeling. Methods Mol. Biol. 673:73–94

73. Shoichet BK (2004) Virtual screening of chemical libraries. Nature 432:862–865

74. Lavecchia A, Giovanni C (2013) Virtual Screening Strategies in Drug Discovery: A Critical Review. Curr Med Chem 20:2839–2860. https://doi.org/10.2174/09298673113209990001

75. Cosconati S, Forli S, Perryman AL, et al (2010) Virtual screening with AutoDock: Theory and practice. Expert Opin. Drug Discov. 5:597–607

76. Oprea TI (2002) Virtual screening in lead discovery: A viewpoint. Molecules 7:51–62. https://doi.org/10.3390/70100051

CHAPTER 4

A meta-analysis of *Cryptosporidium* species in humans from southern Africa (2000 - 2020).

First author- Kehinde Foluke OMOLABI[1]

Second author- Paul Olalekan ODENIRAN[2]

Third author- Mahmoud E. SOLIMAN[1*]

*Corresponding author- Mahmoud E.S SOLIMAN[1]

Email address- soliman@ukzn.ac.za

Affiliations

1. Molecular Bio-Computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban, 4001, South Africa.

2. Department of Veterinary Parasitology and Entomology, Faculty of Veterinary Medicine, University of Ibadan, Nigeria.

4.1 Abstract

Introduction- The epidemiology of cryptosporidiosis in southern Africa is largely unknown. The disease is associated with diarrhea and nutritional deficiencies, leading to severe morbidity and mortality among immune-compromised patients. This study aimed to assess the pooled prevalence of *Cryptosporidium* spp. infection among immune-compromised humans in southern Africa over the past 20 years.

Methods- Reports of *Cryptosporidium* spp. infection in humans published between 2000 and 2020 using Google Scholar, PubMed, Ovid Medline, African Journal Online (AJOL), and Web of Science literature database were obtained. Inclusion criteria of sorted articles for *Cryptosporidium* spp. infection were standardized using preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist. A total of 22 eligible studies were sorted for meta-analysis.

Results- Overall prevalence of *Cryptosporidium* spp. infection in southern African countries with reports was 16.8% (*95%CI*: 9.7-25.3). Sub-group analysis showed a pooled prevalence of 25.2, 20.5, and 17.9% among HIV/AIDS patients, children, and diarrhoeic individuals, respectively. Pooled prevalence was highest in South Africa and lowest in Zimbabwe across examined individuals. The pooled prevalence of *Cryptosporidium* spp. infections in diarrhoeic patients was highest in individuals from Botswana (17.6%) which is significantly different (X^2 = 9.337; P = 0.002) from South Africans (12.7%). South African individuals with HIV/AIDS showed the highest pooled prevalence of Cryptosporidium infections than other countries.

Discussion- The high prevalence of *Cryptosporidium* spp. infections among immune-compromised patients in southern Africa showed that the pathogen is of significant importance in this region. Continuous studies on the genetic characterization of *Cryptosporidium* spp. isolates and associated risk factors are needed across southern Africa to identify the predominant subtypes in humans.

Keywords: Cryptosporidiosis, Zoonotic Disease, Prevalence, HIV, Public Health, Protozoa

4.2 Introduction

One of the most important neglected tropical diseases is cryptosporidiosis caused by *Cryptosporidium* spp. Cryptosporidiosis is primarily a water-borne disease associated with fatal diarrhea and is often reported in immune-compromised individuals. A joint Food and Agriculture Organization (FAO)/World Health Organization (WHO) expert committee has ranked *Cryptosporidium* spp. enteropathogen fifth among the 24 most significant foodborne parasites in a global ranking. [1, 2]. Cryptosporidiosis perception among people in southern Africa is low, while the disease is exacerbated by the widespread human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), rural settlements with poor hygiene, and water shortages [3].

The disease is common in children, and several identified risk factors include contact with animals, malnutrition, HIV status, and water-related activities [4, 5]. Indicators of *Cryptosporidium* spp. infection include presence of watery diarrhoea, nutritional defects, elevated levels of lactoferrin and immune system related defects. Infection can be contracted through direct contact with infected humans, zoonotic (animal), and can likewise be foodborne or waterborne [6, 7].

Water left untreated and still used for domestic purposes such as cooking, drinking, swimming, and bathing in many rural African homes mediates the exposure to waterborne *Cryptosporidium* spp. [8]. *Cryptosporidium* spp. has been identified in animals such as poultry, fish, dogs, horses, sheep, cattle etc. [2, 9, 10]. Only a few studies have been directly conducted in southern Africa on the identification of *Cryptosporidium* spp. species in different vertebrate hosts.

Cryptosporidium spp. enteropathogen genotypes identified in humans from Africa include, *C. andersoni*, *C. bovis, C. canis*, *C. cuniculus, C. felis*, *C. hominis*, *C. meleagridis*, *C. muris, C. parvum, C. suis, C. ubiquitum*, *C. viatorum,* and *C. xiaoi* [5]. However, most important species identified in southern Africa were *C. hominis, C. meleagridis* and *C. parvum*. Some of the subtypes of these species are zoonotic and more studies are needed to identify the most prevalent and distributed species in humans from these regions.

To decrease the prevalence of *Cryptosporidium* spp. among HIV patients, the introduction of antiretrovirals have been reported to slightly restore immune function [11, 12]. Besides, HIV protease inhibitors have been suggested to serve as antiparasitic drugs in cases of

cryptosporidiosis. For example, the drugs indinavir, ritonavir, and saquinavir were confirmed to have anti-cryptosporidial effects in experimental studies [13]. However, their efficacy as an anti-cryptosporidial is limited and cannot be substituted as mainstream antiparasitic drugs.

The diagnosis of cryptosporidiosis in this region is mainly based on the morphological identification of *Cryptosporidium* spp. oocysts in faecal samples by microscopy using acid-fast stains or immunofluorescent antibody staining. More sensitive test such as polymerase chain reaction (PCR) has been limited to research and not diagnostic purposes in Africa due to the cost and expertise required.

Studies conducted on *Cryptosporidium* spp. infection from southern African countries have not been adequately synchronized with its associated risk factors to develop public health implications and the disease's distribution rate. Similarly, meta-analyses conducted have focused on individual studies with a regional explanation of its prevalence with many comparisons rather than improving disease awareness and providing data for policymakers.

Therefore, we investigated the prevalence, risk factors, and sub-group analyses of published data on *Cryptosporidium* spp. infection from humans in southern Africa. Distribution and impact of the pathogen on immune-compromised patients based on reporting was also examined in the study.

4.3 Materials and methods

4.3.1 Strategic searching of items

Literature database (PubMed, Ovid Medline, AJOL, Google Scholar, and Web of Science) were searched for published articles on *Cryptosporidium* spp. infection in southern Africa in the English Language from January 2000 to October 2020. The considered articles were those with full-texts. Keywords for searches include, "Cryptosporid"*, "*parvum*", "*hominis*", "species", "humans", "children", "HIV", "patients", "diarrhoea", "intestinal"*, "prevalence", "epidemiology", "South Africa", "Botswana", "Lesotho", "Malawi", "Angola", "Namibia", "Zambia", "Zimbabwe", "Mozambique", "Swaziland" and "Southern Africa". Missing articles were avoided by carefully examining the references of identified articles, while the authors independently did the extraction process to minimise error.

4.3.2 Inclusion and exclusion criteria

An article's inclusion criterion was primarily identified, provided the conducted study that investigated the prevalence of cryptosporidiosis in humans by a diagnostic tool was a cross-sectional type and such detected positive cases in Southern Africa. Exclusion criteria focused on study review, case reports and letters to the editor, animal studies, duplicated manuscripts, and articles with insufficient information.

4.3.3 Data extraction

All the authors independently examined downloaded articles to avoid bias. Articles with irrelevant objectives were removed after reading through the titles, abstracts, and full texts. Variables were generated for each manuscript and were entered in Microsoft Excel spreadsheet. Some of the variables included record information of the authors, total cases examined, the number of positive cases detected, diagnostic techniques (Microscopy, Polymerase Chain Reaction, and IFAT-indirect fluorescent antibody test), the country where the study was done, study area, age group, and gender. Other attributes were those associated with risk factors such as occupation, contact with animals, water use, frequency of use, disease pathologies etc. Cross evaluation among authors was done to resolve all forms of disagreements in computing the data.

4.3.4 Quality assessment

PRISMA checklist was used to standardize the inclusion criteria of all relevant information in the study analysis [14] (Supplementary file). Each article was subjected to a quality assessment scale (1-10) generated in line with the objective of the study (low quality, <5.0; moderate quality, 5.0-7.5; and high quality, 7.5-10.0). Excluded published articles were those with lesser than acceptable qualities (> 5.0) for the analysis.

4.3.5 Statistical analysis

METAXL® (version 3.1) was used to perform the meta-analysis. The quality effects model was used to obtain the pooled prevalence of the study estimate. The overall prevalence of the quality effects model findings was assessed using the validity of the tests. To evaluate the heterogeneity between test results, Cochran's Q test and the inconsistency (I2) indicator were used. The Luis Frya-Kanamori (LFK) index was obtained by plotting the z-score against the double-arcsin

prevalence of the publications examined. This suggests substantial asymmetry when the LFK index reaches ±2, (publication bias) [15]. WINPEPI (version11.65) was used for the Pearson chi-square, while Tukey's post-hoc multiple pair-wise one-way ANOVA comparison test was used to compare disease trends among countries using GraphPad Prism (version 5). Confidence intervals was set at ninety-five percent with minimum and maximum values. A map was constructed with qGIS (version 2.8.10) to show *Cryptosporidium* spp. infection distribution in southern African countries.

4.4 Results

4.4.1 Study characteristics

A total of 431 articles were selected on *Cryptosporidium* spp. infection from five database and reference lists of relevant studies (Figure 4.1). Excluded studies from title screening and abstracts was 61 from the combined searches of 98 studies. Further excluded studies due to duplication were five studies. The inclusion criteria lists were major standards, and 15 studies did not meet the requirements (these were studies outside the range of selected years). Therefore, 22 studies were included in the meta-analysis (Figure 4.1). Examined cases were observed in South Africa (*n*=7), Botswana (*n*=5), Zambia (*n*=4), Malawi (*n*=3), Zimbabwe (*n*=2) and Angola (*n*=1). It was observed that several southern African countries have not published data for human cryptosporidiosis or published in non-visible journals in recent years as observed from the database. In total, 5932 individuals were considered in the study, and the immune-compromised patients include 3517 children, 1280 diarrheic patients, 1073 HIV/AIDS seropositive individuals. and 589 individuals who were immune-competent. Four studies reported patients who were both children and HIV patients, three studies reported patients who were both children and diarrheic, while a study reported patients who were diarrheic and positive for HIV/AIDS.

4.4.2 Cryptosporidium infections in southern Africa

The pooled prevalence estimates of *Cryptosporidium* spp. spp. in humans (*n*=22) with individual studies from southern Africa on human *Cryptosporidium* spp. infections are shown in a forest plot (Figure 4.2). Studies revealed pooled prevalence of 16.8% (*95%CI*: 9.7-25.3); $I^2 = 97.6$; $Q = 869.9$; $df = 21$; $P < 0.0001$. The prevalence of *Cryptosporidium* spp. of each study varied from 2.4 to

65.7% (median=14.5), with substantial heterogeneity among studies (Figure 4.2). The LFK index of 1.17 showed no asymmetry, indicating that there was no bias. Based on available information from countries with cryptosporidiosis reports, the pooled prevalence was highest in South Africa with 21.8% (*95%CI*: 6.1-42.9), and lowest in Zimbabwe 6.6% (*95%CI*: 4.4-9.3) (Table 4.1). Only one study met the inclusion criteria in Angola, with prevalence of 29.9%. Statistical analysis showed that there is significant increase (X^2 = 67.9; P < 0.0001) in the prevalence of immune-compromised patients (20.3%) compared to immune-competent individuals (6.3%). The forest plot of the prevalence of immune-compromised patients was reported (Figure 4.3). A map showing the recent distribution of *Cryptosporidium* spp. infection studies across southern African countries was illustrated in Figure 4.4.

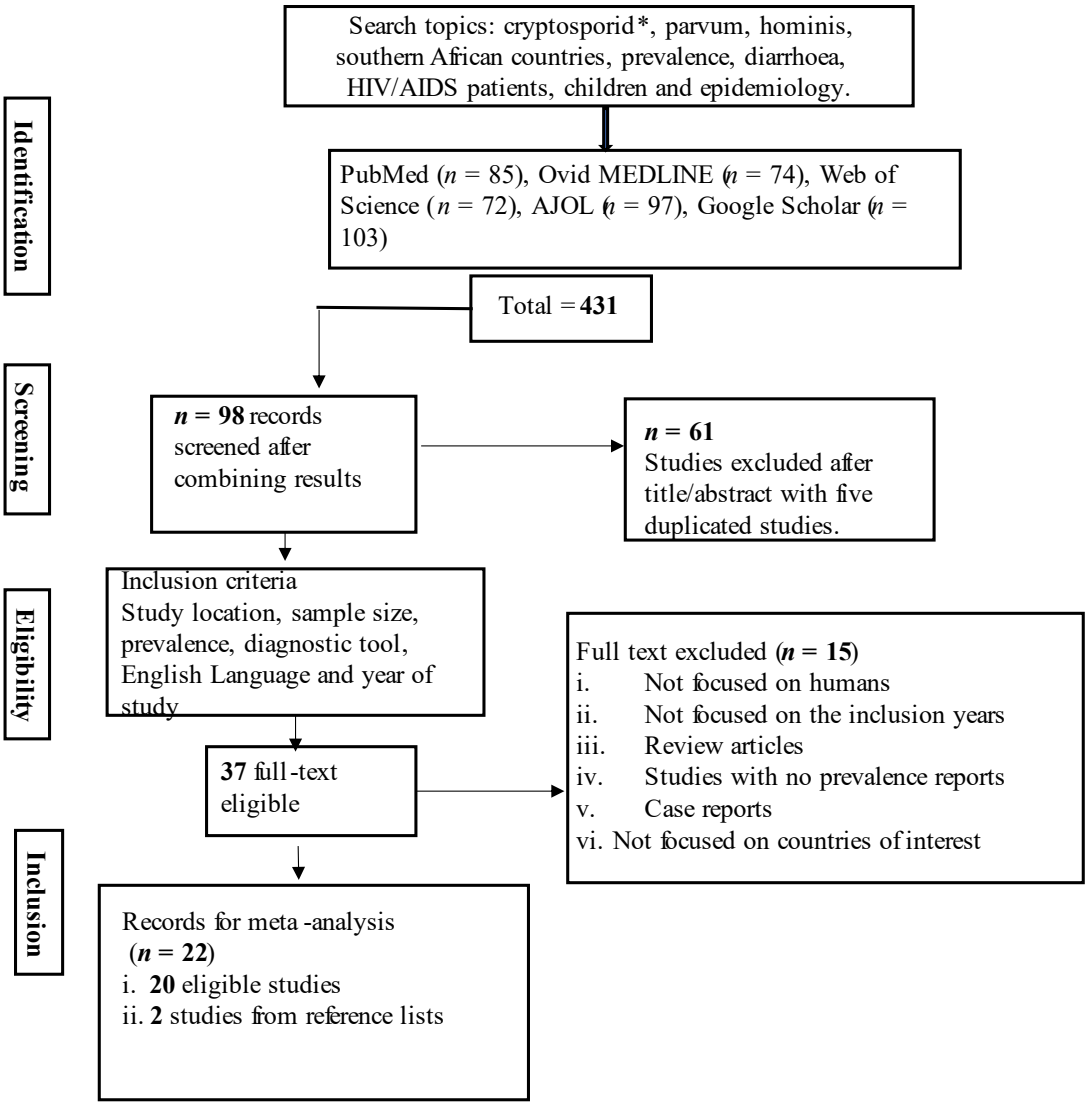


Figure 4.1 Literature database search on *Cryptosporidium* infection of humans in southern Africa.

4.4.3 Pooled prevalence and heterogeneity of Cryptosporidium infections in children

Substantial heterogeneity was observed across studies on *Cryptosporidium* spp. infections in children. The pooled prevalence and heterogeneity variables revealed 20.5% (*95%CI*: 9.6-34.1); I^2 = 98.1; Q = 626.7; P < 0.0001. The LFK (0.02) indicates that there is no asymmetry. Cryptosporidiosis pooled prevalence showed that children in Zambia were most vulnerable with 26.6% (*95%CI*: 13.2-42.5), followed by South African children with 24.1% (*95%CI*: 1.9-56.5), and least in Malawian children 8.5% (*95%CI*: 0.0-100) (Table 4.1). Considering the latter pooled prevalence, the confidence interval showed that the prevalence could rise further depending on the number of studies conducted across the country (Figure 3A).

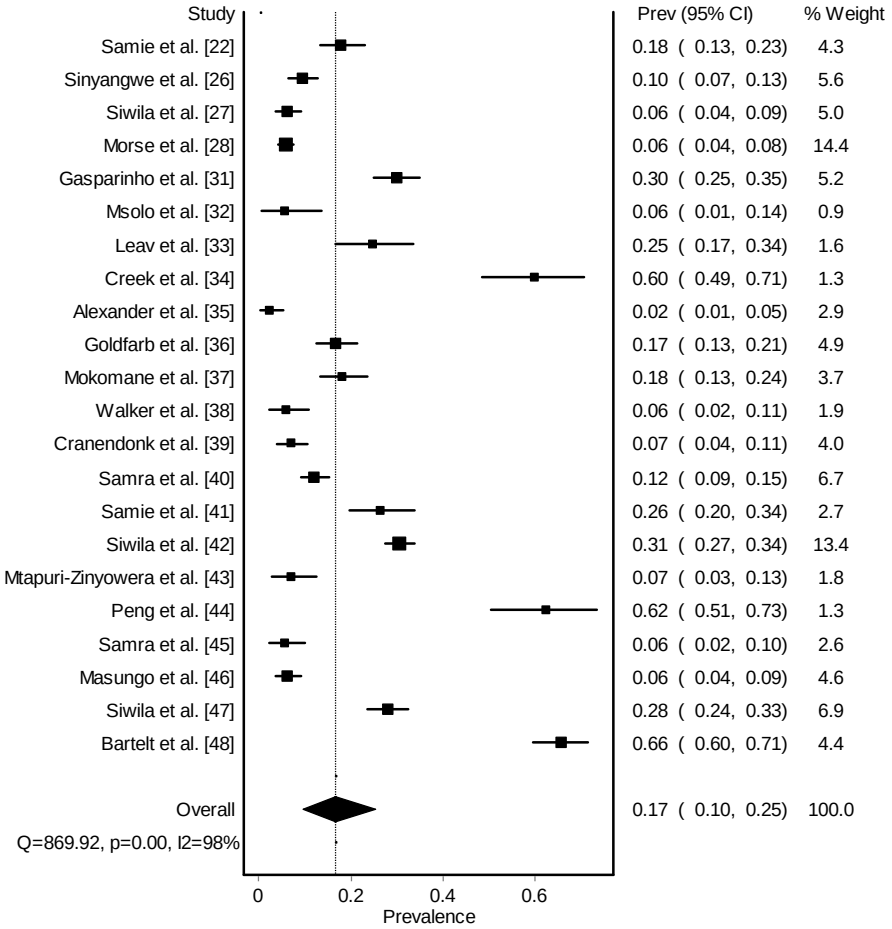


Figure 4.2 Forest plot of *Cryptosporidium* infection among human population in southern Africa between 2000-2020.

Table 4.1: Total report of *Cryptosporidium* infection in humans and associated species in southern Africa between 2000-2020

	Sample size	No positive	Pooled Prev.	95% CI	Cochran's Q	I^2	*df*	$\alpha_{0.05}$	Chi^2, OR/ Anova
Total infection	5932	1121	16.8	9.7-25.3	869.9	97.6	21	<0.0001	
Countries									
South Africa	1385	339	21.8*	6.1-42.6	289.2	97.9	6	<0.0001	$P < 0.0001$
Botswana	849	142	13.3	1.8-31.1	119.3	96.6	4	<0.0001	
Zambia	1804	403	20.3*	7.5-36.8	139.8	97.9	3	<0.0001	
Malawi	1144	109	8.5	0.0-100.0	112.5	99.1	2	<0.0001	
Zimbabwe	413	27	6.6	4.4-9.3	0.1	0.0	1	0.729	
Angola	337	101	-	-	-	-	-	-	
Children with *Cryptosporidium* spp. infection									
Children infection	3959	913	20.5	9.6-34.1	626.7	98.1	12	<0.0001	
Countries									
South Africa	1069	306	24.1*	1.9-56.5	285.9	98.6	4	<0.0001	$P < 0.0001$
Botswana	327	51	15.3	8.0-24.3	2.1	51.8	1	0.1500	
Zambia	1309	362	26.6*	13.2-42.5	42.7	95.3	2	<0.0001	
Malawi	917	93	8.5	0.0-100.0	112.5	99.1	1	<0.0001	
Diarrhoeic patients with *Cryptosporidium* spp. infection									

Diarrhoeic patients	1722	333	17.9	9.8-27.6	126.9	93.7	8	<0.0001	
Countries									
South Africa	592	81	12.7	3.6-25.6	12.1	83.4	2	0.0020	X^2 = 9.337; P = 0.002; OR = 0.63
Botswana	685	138	17.6[#]	3.5-37.9	75.8	96.0	3	<0.0001	
Malawi	108	13	-	-	-	-	-	-	
Angola	337	101	-	-	-	-	-	-	
HIV/AIDS patients with *Cryptosporidium* spp. infection									
HIV/AIDS patients	1073	305	25.2	5.7-51.0	240.5	96.8	4	<0.0001	
Countries									
South Africa	747	274	34.9	10.9-63.3	143.8	97.9	3	<0.0001	
Zambia	326	31	-	-	-	-	-		
Diagnostic technique									
ZN	3676	568	13.5	6.1-23.1	332.7	97.3	9	<0.0001	$F_{2,22}$ P = 0.5646 R^2 = 0.0506
ELISA	1501	413	24.4[#]	7.8-45.8	318.9	98.1	6	<0.0001	
PCR	2486	313	11.0	4.0-20.5	181.1	96.1	7	<0.0001	
Species in humans									
Cryptosporidium spp. species characterised									
C. hominis	186	105	56.5	-	-	-	5	-	
C. parvum	182	58	31.9	-	-	-	4	-	
C. meleagridis	79	4	5.1	-	-	-	2	-	
C. andersoni	50	1	2.0	-	-	-	-	-	

C. parvum/*C. hominis*	50	1	2.0	-	-	-	-	-

Abbreviations: I^2 = level of inconsistency, *df* = degree of freedom, $\alpha_{0.05}$ = level of significance, OR = odd ratio, CI = confidence interval, HIV^+ = human immunodeficiency virus positive, HIV^- = human immunodeficiency virus negative, * = shows significance at $P < 0.05$, # = no significance at $P > 0.05$.

4.4.4 Heterogeneity of Cryptosporidium infections in diarrhoeic patients

A total of 1280 individuals with diarrhea were examined for cryptosporidial infections, of which 279 were reported positive. It was observed that most examined studies (98.3%) with diarrhoeic patients were children. Reported studies were from Botswana, South Africa, Malawi, and Angola. The pooled prevalence and heterogeneity variables of diarrhoeic patients with cryptosporidiosis showed 17.9% (9.8-27.6); $I^2 = 93.7$; $Q = 126.9$; $df = 8$; $P < 0.0001$ as observed from the forest plot (Figure 4.3B). Most studies were retrieved from Botswana, followed by South Africa. The pooled prevalence was highest in Botswana (17.6%) but not significantly different ($X^2 = 0.357$; $P = 0.550$) from South Africa (12.7%) (Table 4.1).

4.4.5 Pooled prevalence of Cryptosporidium infections in HIV/AIDS patients in southern Africa

A total of 305 individuals positive with HIV/AIDS were observed to test positive to *Cryptosporidium* spp. infection out of examined 1073 individuals. The LFK index of 0.59 indicates that there is no asymmetry in the analysed studies. There is substantial heterogeneity with pooled prevalence showing 25.2% (5.7-51.0); $I^2 = 98.3$; $Q = 240.5$; $P < 0.0001$ (Figure 4.3C). Studies were only reported in South Africa and Zambia (Table 4.1). The prevalence in South Africa (34.9%) was significantly higher ($X^2 = 81.024$; $P < 0.0001$) than the study from Zambia (9.5%). The presence of lactoferrin was 59.1% in *Cryptosporidium* spp.-positive patients observed in a study.

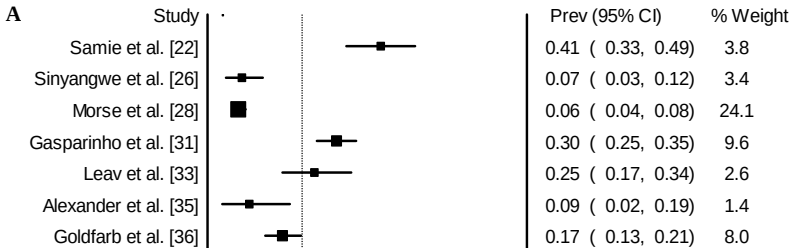


Figure 4.3 Forest plot of *Cryptosporidium* infection among immune-compromised patients. A. *Cryptosporidium* infection in children; B. *Cryptosporidium* infection in diarrhoeic patients; C. *Cryptosporidium* infection in HIV/AIDS patients

4.4.6 Diagnostic techniques

Studies were mostly examined with light microscopy using Ziehl-Neelsen staining techniques. The pooled prevalence based on light microscopy, immunoassay (IFAT, ELISA) and PCR methods was 13.5% (6.1-23.1), 24.4% (7.8-45.8) and 11.0% (4.0-20.5), respectively (Table 4.1). Tukey post-hoc multiple comparison test showed no significant difference (P = 0.5646*;* R^2 = 0.0506) between the three diagnostic techniques. Five countries (Zambia, Malawi, Botswana, South Africa, and Zimbabwe) reported studies using light microscopy and immunoassay, respectively. For the PCR technique, studies were observed in four countries with most reports from South Africa (n = 3), while Malawi and Botswana (n = 2) followed, and least in Zambia (n = 1).

4.4.7 Species variability in southern Africa

A total of eight studies characterized *Cryptosporidium* spp. subtypes with PCR. The characterized *Cryptosporidium* spp. species showed that *C. hominis* was mostly observed with 56.5% distribution, followed by *C. parvum* with 31.9%. *Cryptosporidium* spp. *hominis* sequences analysis with either *HSP70* and *GP60* genes showed the presence of five sub-families (Ia, Ib, Id, Ie and If) across all available studies. The sequence analysis of *GP60* genes of *C. parvum* showed three sub-families (IIb, IIc and IIe) across examined studies. Mixed infection of *C. parvum* and *C. hominis* showed a 2.0% distribution of the examined studies (Table 4.1). Only one sub-family (IId) was identified for *C. meleagridis* (Table 4.2). Of the reviewed studies, *C. meleagridis* showed 5.1% distribution, while *C. andersoni* stands at 2.0%.

Table 4.2: Reported genotypes of *C. hominis*, *C. parvum* and *C. meleagridis* in southern Africa from molecular characterization and sequencing.

Species/ Genotype	Host	Percentage positives	Subfamily	Subtype	References
Total	Country				
C. hominis					
	Malawi	63.6	Ia, Ib, Id, Ie		Peng *et al.,* 2003
		95.3			Morse *et al.,* 2007
		58.1			Gatei *et al.,* 2003
	South Africa	75.0	Ib, Ie	IbA12G3R2; IbA10G2; IeA11G3T3	Samra *et al.,* 2016
		76.0	Ia, Ib, Id, Ie, If	IaA20R3; IaA25G1R3; IaA17R3; IbA9G3; IbA10G1; IdA20; IdA25; IdA26; IdA24; IeA11G3T3b; IfA14G1; IfA12G1	Samra *et al.,* 2013
		75.0	Ia, Ib, Id, Ie		Leav *et al.,* 2002
		20.0			Siwila *et al.,* 2007
		81.8			Samie *et al.,* 2006
	Botswana	41.0			Creek *et al.,* 2010
C. parvum					
	Malawi		IIc, IIe		Peng *et al.,* 2003
	South Africa	20.0	IIb, IIc, IIe	IIbA11; IIcA5G3bb; IIeA12G1	Samra *et al.,* 2013
		25.0	IIc		Leav *et al.,* 2002
		80.0			Siwila *et al.,* 2007
		18.2			Samie *et al.,* 2006
	Botswana	50.0			Creek *et al.,* 2010
C. meleagridis					

Malawi	4.7			Morse *et al.,* 2007
South Africa	4.0	IId	IIIdA4	Samra *et al.,* 2013

4.4.8 Sensitivity test and limitations

The stability and reliability of the analysed results were examined from the sensitivity tests of individual data analysed in the METAXL. The analysis from the funnel plot within the 95% confidence interval and doi plot showed LFK index of 1.17 on the overall analysis, which ruled out significant bias risk of the analysed studies on human cryptosporidiosis in southern Africa. The LFK indexes of sub-group analysis have been reported in each sub-section of the result. In cases of publication bias, there is major asymmetry that could result from few cases examined in sub-group assessment.

4.5 Discussion

Cryptosporidiosis is a disease of the immune-compromised individuals, and it is often endemic in areas with poor social infrastructures such as lack of safe drinking water and sanitation problems. Studies on *Cryptosporidium* spp. infection in southern Africa within the past 20 years have been scanty, despite the burden of cryptosporidiosis and its associated risk factors. Southern Africa is disproportionately affected by cryptosporidiosis. This may be due to the fact that approximately 70% of its population live with HIV/AIDS, 31% of newly infected individuals with HIV, and 34% of individuals dying from AIDS are in the region [3, 16]. South Africa has the largest population living with HIV [17].

The result from this meta-analysis showed that the overall prevalence of 16.8% *Cryptosporidium* spp. infection among examined humans is high in southern Africa. The published articles showed that perception and awareness of the pathogen are low. The high number of immune-compromised individuals could be correlated with the high prevalence of *Cryptosporidium* spp. infection. For instance, the prevalence rate of HIV individuals was estimated between 1.0-26.5% in this region, while adult HIV prevalence exceeded 20% in Lesotho, Swaziland and Botswana. *Cryptosporidium*

spp. infection was highest in South Africa compared to other pooled studies from other countries from this study. This could be associated with the high population of immune-compromised individuals within the country.

Notably, *Cryptosporidium* spp. infection was highest in children less than five years old. Meanwhile, children are disproportionately affected by cryptosporidiosis as observed in this study. The watery diarrhoea accompanied by dehydration and weakness, could lead to death easily. The global estimated 800,000 deaths due to cryptosporidiosis have been reported annually, mostly occurring in sub-Saharan Africa and south Asia [2,18].

The high prevalence of *Cryptosporidium* spp. infection observed in diarrhoeic patients, especially studies from Botswana and South Africa, could be linked to the anthroponotic nature of some subtypes and hygiene. For instance, *Cryptosporidium* spp. oocysts and isolates have been reported to be present in the surface waters of the Vaal Dam, treated effluents, drinking water in South Africa [19], and piped-water in Zambia [20]. The burden of diarrhea is particularly high in children, which could be linked with water-borne cryptosporidiosis. This protozoon is the second cause of severe diarrhea and the leading cause of death in children [18, 21]. The elevated lactoferrin level observed in *Cryptosporidium* spp. positive individuals in a study [22] indicated that inflammation is likely present. However, more studies are needed to correlate elevated lactoferrin and cryptosporidiosis in immune-compromised patients.

Low sensitivity diagnostic techniques are major problems in southern African countries. Even though molecular studies were done, genetic characterization of *Cryptosporidium* spp. species is lacking, which has limited our knowledge of *Cryptosporidium* spp. species and subtypes across the region. The high prevalence observed from studies examined with immunoassay, could be due to circulating IgG antibodies or sensitivity of the rapid diagnostic kits used for diagnosis.

Most characterized species with *GP60 genes* revealed *C. hominis* followed by *C. parvum* across southern Africa [5]. Moreover, *C. hominis* and anthroponotic *C. parvum* subtypes isolated revealed both zoonotic and anthroponotic transmission in southern Africa. Earlier, there have been reports of *C. hominis* subtypes dominance infecting humans in several studies from Africa, regardless of the immune status [23-25]. Although only a study from Zambia reported an animal contact rate of 11.4% [26], the association between human and animal with diarrhea, could be responsible for the

increasing zoonotic danger of *C. parvum* isolates [27], particularly among the immune-compromised patients in southern Africa. Some *C. parvum* subtypes have been reported to be human-adapted subtypes, that could be transmitted from person to person, with origins from humans [28]. The reporting of *C. meleagridis* in this study is not surprising, as the African population has a growing immune-compromised population with a previous report of 21% in these populations prone to infection by this pathogen [25, 29, 30]. Cryptosporidial infection was observed to be highest among HIV/AIDS patients, followed by children and then diarrhoeic individuals. The pooled prevalence across these groups was higher than 20%, which indicates that infection is more likely within the immune-compromised patients.

Recent studies showed that the absence of reports for four countries (Lesotho, Namibia, Mozambique, and Swaziland) could be due to low awareness of the importance of cryptosporidiosis or local publication of *Cryptosporidium* spp. infection reports (Figure 4)

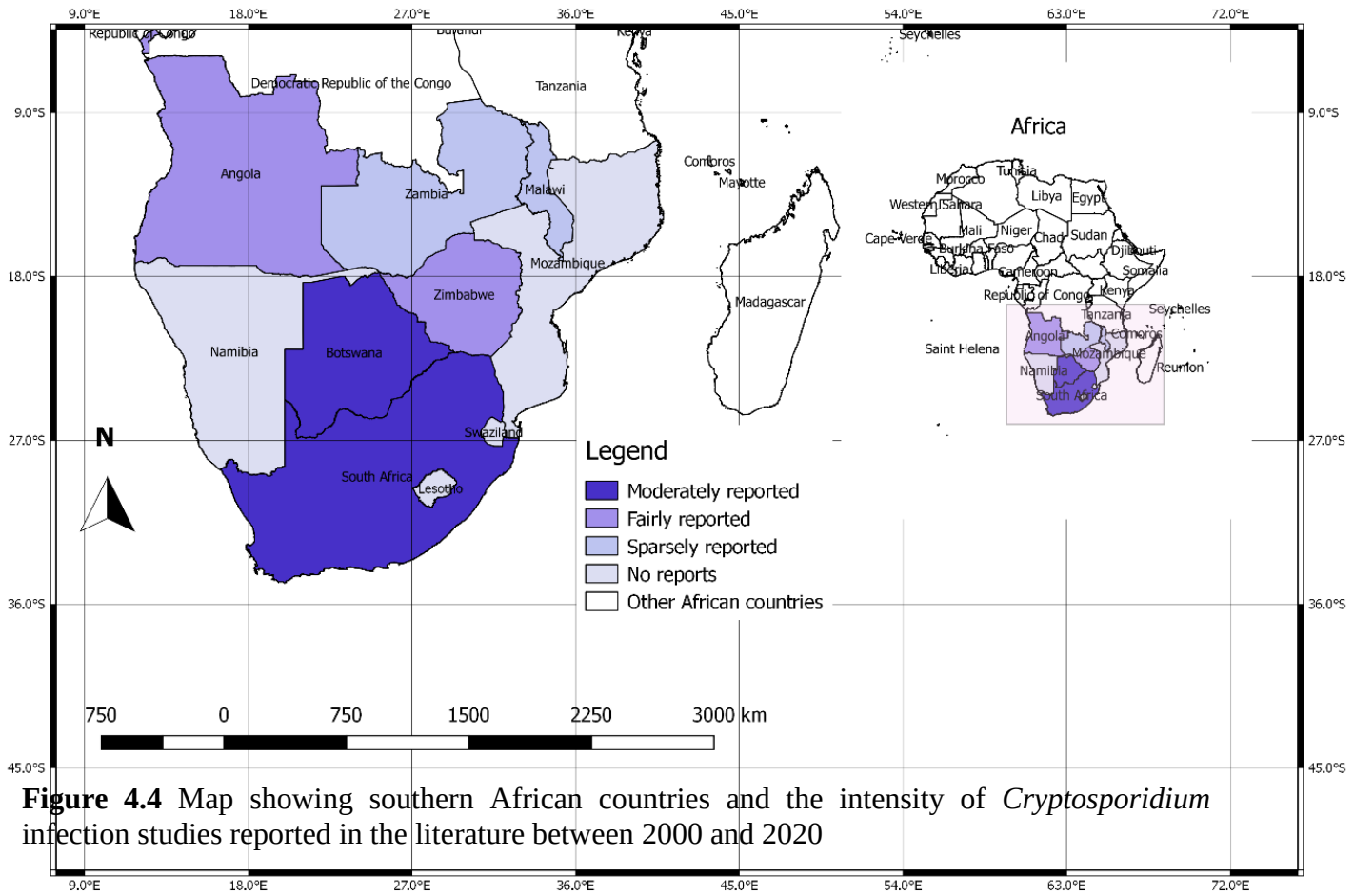


Figure 4.4 Map showing southern African countries and the intensity of *Cryptosporidium* infection studies reported in the literature between 2000 and 2020

4.6 Conclusions

Cryptosporidium spp. infection was highest in HIV/AIDS patients (25.2%), followed by children (20.2%) and diarrheic patients (17.9%). Although, there is no significant difference in the prevalence. These categories of individuals could be at risk of infection if prompt awareness and treatment are neglected. Areas with a high density of livestock and the presence of humans with *Cryptosporidium* spp. infection, especially immune-compromised individuals, could mean that routine screening for opportunistic infections should be prioritized to avoid mortalities and spread of infection. More studies are needed to be conducted in southern Africa, such as correlating the significance of highly active antiretroviral therapy (HAART) positive or negative individuals and evaluating the $CD4^+$ cell counts with *Cryptosporidium* spp. infections. Awareness on the zoonotic and anthroponotic nature of several *Cryptosporidium* spp. subtypes is needed.

Acknowledgements

The authors appreciate the University of KwaZulu Natal, Durban, South Africa, for access to their literature database.

Conflict of interest

The authors declare that there is no conflict of interest.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

All authors agree to publish the article

4.7 References

[1] Delahoy MJ, Omore R, Ayers TL, Schilling KA, Blackstock AJ, Ochieng JB, Moke F, Jaron P, Awuor A, Okonji C et al. (2018) Clinical, environmental, and behavioral characteristics associated with *Cryptosporidium* infection among children with moderate-to-severe diarrhoea in rural western Kenya 2008–2012: the Global Enteric Multicenter Study (GEMS), PLoS Negl Trop Dis 12: e0006640. doi.org/10.1371/journal.pntd.0006640

[2] Odeniran PO, Ademola IO (2019) Epidemiology of *Cryptosporidium* infection in different hosts in Nigeria: A meta-analysis. Parasitology Int 71: 194-206. doi.org/10.1016/j.parint.2019.04.007

[3] Ojuromi TO, Ashafa AOT (2018) Cryptosporidiosis in southern Africa: review of prevalence and molecular epidemiology of a neglected disease. Annals Trop Med Publ Hlth 11(4): 108-118. doi: 10.4103/atmph.atmph_29_17

[4] Gendrel D, Treluyer JM, Richard-Lenoble D (2003) Parasitic diarrhea in normal and malnourished children. Fundam Clin Pharmacol 17(2):189–197. doi: 10.1046/j.1472-8206.2003.00169.x

[5] Squire SA, Ryan U (2017) *Cryptosporidium* and *Giardia* in Africa: current and future challenges. Parasit Vect 10:195. doi:10.1186/s13071-017-2111-y

[6] Xiao L (2010) Molecular epidemiology of cryptosporidiosis: an update. Exp Parasitol 124: 80–89. doi: 10.1016/j.exppara.2009.03.018

[7] Burnet JB, Penny C, Ogorzaly L, Cauchie HM (2014) Spatial and temporal distribution of *Cryptosporidium* and *Giardia* in a drinking water resource: implications for monitoring and risk assessment. Sci Total Environ 472:1023–1035. doi: 10.1016/j.scitotenv.2013.10.083

[8] Centers for Disease Control and Prevention. Domestic Water, Sanitation, and Hygiene Epidemiology, Atlanta, USA.. 2015. http://www.cdc.gov/ncezid/dfwed/waterborne/global.html. Accessed 25 Sept 2016.

[9] Amer S, Zidan S, Feng Y, Adamu H, Li N, Xiao L (2013) Identity and public health potential of *Cryptosporidium* spp. in water buffalo calves in Egypt. Vet Parasitol 191:123–127. doi: 10.1016/j. vetpar.2012.08.015

[10] Bodager JR, Parsons MB, Wright PC, Rasambainarivo F, Roellig D, Xiao L et al (2015) Complex epidemiology and zoonotic potential for *Cryptosporidium suis* in rural Madagascar. Vet Parasitol 207(1):140–143. doi: 10.1016/j.vetpar.2014.11.013

[11] Missaye A, Dagnew M, Alemu A, Alemu A (2013) Prevalence of intestinal parasites and associated risk factors among HIV/AIDS patients with pre-ART and on-ART attending Dessie hospital ART clinic, Northeast Ethiopia. AIDS Res Ther 10(1):7. doi.org/10.1186/1742-6405-10-7.

[12] Kiros H, Nibret E, Munshea A, Kerisew B, Adal M (2015) Prevalence of intestinal protozoan infections among individuals living with HIV/AIDS at Felegehiwot Referral Hospital, Bahir Dar, Ethiopia. Int J Infect Dis 35:80–86. doi: 10.1016/j.ijid.2015.04.012

[13] Mele R, Morales MAG, Tosini F, Pozio E (2003) Indinavir reduces *Cryptosporidium parvum* infection in both *in vitro* and *in vivo* models. Int J Parasitol 33(7):757–64. doi: 10.1016/s0020-7519(03)00093-6.

[14] Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP 2010. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, Int. J. Surg. 8: 336–441. doi: 10.1016/j.ijsu.2010.02.007

[15] Barendregt JJ, Doi SA (2015) MetaXL User Guide Version 3.1, Epigear International Pty Ltd, Queensland. www.epigear.com

[16] WHO (2015) World Health Organisation, Geneva Switzerland. World Health Statistics. http://apps.who.int/iris/bitstream/10665/170250/1/9789240694439_eng.pdf. Accessed 25 Sept 2016.

[17] CIA (2016) Central Intelligence Agency. World Fact Book; 2016. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2156rank.html

[18] Sow SO, Muhsen K, Nasrin D, Blackwelder WC, Wu Y, Farag TH, et al (2016) The burden of *Cryptosporidium* diarrheal disease among children < 24 months of age in moderate/high

mortality regions of sub-Saharan Africa and South Asia, utilizing data from the Global Enteric Multicenter Study (GEMS). PLoS Negl Trop Dis 10(5): e0004729. doi: 10.1371/journal.pntd.0004729

[19] Dungeni M, Momba M (2010) The abundance of *Cryptosporidium* and *Giardia* spp. in treated effluents produced by four wastewater treatment plants in the Gauteng Province of South Africa. Water SA 36: 425–431. doi: 10.4314/wsa.v36i4.58413

[20] Nchito M, Kelly P, Sianongo S, Luo NP, Feldman R, Farthing M, Baboo KS (1998) Cryptosporidiosis in urban Zambian children: an analysis of risk factors. Am J Trop Med Hyg 59:435–437. doi: 10.4269/ajtmh.1998.59.435

[21] Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D, Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL et al. (2013) Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet 382:209–222. doi: 10.1016/S0140-6736(13)60844-2

[22] Samie A, Bessong PO, Obi CL, Sevilleja JE, Stroup S, Houpt E, et al (2006) *Cryptosporidium* species: preliminary descriptions of the prevalence and genotype distribution among school children and hospital patients in the Venda region, Limpopo province, South Africa. Exp Parasitol 114(4): 314–22. doi: 10.1016/j.exppara.2006.04.007

[23] Maikai BV, Umoh JU, Lawal IA, Kudi AC, Ejembi CL, Xiao L (2012) Molecular characterizations of *Cryptosporidium*, *Giardia*, and *Enterocytozoon* in humans in Kaduna State, Nigeria. Exp Parasitol 131(4):452–6. doi.org/10.1016/j.exppara.2012.05.011

[24] Helmy YA, Krücken J, Nöckler K, von Samson-Himmelstjerna G, Zessin KH (2013) Molecular epidemiology of *Cryptosporidium* in livestock animals and humans in the Ismailia province of Egypt. Vet Parasitol 193:15–24. doi: 10.1016/j.vetpar.2012.12.015

[25] Aldeyarbi HM, Abu El-Ezz NMT, Karanis P (2016) *Cryptosporidium* and cryptosporidiosis: the African perspective. Environ Sci Pollut Res 23(14):13811–21. doi: 10.1007/s11356-016-6746-6

[26] Sinyangwe NN, Siwila J, Muma JB, Chola M, Michelo C (2020) Factors associated with *Cryptosporidium* infection among adult HIV positive population in contact with livestock in Namwala district, Zambia. Frontiers in Public Health 8: 74. doi: 10.3389/fpubh.2020.00074

[27] Siwila J, Phiri IGK, Vercruysse J, Goma F, Gabriel S, Claerebout E, Geurden T (2007) Asymptomatic cryptosporidiosis in Zambian dairy farm workers and their household members. Trans R Soc Trop Med Hyg 101:733–734. doi: 10.1016/j.trstmh.2007.01.006

[28] Morse TD, Nichols RAB, Grimason AM, Campbell BM, Tembo KC, Smith HV (2007) Incidence of cryptosporidiosis species in paediatric patients in Malawi. Epidemiol Infesict 135:1307–1315. doi: 10.1017/S0950268806007758

[29] Morgan U, Weber R, Xiao L, Sulaiman I, Thompson RC, Ndiritu W, Lal A, Moore A, Deplazes P (2000) Molecular characterization of *Cryptosporidium* isolates obtained from human immunodeficiency virus-infected individuals living in Switzerland, Kenya, and the United States. J Clin Microbiol 38:1180–1183. doi: 10.1128/jcm.38.3.1180-1183.2000

[30] Ben Abda I, Essid R, Mellouli F, Aoun K, Bejaoui M, Bouratbine A (2011) *Cryptosporidium* infection in patients with major histocompatibility complex class II deficiency syndrome in Tunisia: description of five cases. Arch Pediatr 18:939–944. doi: 10.1016/j.arcped.2011.06.015

[31] Gasparinho C, Mirante MC, Centeno-Lima S, Istrate C, Mayer AC, Tavira L, Nery SV, Brito M (2016) Etiology of diarrhea in children younger than 5 years attending the Bengo General Hospital in Angola. Pediatr Infect Dis J 35(2): e28-e34. doi: 10.1097/INF.0000000000000957

[32] Msolo L, Iweriebor BC, Okoh AI (2020) Rotavirus and *Cryptosporidium* pathogens as etiological proxies of gastroenteritis in some pastoral communities of the Amathole district municipality, Eastern Cape, South Africa. BMC Res Notes 13: 187. https://doi.org/10.1186/s13104-020-05024-2

[33] Leav BA, Mackay MR, Anyanwu A, O' Connor RM, Cevallos AM, Kindra G, et al. (2002) Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among

Cryptosporidium isolates from human immunodeficiency virus-infected children in South Africa. Infect Immun 70: 3881-3890. doi: 10.1128/iai.70.7.3881–3890.2002

[34] Creek TL, Kim A, Lu L, Bowen A, Masunge J, Arvelo W, et al. (2010) Hospitalization and mortality among primarily nonbreastfed children during a large outbreak of diarrhea and malnutrition in Botswana, 2006. J Acquir Immune Defic Syndr 53: 14-19. doi: 10.1097/qai.0b013e3181bdf676

[35] Alexander KA, Herbein J, Zajac A (2012) The occurrence of *Cryptosporidium* and *Giardia* infections among patients reporting diarrheal disease in Chobe District, Botswana. Adv Infect Dis 2: 143-147. doi.org/10.4236/aid.2012.24023

[36] Goldfarb DM, Steenhoff AP, Pernica JM, Chong S, Luinstra K, Mokomane M, et al. (2014) Evaluation of anatomically designed flocked rectal swabs for molecular detection of enteric pathogens in children admitted to hospital with severe gastroenteritis in Botswana. J Clin Microbiol 52: 3922-3927. doi:10.1128/JCM.01894-14

[37] Mokomane M, Kasvosve I, Gaseitsiwe S, Steenhoff AP, Pernica JM, Lechiile K, Luinstra K, Smieja M, Goldfarb DM (2016) A comparison of flocked swabs and traditional swabs, using multiplex real time PCR for detection of common gastroenteritis pathogens in Botswana. Diagn Microbiol Infect Dis 86(2): 142-143. doi: 10.1016/j.diagmicrobio.2016.07.007

[38] Walker CR, Lechiile K, Mokomane M, Steenhoff AP, Arscott-Mills T, Pernica JM, Goldfarb DM (2019) Evaluation of anatomically designed flocked rectal swabs for use with the biofire filmarray gastrointestinal panel for detection of enteric pathogens in children admitted to hospital with severe gastroenteritis. J Clin Microbiol 57: e00962-19. doi.org/10.1128/JCM.00962-19.

[39] Cranendonk RJ, Kodde CJ, Chipeta D, Zijlstra EE, Sluiters JF (2003) *Cryptosporidium parvum* and *Isospora belli* infections among patients with and without diarrhoea. East Afr Med J 80: 398-401. PMID: 14601779

[40] Samra NA, Thompson PN, Jori F, Frean J, Poonsamy B, du Plessis D, et al. (2013) Genetic characterization of *Cryptosporidium* spp. in diarrhoeic children from four provinces in

South Africa. Zoonoses Public Health 60: 154-159. doi: 10.1111/j.1863-2378.2012.01507.x

[41] Samie A, Makuwa S, Mtshali S, Potgieter N, Thekisoe O, Mbati P, et al. (2014) Parasitic infection among HIV/AIDS patients at bela-bela clinic, Limpopo province, South Africa with special reference to *Cryptosporidium*. Southeast Asian J Trop Med Public Health 45: 783-795. PMID: 25427345

[42] Siwila J, Phiri IG, Enemark HL, Nchito M, Olsen A (2011) Seasonal prevalence and incidence of *Cryptosporidium* spp. and *Giardia duodenalis* and associated diarrhoea in children attending pre-school in Kafue, Zambia. Trans R Soc Trop Med Hyg 105: 102-108. doi: 10.1016/j.trstmh.2010.10.004

[43] Mtapuri-Zinyowera S, Ruhanya V, Midzi N, Berejena C, Chin'ombe N, Nziramasanga P, et al. (2014) Human parasitic protozoa in drinking water sources in rural Zimbabwe and their link to HIV infection. Germs 4: 86-91. doi: 10.11599/germs.2014.1061

[44] Peng MM, Meshnick SR, Cunliffe NA, Thindwa BD, Hart CA, Broadhead RL, et al. (2003) Molecular epidemiology of cryptosporidiosis in children in Malawi. J Eukaryot Microbiol 50 (Suppl) :557-559. doi: 10.1111/j.1550-7408.2003.tb00628.x

[45] Samra NA, Jori F, Caccio SM, Frean J, Poonsamy B, Thompson PN (2016) *Cryptosporidium* genotypes in children and calves living at the wildlife or livestock interface of the Kruger National Park, South Africa. Onderste J Vet Res 83(1), a1024. http://dx.doi.org/10.4102/ojvr.v83i1.1024

[46] Masungo P, Dube T, Makaka C (2010) A survey of the diversity of human enteric protoctistan parasites and the associated risk factors in urban Zvishavane, Zimbabwe. Agric Biol J N Am 1(5): 985-991. doi:10.5251/abjna.2010.1.5.985.991

[47] Siwila J, Phiri IGK, Enemark HL, Nchito M, Olsen A (2010) Intestinal helminths and protozoa in children in pre-schools in Kafue district, Zambia. Trans R Soc Trp Med Hyg 104: 122-128. doi: 10.1016/j.trstmh.2009.07.024

[48] Bartelt LA, Sevilleja JE, Barrett LJ, Warren CA, Guerrant RL, Bessong PO, et al. (2013) High anti-*Cryptosporidium parvum* IgG seroprevalence in HIV-infected adults in Limpopo, South Africa. Am J Trop Med Hyg 89: 531-534. doi: 10.4269/ajtmh.12-0550.

CHAPTER 5

Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity – A Structural, Dynamical and Mechanistic Stance.

Kehinde F. Omolabi[a], Clement Agoni[a], Fisayo A. Olotu[a] and Mahmoud E.S. Soliman[a*]

[a]Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Lab webpage: http://soliman.ukzn.ac.za/

5.1 Abstract

Cryptosporidiosis accounts for a surge in infant (< 5 years) mortality and morbidity. To date, several drug discovery efforts have been put in place to develop effective therapeutic options against the causative parasite. Based on a recent report, P131 spares Inosine Monophosphate Dehydrogenase (IMPDH) in a eukaryotic model (mouse IMPDH - *m*IMPDH) while binding selectively to the NAD^+ site in *Cryptosprodium parvum*-*Cp*IMPDH. However, no structural detail exists on the underlining mechanisms of P131-*Cp*IMPDH selectivity to date. To this effect, we investigate the selective inhibitory dynamics of P131 in *Cp*IMPDH relative to *m*IMPDH via molecular biocomputation methods.

Pairwise sequence alignment revealed prominent variations at the NAD^+ binding regions of both proteins that accounted for disparate P131 binding activities. The influence of these variations was further revealed by the MM/PBSA energy estimations coupled with per-residue energy decomposition which monitored the systematic binding of the compound. Furthermore, relative high-affinity interactions occurred at the *Cp*IMPDH NAD^+ site which were majorly mediated by SER22, VAL24, PRO26, SER354, GLY357 and TYR358 located on chain D. These residues are unique to the parasite IMPDH form and not in the eukaryotic protein, highlighting variations that account for preferential P131 binding. Molecular insights provided herein corroborate previous experimental reports and further underpin the basis of *Cp*IMPDH inhibitor selectivity.

Findings from this study could present attractive prospects towards the design of novel anticryptosporidials with improved selectivity and binding affinity against parasitic targets.

Keywords: Cryptosporidiosis; Inosine Monophosphate Dehydrogenase; P131; target selectivity; Molecular dynamics simulation; structural mechanisms; Per residue energy decomposition; sequence alignment

5.2 Introduction

Cryptosporidium is gaining increasing recognition as the most significant protozoan parasite causing diarrhea in animals and humans [1]. Presently, it is reputed to be the principal cause of mortality in children less than five years old, claiming 800,000 lives annually and close to three million cases of detected infection [2]. The bulk of the burden of this infection falls in sub-Saharan Africa and South Asia [3–6]. A joint expert committee from the Food and Agricultural Organization (FAO) and World Health Organization (WHO), undertook a worldwide ranking which placed *Cryptosporidium* as fifth among the twenty-four most important foodborne parasites [7, 8]

The infection is self-limiting in immunocompetent individuals and typically would resolve within two weeks [9]. This is converse to the outcome in immunodeficient individuals (Human Immunodeficiency Virus /Acquired Immunodeficiency Syndrome (HIV/AIDS), organ transplant patients, etc., where infection can cause severe and chronic diarrhea, wasting and most times death [10]. Symptoms are most times aggravated when HIV/AIDS is left untreated, especially in low-resource countries where antiretroviral drugs are not easily accessible and affordable [11].

The causative agent of human cryptosporidiosis is either Cryptosporidium *hominis* or Cryptosporidium *parvum*. [12, 13] The latter has been found to infect several mammals and therefore responsible for most zoonotic infections [7]. Some risk factors include contaminated water facilities and foods, unhygienic environments, and malnutrition [14]. South Africa is not spared from this menace as *Cryptosporidium* species are frequently isolated from her surface waters used for domestic purposes [15]. The circumstance poses more problems as HIV/AIDS is endemic in South Africa, with the prevalence rates among the highest in the world [16].

There is a shortage of therapeutic agents employed in the treatment of cryptosporidiosis. Nitazoxanide remains the only Food and Drug Administration (FDA)-approved drug [17]. As it stands, nitazoxanide has an inexplicit mechanism of action, and it is not especially effective in immunocompromised individuals [18]. Nitazoxanide administration is also punctuated with adverse side effects which include hives, decreased liver function, respiratory problems etc [19]. Apart from nitazoxanide, several drugs though not FDA approved, have been deployed in the treatment of cryptosporidiosis. However, they also have limited efficacy as they failed in

controlled trials in AIDS patients [20, 21]. Examples include CD40 agonist antibody [22], paromomycin [23, 24], spiramycin, azithromycin and bovine anti-*Cryptosporidium* immunoglobulin [20, 25, 26], rifamycin [27], rifaximin [28, 29]. Better responses have been documented when nitazoxanide, azithromycin and/or paromomycin are combined [30–32]. In the wake of the ineffectiveness of the drugs, as mentioned above, the discovery of novel anticryptosporidial drugs is pressing.

Inosine monophosphate dehydrogenase (IMPDH) is a crucial enzyme in *Cryptosporidium* and has been a target for some antibiotic discovery [33–36]. It is responsible for the synthesis of guanine nucleotides. It catalyzes the initial and rate-limiting step of the oxidation of inosine 5'-monophosphate (IMP) to xanthosine 5'-monophosphate (XMP) and simultaneously reducing NAD^+ to NADH [37]. Its inhibition results in the diminution of guanine nucleotide reservoir and consequently, cell death [38] . It is worthy of note that *Cryptosporidium* though a eukaryote has a prokaryotic IMPDH, which was made possible by a horizontal gene transfer from a bacterium [9, 39]. Prokaryotic and eukaryotic IMPDHs are quite diverge in both their structural and kinetic properties, and these account for the selectivity in action of anticryptosporidial agents [40].

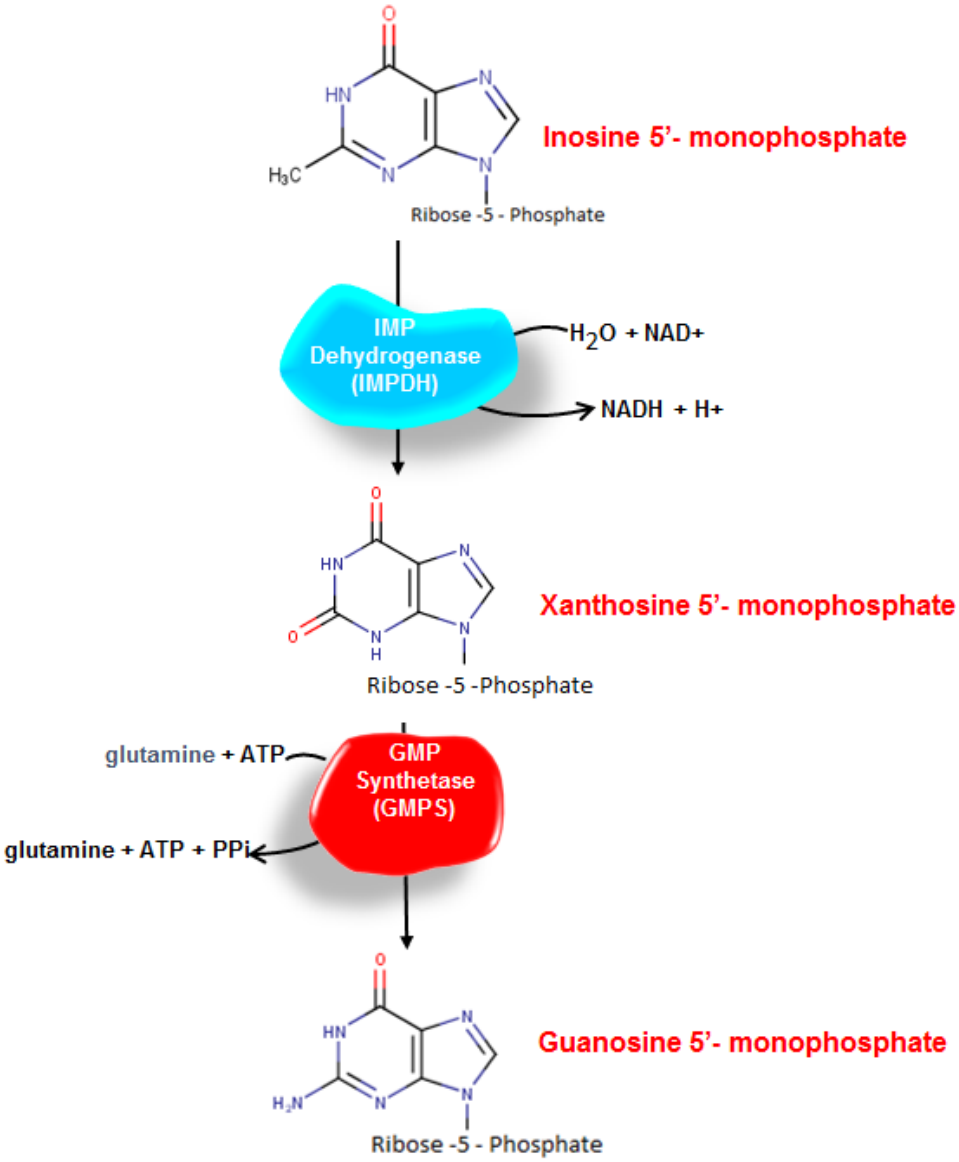


Figure 5.1 Reaction Pathway showing the conversion of Inosine Monophosphate (IMP) to Guanine monophosphate (GMP). IMP dehydrogenase (IMPDH) catalyzes the rate-limiting step in the conversion to GMP [41]

IMPDH are homotetramers with a D4 square symmetry [42]. It is made up of an IMP-binding site, a co-factor site (NAD^+ binding site), and a mobile flap [43]. The four IMP and co-factor binding sites are situated close to the subunit interfaces [44]. The IMP site is confined within a monomer, and it is highly conserved in the IMPDH of all organisms [44]. On the other hand, the NAD+ binding site sequence is diverged.

The NAD^+ binding site has been consistently targeted for the design of IMPDH inhibitors [33, 44, 45]. Three subdomains make up the NAD^+-binding site, namely nicotinamide riboside-binding site (N-subdomain), the pyrophosphate-binding site (P-subdomain); and the adenosine-binding site (A-subdomain) [45]. The N-subdomain lies in the same monomer as IMP, and it is also strongly conserved, being the site of chemical transformation [43]. On the contrary, the adenosine subsite is highly divergent in different organisms [42]. In eukaryotic organism, it is in the same monomer as IMP [46] while in the prokaryote, it interacts with a pocket in the adjacent monomer [42].

A *Cryptosporidium* drug discovery program targeting IMPDH was undertaken by Hedstrom *et al.* research group, which led to the synthesis and production of a battery of *Cryptosporidium* inhibitors [9, 47–56]. Of all these inhibitors, in a murine model, P131 at a single dose demonstrated an equivalent activity when administered at 250mg/kg bodyweight in comparison to the control group, which was treated with paromomycin at 2000mg/kg body weight. Following a thrice-daily administration of both drugs at the same concentration stated earlier, P131 elicited a superior parasiticidal activity when compared to paromomycin. This was followed by a crystallographic determination of the structure of *Cp*IMPDH with P131 bound at the NAD^+ binding site [42].

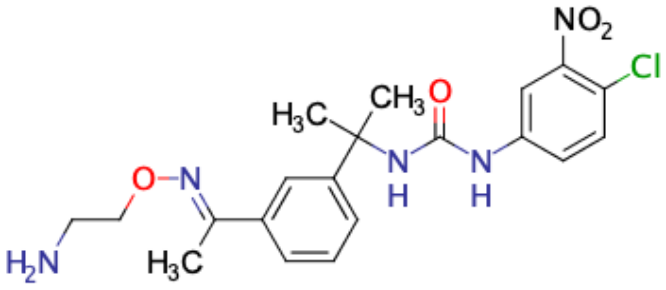


Figure 5.2 2D structure of P131, a novel *Cryptosporidium* inhibitor that elicited a superior parasiticidal activity when compared to paromomycin

Although the selectivity of anticryptosporidial drugs for *Cryptosproridium parvum* (*Cp*) IMPDH has been established [34, 43], the molecular mechanism of non-selectivity towards the hosts' IMPDH unaffected is unclear. Herein, we attempted to explain from bioinformatics and molecular bio-computation point of view, the molecular mechanisms associated with the selectivity of P131 towards *Cp*IMPDH. We further compare the structural and conformational dynamics of P131 when bound to the NAD^+ site of *Cp*IMPDH and mouse IMPDH (*m*IMPDH). Findings from this

study could present an attractive prospect in informing the design of new and potential inhibitors that possess improved binding, pharmacokinetics, and pharmacodynamics properties and selectivity for parasitic targets, that are able to overcome the problems of drug resistance.

5.3 Computational Methodology

5.3.1 Sequence Mapping and Alignment

There was a need to determine the disparity among the binding site residues of *Cryptosporidium parvum* IMPDH (*Cp*IMPDH) and Mouse IMPDH (*m*IMPDH) as regards P131. The divergence in these residues may unravel the basis for the differential drug activity. Clustal Omega was used in performing sequence mapping and alignment. The well-suited match for the selected sequences is calculated and presented in a form in which their identities, similarities, and differences can be seen. [57]. The FASTA sequences of the IMPDH of the two different hosts were retrieved from UNIPROT. *Cp*IMPDH (ID:Q5CPK7) [58] and *m*IMPDH2 (ID: P24547) [59]. All these sequences were fed into the Clustal Omega webserver, and the output was retrieved. Residues that make up the P131 binding sites were then investigated for their variations.

5.3.2 Systems Retrieval, Preparation and Molecular Dynamics (MD) Simulations

Two systems were constituted (*m*IMPDH and *Cp*IMPDH) as investigated in the experimental paper [49]. X-ray crystal structure of the catalytic domain of *Cp*IMPDH co-crystallized with inhibitor P131, was obtained from the RSCB Protein Data Bank (Entry code: 4RV8) [42]. The protein (4RV8) retrieved was a tetramer however, we only made use of a dimer (Chain A and D) to minimize computational cost and reflect experimental study [42]. The dimer used had 652 residues in total. The crystal structure of *m*IMPDH is not available therefore we retrieved the FASTA sequence from UNIPROT (ID: P24547) [59] for homology modeling using the SwissModel algorithm [60]. Moreover, the crystal structure of *Cricetulus griseus* IMPDH in complex with Mycophenolic Acid (PDB entry: 1JR1.A) [61] which has a 99.22% identity with *m*IMPDH was used as the template for modeling. The modelled structure was then validated using RAMPAGE, PROCHECK and VERIFY3D webservers [62–64].

Using the Clustal Omega Sequence Alignment, *Cp*IMPDH and *m*IMPDH have 35.1% sequence similarity. It is important to note that *m*IMPDH shares 98.83% sequence similarity to human

IMPDH (*h*IMPDH). The dissimilarities in both sequences *h*IMPDH and *m*IMPDH are in Ile192, Asp215, Gln265, Asp292 and Asn296 in *h*IMPDH which are replaced by Val192, Asn215, Leu265, Glu292 and Ser296 in *m*IMPDH. The two proteins also share 100% sequence identity in the co-factor binding site. This could explain the choice of selection of murine subject as a representative eukaryotic model in the report by Gorla *et al* [49]. Although, the *Cg*IMPDH on which the *m*IMPDH was built shared a higher similarity with *h*IMPDH (99.6%) than *m*IMPDH did (98.83%), we made use of the putative structure of *m*IMPDH rather than *Cg*IMPDH to keep in line with the experimental model used in the investigating the *in vivo* efficacy of P131 as reported by Gorla *et al* [49].

In *Cp*IMPDH, the binding of P131 has been established to span both adenosine and nicotinamide subdomains of the co-factor (NAD^+) binding site in IMPDH. Having a prokaryotic antecedent, the binding of P131 at the adenosine subsite interacts with chain D of the IMPDH [42]. The target protein structures (*Cp*IMPDH and *m*IMPDH) were prepared using the Graphical User Interface (GUI) of UCSF Chimera [65], which involves the removal of ions, crystal waters, and non-standard residues. Missing residues were added using MODELLER [66]. The co-crystallized P131 in *Cp*IMPDH (4RV8) was retained and its binding orientation to chains A and D was used as the starting structure for MD simulation. However, ligand P131 retrieved from 4RV8 was docked to *m*IMPDH using Autodock Vina [67], which has been widely used over other docking algorithms due to its high scoring power [68]. The co-factor binding site of the *m*IMPDH was defined using the coordinates of co-crystallized P131 in the NAD^+ binding site of *Cp*IMPDH achieved by the superimposition of *Cp*IMPDH and *m*IMPDH. The grid-box was defined as centre (X=2.6487, Y=21.9974, Z=77.3995) and size (X=11.1647, Y=9.4625, Z=7.3236). The best docked pose (most negative score) was aligned with the co-crystallized P131 in *Cp*IMPDH to obtain similar binding orientation for P131 in *m*IMPDH, which was also used as the starting MD structure [65]. The proteins were prepared further by Molegro Molecular Viewer software prior to MD simulation.

5.3.3 Molecular Dynamics simulation

This was performed using standard simulation protocol which has been extensively employed in previous studies [69–71]. Each system was further subdivided into apo (IMPDH bound to IMP in its active site) and complexes which comprised of apo bound by P131 in the co-factor site. Afterward, these were set up for MD simulations according to previously reported protocols [72–

74]. MD simulation was performed using Graphical Processor Unit (GPU) version of the Particle Mesh Ewald Molecular Dynamics (PMEMD) engine in AMBER18 suite coupled with integrated modules [75]. FF14SB forcefield was used in defining protein parameters. P131 parameterization was done using the ANTECHAMBER module, which generated atomic partial charges (Gasteiger - gaff) through the bcc charge scheme [76]. Sequel to this is the generation of topology and parameter files for complexes using the LEAP module, which neutralized the complexes by adding counter ions at a constant pH (cpH) and solvated them in a 10 Å TIP3P water box. Partial minimization was executed in 2500 steps, using a 500 kcal/mol Å restraint potential followed by full minimization for 5000 steps with no energy restraints. The systems were heated for 50 ps from 0 to 300 k in an NVT canonical ensemble using a Langevin thermostat [77] and a harmonic restraint of 5 kcal/mol $Å^2$. The systems were then equilibrated at 300 k for 1000 ps without energy restraints while the Berendsen barostat was used to maintain atmospheric pressure at 1 bar [78]. The MD production run was carried out for 280 ns with the SHAKE algorithm used in constricting all atomic hydrogen bonds [79]. The trajectories obtained were analyzed by the integrated CPPTRAJ and PTRAJ modules [80]. Origin data analytical tool was used in creating the needed plots [81]. 3D visualization of the structures and corresponding analyses were carried out on the GUI of UCSF Chimera.

5.3.4 Post Dynamic Analysis

5.3.4.1 Thermodynamics Calculations

This calculation was used to probe the selective binding of P131 to *Cp*IMPDH and *m*IMPDH. The estimation of free binding energy is vital as it gives insight into the binding affinity and stability of ligands in a complex. Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method, was used in the estimation of the binding free energy. MM/PBSA is a reliable analytical tool widely used for evaluating the interaction of ligands with biological macromolecules [82–85]. Mathematically, binding free energy is depicted by the following equation;

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{inhibitor}) \quad (1)$$

$$\Delta G_{bind} = \Delta G_{gas} + \Delta G_{sol} - T\Delta S \quad (2)$$

$$\Delta G_{gas} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (3)$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{np,sol} \quad (4)$$

$$\Delta G_{np,sol} = \gamma SASA + \beta \quad (5)$$

From the above equation, ΔG_{gas} depicts the gas-phase energy while the internal energy is represented as ΔE_{int}. In the same vein, the coulomb and van der Waals energies are represented as ΔE_{ele} and ΔE_{vdw} respectively. In addition, ΔG_{sol} depicts the free energy of solvation while the polar solvation contribution is represented as ΔG_{PB}. On the other hand, $\Delta G_{np,sol}$ depicts the non-polar contribution and is determined by the linear relationship between the surface tension proportionality constant (γ), solvent accessible surface area, SASA ($Å^2$), and a constant β. The surface tension proportionality constant (γ) is given as 0.0072 kcal/(mol·$Å^2$) and β as 0.92 kcal/mol.

which is estimated by the solvent accessible surface area (SASA) that is determined by using a water probe of radius 1.4Å with a surface tension constant, γ of 0.0072 kcal/(mol·$Å^2$). Per residue decomposition analysis was also carried out to obtain the different energies each binding site residues contributed to ligand affinity and stabilization.

5.4 Result and Discussion

5.4.1 Structural validation of the putative model of *m*IMPDH

The modeled structure of *m*IMPDH retrieved from SwissModel webserver was validated and cross-validated on three online platforms- RAMPAGE, PROCHECK and VERIFY3D. From Ramchandran plot, 85.6% of the total residues fell in the favored region, 12.6% of the residues were in the allowed region, 0.4% was found in the generally allowed region while 1.4% of the total residues were found in the outlier region. This result improved the confidence in the modeled structure. We cross-validated with ProSA and the overall quality of modeled *Tc*HK had a Z-score of -7.43. According to Suppl Figure 5.1, the Z-score is within the range found for other experimentally determined proteins of similar size. Verify3D estimated 80.52% of the residues to have averaged 3D-1D score >= 0.2 score which corroborated the good quality of the modeled protein. The graphical representations of the quality of *Tc*HK on VERIFY3D and PROSA webservers are represented in Suppl Figure 5.1.

5.4.2 Relative sequence analysis and its effect on P131 selectivity

The basis of selectivity has been the subject of numerous researches with the intent to underpin the targetability of small-molecule inhibitors of *Cp*IMPDH. For instance, previous studies by Gollapalli *et al*, Gorla *et al* and MacPherson *et al*, have established that the selectivity of *Cp*IMPDH inhibitors is based on the extension of the inhibitor binding across the subunit interface into a pocket in the adjacent monomer.[34, 42, 48, 56] . These findings therefore serve as the premise for this study.

Pairwise sequence alignment between *Cp*IMPDH and *m*IMPDH revealed variations in their amino acid sequences. Zoning in on the co-factor (NAD^+) binding regions of both proteins, we further identified variations among residues involved in direct contact or interaction with P131 upon binding. As observed, most variations unique to *Cp*IMPDH occurred at chain D which were, Ser22, Val24, Leu25, Pro26, Ser354, Gly357, Tyr358, Lys73 while some occurred at the main chain (chain A) which were Ala165, His166, Asn171, Lys200, Ile213, Val305, Met326, Pro328, Glu329. Corresponding residues in *m*IMPDH are shown in Figure 5.3.

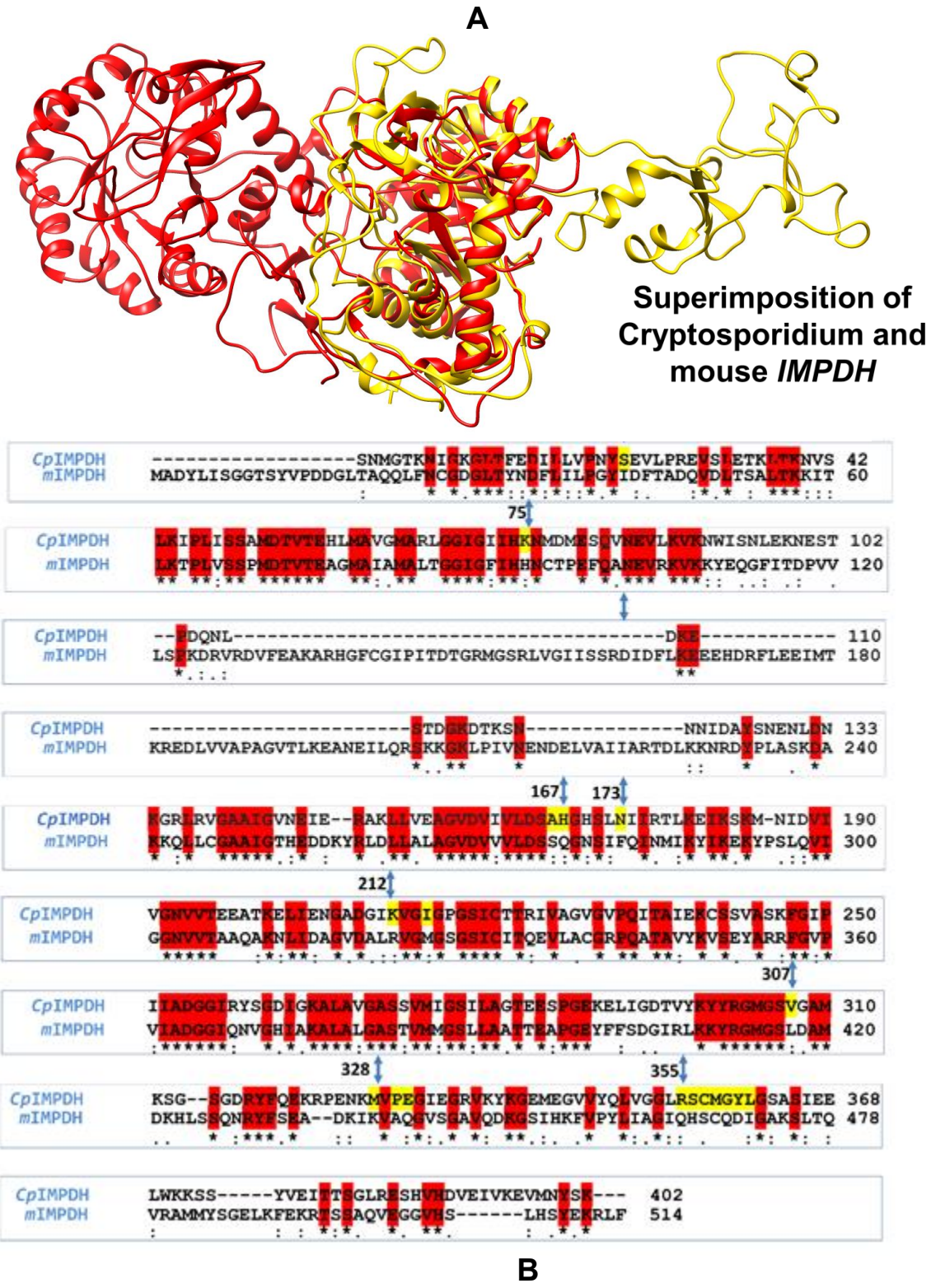


Figure 5.3 (**A**) Superimposed structures of the of *Cp*IMPDH and *m*IMPDH (colored red and gold respectively). (**B**) ClustalW sequence alignment of *Cp*IMPDH and *m*IMPDH. Red box highlight residues

that are identical between both organisms. Binding site amino acid sequences that interacted directly with P131 but were peculiar to only *Cp*IMPDH are highlighted in yellow The position of non-identical sequences at both *Cp*IMPDH and *m*IMPDH are without annotations, while the position of highly conserved sequences are denoted by (*****). Position of less conserved sequences are denotated by (**:**) and (**.**) respectively. These annotations are defined by the ClustalW prediction method.

We further explored the differential binding modes of P131 at the cofactor (NAD^+) site of *m*IMPDH and *Cp*IMPDH as it involved varying residues among both proteins as earlier mentioned. As shown in figure 5.4, using averaged snapshot generated across the simulation period, the molecular interaction profiles of P131-*Cp*IMPDH relative to *m*IMPDH were dissimilar. These differences in interaction modes could in turn influence the binding affinity and stability of P131 at the NAD^+ regions of both proteins.

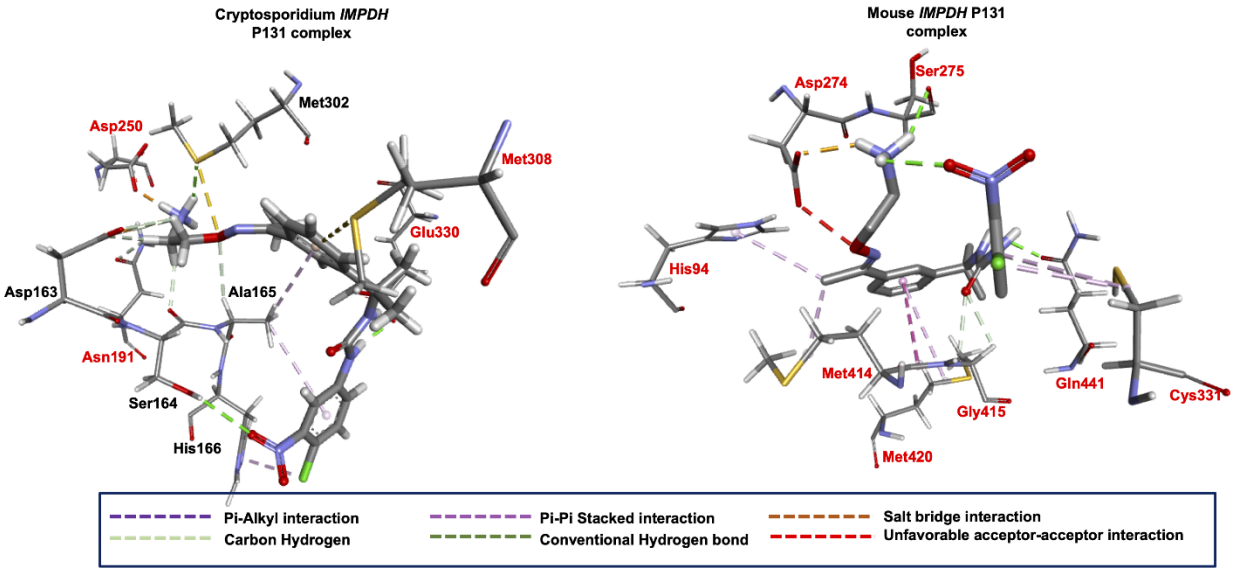


Figure 5.4 Interaction analyses showing differences in the types and nature of interactions mediated by P131 at the NAD^+ sites of *Cp*IMPDH and *m*IMPDH. Unique residues involved in P131 binding at the NAD^+ site of *Cp*IMPDH are indicated in black text.

More specifically, unique *Cp*IMPDH residues such as ALA165, HIS166, GLU329, PRO26' TYR358' ('denotes chain D) facilitated strong hydrogen and salt bridge interactions with P131, which were not observed in *m*IMPDH. These high-affinity and favorable interactions observed in *Cp*IMPDH could explain the preferential binding of P131 over *m*IMPDH. Also, unfavorable O---O interactions seen in the *m*IMPDH-P131 complex could further reduce the binding affinity of the compound at the NAD^+ site. Thus, this particular sequence variation, with accompanying distinct

interactions, could explain the basis of selectivity and favorable binding of P131 towards *Cp*IMPDH. Taken together, these findings also suggest that the interactions of P131 with residues of chain D of the *Cp*IMPDH could be crucial to the selectivity of P131 due to the sequence variation observed on this chain as shown. These finding lend credence to previous reports which posited that interaction of the adenosine subsite with adjacent monomer in prokaryotic IMPDH and the absence of this in eukaryotic IMPDH could underpin basis of selectivity of IMPDH inhibitors [34, 42, 48, 56].

5.4.3 Peculiar Sequence Variation Favors P131 binding to CpIMPDH

The MM/PBSA method was used to estimate the binding energies (***ΔG***) of P131 in *Cp*IMPDH, and *m*IMPDH. As presented in Table 1, P131 had a total ***ΔG*** of -37.0 kcal/mol when bound to *Cp*IMPDH and -17.0 kcal/mol in *m*IMPDH. This notable energy difference could suggest that P131 was bound more favorably to *Cp*IMPDH which could involve more high-affinity interactions, particularly with residues unique to the NAD^+ site of *Cp*IMPDH as earlier indicated in figure 5.4. To quantify the roles of these residues to the binding and stability of P131 in *Cp*IMPDH, we decomposed the overall energy and measured the per-residue contributions compared to *m*IMPDH.

Table 5.1 MM/PBSA-based binding free energy profile of P131 towards *Cp*IMPDH and *m*IMPDH

Complexes	**Energy components (kcal/mol)**				
	ΔE_{vdw}	**ΔE_{ele}**	**ΔG_{gas}**	**ΔG_{sol}**	**ΔG_{bind}**
***Cp*IMPDH**	-49.82 ±0.08	-61.69 ± 0.32	-111.51 ± 0.31	74.5±0.28	-37.00±0.11
***m*IMPDH**	-21.70 ±0.48	-28.55 ± 0.66	-50.26 ± 1.09	33.25±0.73	-17.01±0.38

ΔE_{ele} = electrostatic energy; **ΔE_{vdW}** = van der Waals energy; **ΔG_{bind}** = total binding free energy; **ΔG_{sol}** = solvation free energy **ΔG_{gas}** = gas phase free energy.

Per-residue energy plots are presented in Figure 5.5, and accordingly, residues with energies > -1 kcal/mol were considered favorable. In *Cp*IMPDH and their respective non-conserved

counterparts in *m*IMPDH, identified residues include *Cp*GLU329 (-3.26 kcal/mol/ *m*GLN441 (-3.46 kcal/mol), *Cp*ALA165 (-2.89 kcal/mol)/*m*SER276 (-2.63 kcal/mol), *Cp*HIS166 (-1.11 kcal/mol)/*m*GLN277 (-0.49 kcal/mol). The conserved residual energy contributions in *Cp*IMPDH/*m*IMPDH are MET302 (-1.82 kcal/mol)/MET414 (-0.6 kcal/mol), GLY303 (-1.63 kcal/mol)/GLY417 (-0.49 kcal/mol). TYR358 (-2.69 kcal/mol) and PRO26 (-1.67 kcal/mol) are present in the chain D of *Cp*IMPDH and have no corresponding residues in *m*IMPDH as the P131 binding is confined to one chain in the eukaryotes.

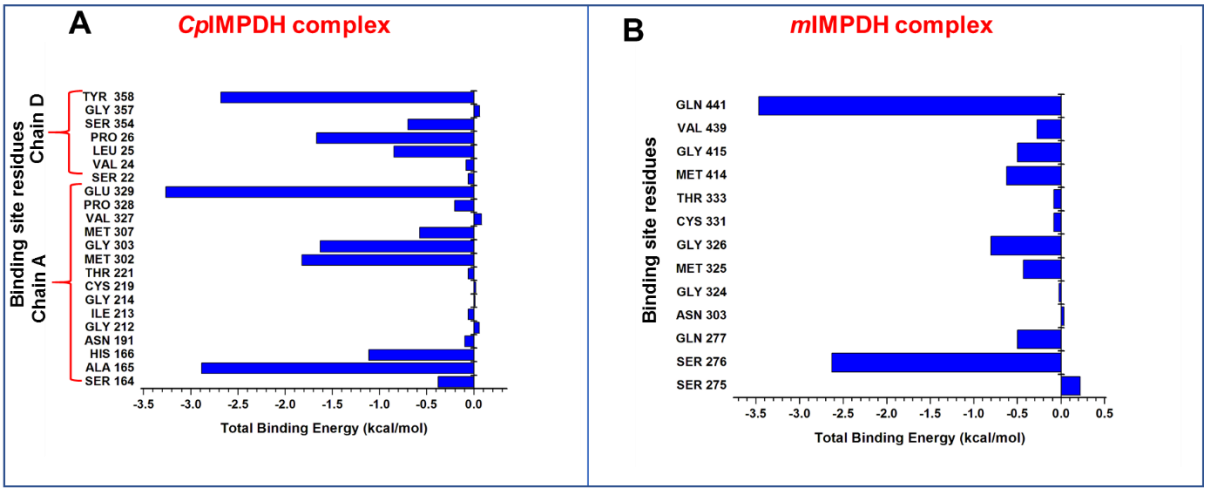


Figure 5.5 A) Per-residue energy decomposition analysis of P131 binding site residues in *Cp*IMPDH. **B**) Per-residue energy decomposition analysis of P131 binding site residues in *m*IMPDH. Both figures represent the energies contributed by the binding sites residues to the total binding free energy of P131 in *Cp*IMPDH and *m*IMPDH.

SER276 and GLN441 were the only residues in *m*IMPDH that contributed above -1 kcal/mol to the binding of P131. This was different in *Cp*IMPDH where we had seven amino acid residues contributing above -1 kcal/mol to the binding of P131. This disparity could have facilitated the higher binding free energy of P131 in *Cp*IMPDH compared to the weaker binding free energy of P131 in *m*IMPDH. Interestingly, almost all the residues on chain D of *Cp*IMPDH that interacted with P131 contributed significantly towards the binding, as shown in Figure 5.5 (LEU25 = -0.85 kcal/mol, PRO26 = -1.67 kcal/mol, SER354 = -0.7 kcal/mol and TYR358 = -2.69 kcal/mol). This further suggests they could be critical to P131 selectivity in *Cp*IMPDH.

5.4.4 Peculiar P131 binding within parasite IMPDH active site favors selectivity

The conformational dynamics and orientations of P131 upon binding to either *Cp*IMPDH, and *m*IMPDH could be crucial in its preferential binding towards *Cp*IMPDH and the resultant therapeutic effect. This is because the varying orientations and the mobility of P131 could somewhat influence the motion of P131 into deep regions of the binding pocket of the enzyme to elicit its therapeutic activity. We, therefore, analyzed the differential binding modes of P131 at the binding site of *Cp*IMPDH over the simulation period, as shown in Figure 5.6. The first representative snapshots (P131-*m*IMPDH and *Cp*IMPDH) were selected at 100ns, indicative of the starting post-equilibrated timeframes. Subsequent snapshots were selected at 220ns and 280ns representing intermediate and final post-equilibrated timeframes across both proteins. Likewise, associated binding pocket amino acids that interacted with P131 over the MD simulation were also explored. It was observed that in *m*IMPDH, a relatively fewer number of residues engaged in interactions with P131 over the simulation period as shown Figure 5.6A.

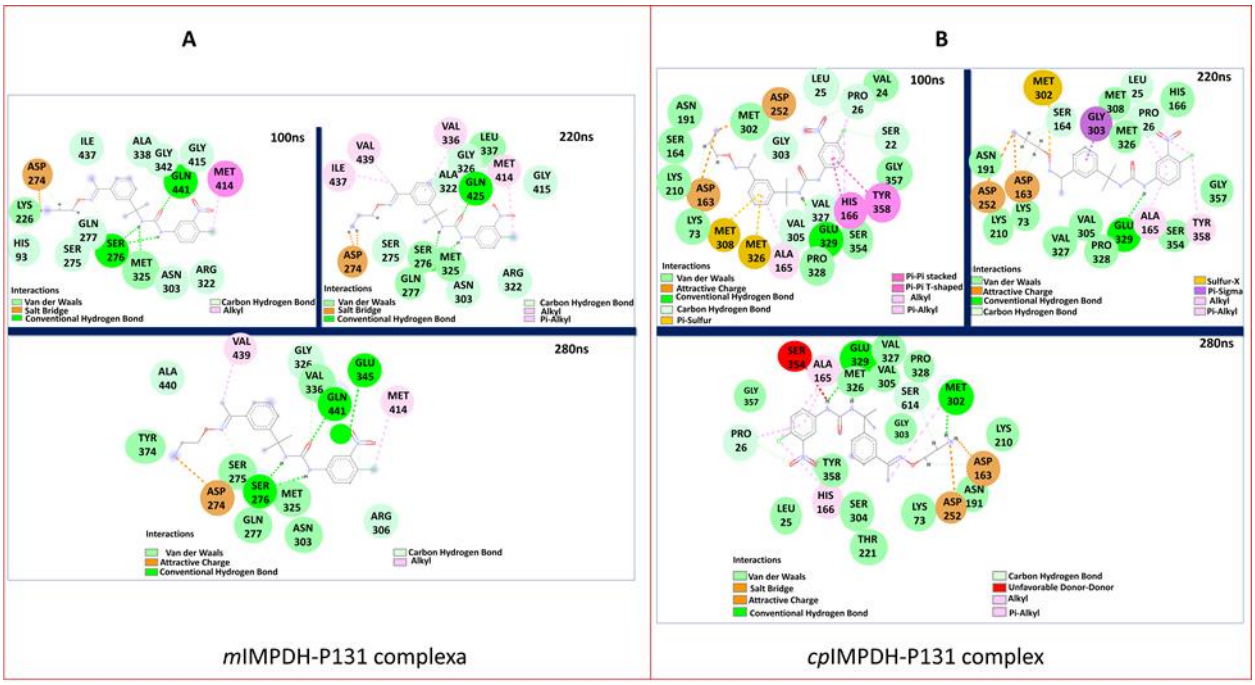


Figure 5.6 Binding interaction dynamics of P131 at the NAD^+ sites of *m*IMPDH and *Cp*IMPDH. A). Time-based analyses of P131 interactions with constituent residues of *m*IMPDH NAD^+ sites at 100ns, 220ns and 280ns. B). Time-based analyses of the interaction of P131 with constituent residues of *Cp*IMPDH NAD^+ site at 100ns, 220ns and 280ns. As shown, binding site residues involved in P131 binding are more in *Cp*IMPDH complex across the simulation time-frames compared to *m*IMPDH. Legends representing interaction types are also shown.

On the contrary, there was a consistently higher number of residues interacting with P131 in *Cp*IMPDH as the simulation proceeded, which could have favored the selective binding of P131 towards *Cp*IMPDH as shown in Figure 5.6B. Prominent and recurring residues towards the binding of P131 during the simulation period include ASP163, SER164, ASP252, MET302. Again, these peculiar residues were unique to only *Cp*IMPDH, suggesting their essential contributions to P131 stability, possibly its eventual selectivity. As shown in Figure 6.6B, the peculiar orientations of PI31 and the unique interacting residues likely defined the motion of P131 into the deep regions of *Cp*IMPDH hydrophobic pocket unlike in *m*IMPDH

The binding of a small molecule inhibitor can induce a significant conformational change in a protein structure, which could consequently influence its known function[86–89]. As such, the selective binding of P131 towards *Cp*IMPDH could be affected by peculiar structural dynamics and conformational changes on *Cp*IMPDH. Using known computational techniques, root-mean-square deviation (RMSD), Solvent accessible surface area (SASA), root-mean-square fluctuation (RMSF) and radius of gyration (RoG), we explore the varying structural dynamics associated with P131 binding to both *Cp*IMPDH and *m*IMPDH. RMSD calculations allow for the estimation of the structural deviation and stability of a IMPDH structures [90–93]. From the RMSD plot (Figure 5.7A and B), both protein systems had high deviations at starting time-frames of the MD run but attained stability at about 250ns in both systems. The average RMSD for *Cp*IMPDH-P131 complex and *m*IMPDH-P131 complex was found to be 2.84Å and 2.92Å, respectively, as shown in Figure 5.7. The RMSD plot suggests that the binding of PI31 stabilizes the *Cp*IMPDH structure and leads to a few conformational changes relative to a more unstable structure in the mouse IMPDH (Figure 5.7). The initial jump in average RMSD for both simulated models could be attributed to the initial orientation of P131 in the binding pocket of both *Cp*IMPDH and *m*IMPDH.

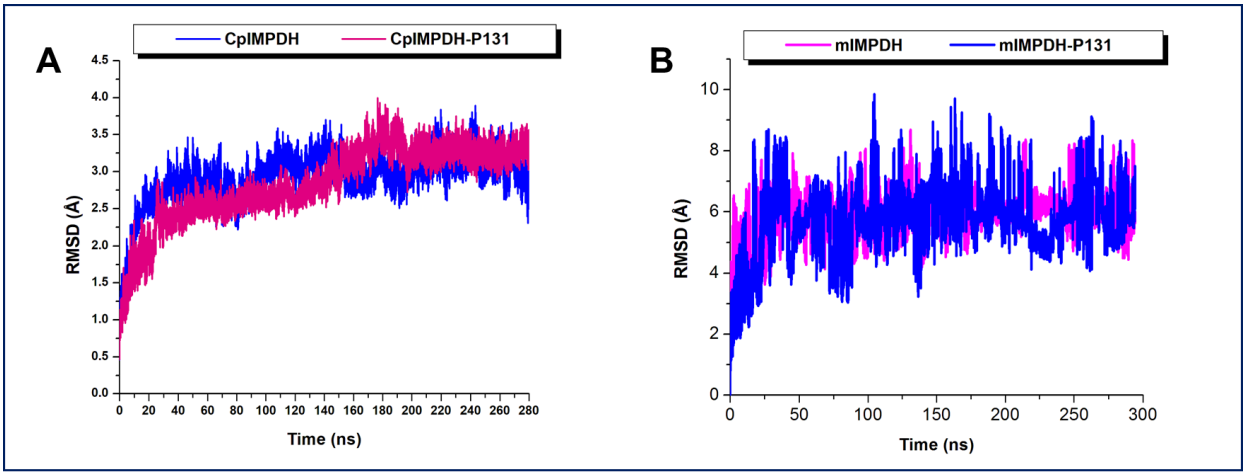


Figure 5.7 A) Comparative RMSD plots of C-a atoms of bound and unbound conformations of *Cp*IMPDH over the simulation period. B) Comparative RMSD plots of C-a atoms of bound and unbound conformations of *m*IMPDH over the simulation period.

Solvent accessible surface area (SASA) of the bound enzymes were also calculated to ascertain how both enzymes interact with surrounding solvents [94, 95]. SASA gives information about the mobility of the residues and side reorientation as it alternates between the hydrophilic and hydrophobic phases [96, 97]. The repositioning of the side chains from the hydrophilic to the hydrophobic phase signifies protein folding while the converse depicts protein unfolding. The binding of P131 to *Cp*IMPDH increased the SASA values minimally when compared to the apo (Figure 5.8). The mean SASA values were 25630A^2 and 24986A^2. In the *m*IMPDH model, the binding of p131 did not significantly impact the reorientation and alternation of the protein side chains.

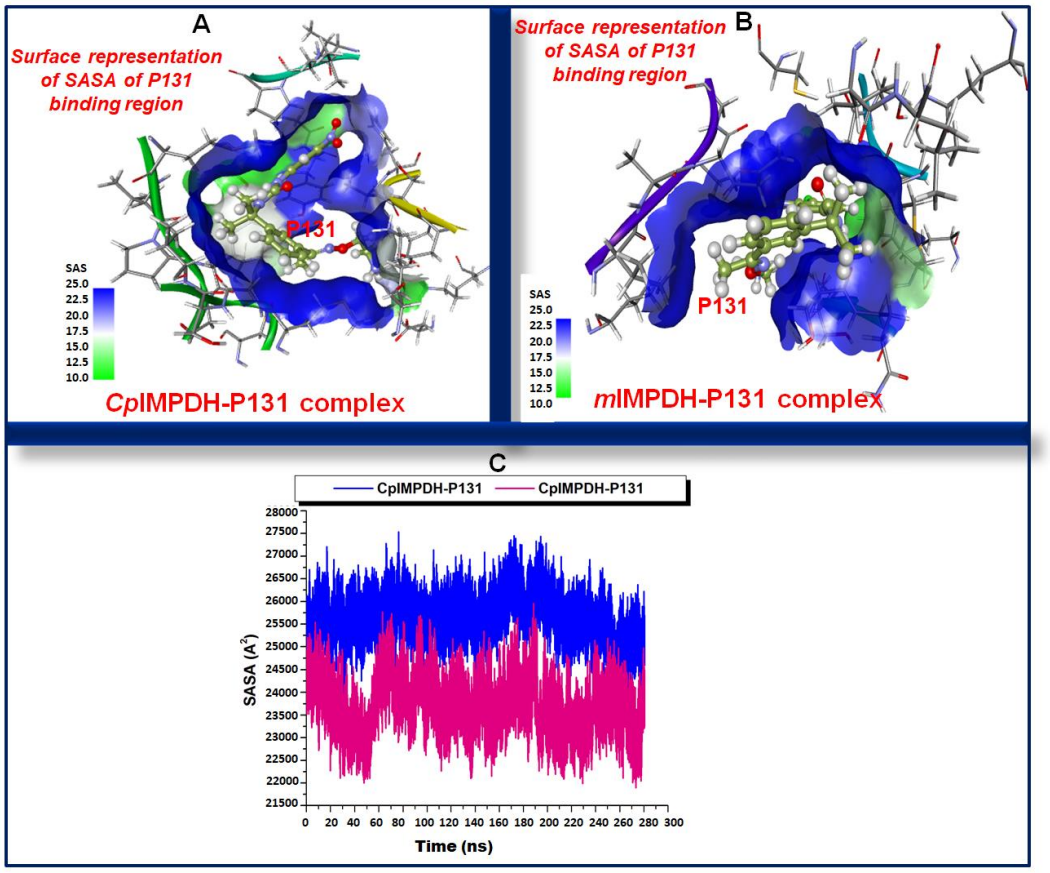


Figure 5.8 (**a**) A 3D surface representation of SASA of the inhibitor binding pocket structure of *Cp*IMPDH upon P131 binding. (**b**) A 3D surface representation of SASA of the inhibitor binding pocket structures of *m*IMPDH. (**c**) Comparative solvent accessible surface area of *Cp*IMPDH (blue), and *m*IMPDH (violet) upon binding of P131.

We further investigated the local vibrations in both *Cp*IMPDH and *m*IMPDH upon the binding of P131 by estimating the average residual fluctuations. These were plotted as root mean square fluctuation (RMSF) as shown in Figure 5.9. The RMSF plot highlighted several residual fluctuations in different regions of both *Cp*IMPDH and *m*IMPDH. Nonetheless, the average fluctuation of *m*IMPDH in the presence of P131 was relatively higher than the *Cp*IMPDH, suggesting the binding of P131 increased the vibration of individual residues. The lower RMSF in the *Cp*IMPDH models suggests that P131 possibly engaged in stronger and steady interactions which could have impeded individual residue motions as consistent with the minimal deviations observed in the RMSD calculations. An average RMSF of 8.57Å and 38.34Å was estimated for the *Cp*IMPDH-P131 and *m*IMPDH-P131 complexes, respectively.

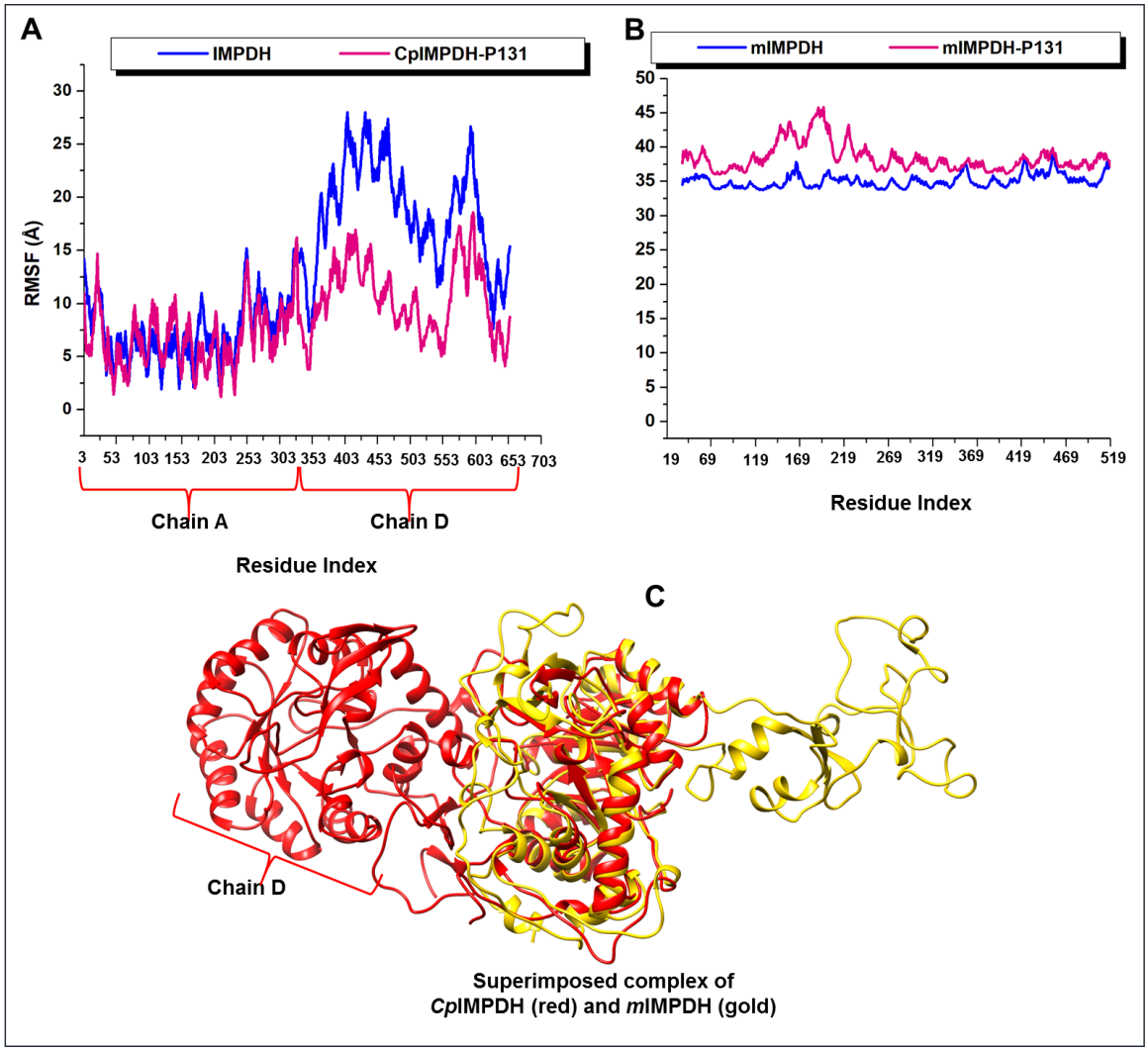


Figure 5.9 (**a**) Comparative RMSF plots of individual residues of bound and unbound conformations of *Cp*IMPDH over the simulation period. (**b**) Comparative RMSF plots of individual residues of bound and unbound conformations of *m*IMPDH over the simulation period. (**c**) 3D representation of superimposed post-MD X-ray crystal structures of bound *Cp*IMPDH (red) and *m*IMPDH (gold)

To further explore the conformational dynamics of both *Cp*IMPDH and *m*IMPDH upon binding of P131, we calculated the radius of gyration (RoG) of each enzyme. RoG provides insights into stability and folding patterns of studied models [98, 99]. Relative to their respective unbound models, the bound *Cp*IMPDH and *m*IMPDH systems exhibited slightly higher average RoG values throughout the simulation. Here, no conformational shift was observed in the RoG plot as shown in Suppl Figure 5.2, which suggests an insignificant structural deviation or unfolding in both *Cp*IMPDH and *m*IMPDH upon P131 binding. The apo and complex *Cp*IMPDH exhibited mean RoG values of 27.63Å and 28.07Å, respectively. The average RoG for unbound and bound

*m*IMPDH were also found to be 26.01Å and 26.55Å, respectively. The juxtaposed RoGs of P131-bound *Cp*IMPDH and *m*IMPDH is represented graphically in Suppl Figure 5.2.

5.5 Conclusion

Molecular dynamics simulation, bioinformatics and advanced post-MD tools were employed in this study to explore the structural basis of P131 selectively towards *Cp*IMPDH rather than *m*IMPDH as reported by previous *in vitro* reports, regardless of the substantial structural and sequence similarity of all the organisms. Using MD simulation, Clustal Omega sequence alignment techniques, MM/PBSA analysis of binding free energies, and per-residue decomposition, the structural basis of P131 selectivity was provided. Sequence alignment of the P131 binding sites of both *Cp*IMPDH and *m*IMPDH revealed variations among constituent amino acids. This allowed the identification of crucial residues within the binding pockets of *Cp*IMPDH that are required binding for P131binding, however, absent in *m*IMPDH. Some of the residues though identical in both organisms, it was noted that in *m*IMPDH, they contributed only little to the binding free energy of P131, suggestive of a weaker binding affinity. Per-residue decomposition plots further consolidated this finding by giving information about the energy contributions of each binding site residue to the affinity and stabilization of P131 within the binding pocket. Amino acid residues such as GLU329 (-3.26kcal/mol), ALA165 (-2.89kcal/mol), HIS166A (-1.11kcal/mol), TYR358 (-2.69kcal/mol), MET302 (-1.82kcal/mol), GLY303A (-1.63kcal/mol) and PRO26 (-1.67kcal/mol) in the *Cp*IMPDH binding site had higher energy contributions. However, compared to the *m*IMPDH complex, energy contributions of binding site residues towards P131 binding were relatively lower, consistent with weaker binding. Calculated binding free energy revealed that P131 binds stronger to *Cp*IMPDH than *m*IMPDH with total binding free energy of - 37.00 kcal/mol and -17.01 kcal/mol, respectively. Analysis of the binding modes and orientations of P131 using representative snapshots showed that the orientations of P131 at the binding site of *Cp*IMPDH over the simulation period possibly enhanced its movement into the deep regions of the P131 binding pocket. This was corroborated by the peculiar and higher number of interacting residues observed as the simulation progressed. On the contrary, the movement of P131 appeared to be restricted in *m*IMPDH binding sites as the simulation progressed, since interacting residues slightly changed and were relatively fewer. Altogether, these findings provide structural insights into the selectivity of P131 towards *Cp*IMPDH. These findings

are in line with earlier reports which showed that interaction of the adenosine subsite with adjacent monomer in prokaryotic IMPDH and the absence of this in eukayotic IMPDH could underpin basis of selectivity and potency of IMPDH inhibitors. Findings could also aid the optimization of P131 with the aim of designing new P131 analogs with improved inhibitory activity and selectivity.

Acknowledgments

Appreciation goes to Centre for High Performance Computing, Cape Town, South Africa for providing computational resources.

Compliance with Ethical Standards

Conflicts of Interest: The authors declare no conflict of interest

5.6 References

1. Bouzid, M., Kintz, E., & Hunter, P. R. (2018). Risk factors for Cryptosporidium infection in low and middle income countries: A systematic review and meta-analysis. *PLoS Neglected Tropical Diseases*, *12*(6). https://doi.org/10.1371/journal.pntd.0006553

2. Fischer Walker, C. L., Aryee, M. J., Boschi-Pinto, C., & Black, R. E. (2012, January 3). Estimating diarrhea mortality among young children in low and middle income countries. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0029151

3. Valentiner-Branth, P., Steinsland, H., Fischer, T. K., Perch, M., Scheutz, F., Dias, F., … Sommerfelt, H. (2003). Cohort study of Guinean children: Incidence, pathogenicity, conferred protection, and attributable risk for enteropathogens during the first 2 years of life. *Journal of Clinical Microbiology*, *41*(9), 4238–4245. https://doi.org/10.1128/JCM.41.9.4238-4245.2003

4. Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., … Black, R. E. (2012). Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000. *The Lancet*, *379*(9832), 2151–2161.

https://doi.org/10.1016/S0140-6736(12)60560-1

5. Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., … Levine, M. M. (2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *The Lancet*, *382*(9888), 209–222. https://doi.org/10.1016/S0140-6736(13)60844-2

6. Sow, S., Muhsen, K., & Nasrin, D. (2016). The Burden of Cryptosporidium Diarrheal Disease among Children &lt.

7. Odeniran, P. O., & Ademola, I. O. (2019). Epidemiology of Cryptosporidium infection in different hosts in Nigeria: A meta-analysis. *Parasitology International*, *71*(May), 194–206. https://doi.org/10.1016/j.parint.2019.04.007

8. Delahoy, M. J., Omore, R., Ayers, T. L., Schilling, K. A., Blackstock, A. J., Ochieng, J. B., … O'Reilly, C. E. (2018). Clinical, environmental, and behavioral characteristics associated with Cryptosporidium infection among children with moderate-to-severe diarrhea in rural western Kenya, 2008–2012: The Global Enteric Multicenter Study (GEMS). *PLoS Neglected Tropical Diseases*, *12*(7). https://doi.org/10.1371/journal.pntd.0006640

9. Umejiego, N. N., Li, C., Riera, T., Hedstrom, L., & Striepen, B. (2004). Cryptosporidium parvum IMP dehydrogenase: Identification of functional, structural, and dynamic properties that can be exploited for drug design. *Journal of Biological Chemistry*, *279*(39), 40320–40327. https://doi.org/10.1074/jbc.M407121200

10. Putignani, L., & Menichella, D. (2010). Global distribution, public health and clinical impact of the protozoan pathogen cryptosporidium.

11. O'connor, R. M., Shaffie, R., Kang, G., & Ward, H. D. (2011). Cryptosporidiosis in patients with HIV/AIDS. *AIDS (London, England)*, *25*(5), 549–60. https://doi.org/10.1097/QAD.0b013e3283437e88

12. Chalmers, R. M., & Davies, A. P. (2010). Minireview: clinical cryptosporidiosis.

Experimental parasitology, *124*(1), 138–46. https://doi.org/10.1016/j.exppara.2009.02.003

13. Kang, G., Sarkar, R., & Desai, N. (2012). Cryptosporidiosis: An under-recognized public health problem. *Tropical Parasitology*, *2*(2), 91. https://doi.org/10.4103/2229-5070.105173

14. Putignani, L., & Menichella, D. (2010). Global distribution, public health and clinical impact of the protozoan pathogen cryptosporidium. *Interdisciplinary perspectives on infectious diseases*, *2010*. https://doi.org/10.1155/2010/753512

15. Aldeyarbi, H. M., Abu El-Ezz, N. M. T., & Karanis, P. (2016). Cryptosporidium and cryptosporidiosis: the African perspective. *Environmental Science and Pollution Research*, *23*(14), 13811–13821. https://doi.org/10.1007/s11356-016-6746-6

16. Abdool Karim, S. S., Churchyard, G. J., Karim, Q. A., & Lawn, S. D. (2009). HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet (London, England)*, *374*(9693), 921–33. https://doi.org/10.1016/S0140-6736(09)60916-8

17. Checkley, W., Jr., A. W., & Jaganath, D. (2015). A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium.

18. Mead, J. R. (2002). Cryptosporidiosis and the challenges of chemotherapy. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, *5*(1), 47–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12127863

19. Lee, S., Harwood, M., Girouard, D., Meyers, M. J., Campbell, M. A., Beamer, G., & Tzipori, S. (2017). The therapeutic efficacy of azithromycin and nitazoxanide in the acute pig model of Cryptosporidium hominis. *PLOS ONE*, *12*(10), e0185906. https://doi.org/10.1371/journal.pone.0185906

20. Cabada, M. M., & White, A. C. (2010). Treatment of cryptosporidiosis: do we know what we think we know? *Current opinion in infectious diseases*, *23*(5), 494–9. https://doi.org/10.1097/QCO.0b013e32833de052

21. Sparks, H., Nair, G., Castellanos-Gonzalez, A., & White, A. C. (2015). Treatment of Cryptosporidium: What We Know, Gaps, and the Way Forward. *Current Tropical Medicine Reports*, *2*(3), 181–187. https://doi.org/10.1007/s40475-015-0056-9

22. Fan, X., Upadhyaya, B., Wu, L., Koh, C., Santín-Durán, M., Pittaluga, S., … Jain, A. (2012). CD40 agonist antibody mediated improvement of chronic Cryptosporidium infection in patients with X-linked hyper IgM syndrome. *Clinical Immunology*, *143*(2), 152–161. https://doi.org/10.1016/j.clim.2012.01.014

23. Hewitt, R. G., Yiannoutsos, C. T., Higgs, E. S., Carey, J. T., Geiseler, P. J., Soave, R., … Bender, J. F. (2000). Paromomycin: no more effective than placebo for treatment of cryptosporidiosis in patients with advanced human immunodeficiency virus infection. AIDS Clinical Trial Group. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *31*(4), 1084–92. https://doi.org/10.1086/318155

24. Hussien, S. M. M., Abdella, O. H., Abu-Hashim, A. H., Aboshiesha, G. A., Taha, M. A. A., El-Shemy, A. S., & El-Bader, M. M. (2013). Comparative study between the effect of nitazoxanide and paromomycine in treatment of cryptosporidiosis in hospitalized children. *Journal of the Egyptian Society of Parasitology*, *43*(2), 463–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24260825

25. Allam, A. F., & Shehab, A. Y. (2002). Efficacy of azithromycin, praziquantel and mirazid in treatment of cryptosporidiosis in school children. *Journal of the Egyptian Society of Parasitology*, *32*(3), 969–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12512828

26. Raja, K., Abbas, Z., Hassan, S. M., Luck, N. H., Aziz, T., & Mubarak, M. (2014). Prevalence of cryptosporidiosis in renal transplant recipients presenting with acute diarrhea at a single center in Pakistan. *Journal of Nephropathology*, *3*(4), 127–131. https://doi.org/10.12860/jnp.2014.25

27. Holmberg, S. D., Moorman, A. C., Von Bargen, J. C., Palella, F. J., Loveless, M. O., Ward, D. J., & Navin, T. R. (1998). Possible effectiveness of clarithromycin and rifabutin for

cryptosporidiosis chemoprophylaxis in HIV disease. HIV Outpatient Study (HOPS) Investigators. *JAMA*, *279*(5), 384–6. https://doi.org/10.1001/jama.279.5.384

28. Amenta, M., Dalle Nogare, E. R., Colomba, C., Prestileo, T. S., Di Lorenzo, F., Fundaro, S., … Ferrieri, A. (1999). Intestinal protozoa in HIV-infected patients: effect of rifaximin in Cryptosporidium parvum and Blastocystis hominis infections. *Journal of chemotherapy (Florence, Italy)*, *11*(5), 391–5. https://doi.org/10.1179/joc.1999.11.5.391

29. Gathe, J. C., Mayberry, C., Clemmons, J., & Nemecek, J. (2008). Resolution of severe cryptosporidial diarrhea with rifaximin in patients with AIDS. *Journal of acquired immune deficiency syndromes (1999)*, *48*(3), 363–4. https://doi.org/10.1097/QAI.0b013e31817beb78

30. Legrand, F., Grenouillet, F., Larosa, F., Dalle, F., Saas, P., Millon, L., … Rohrlich, P. S. (2011). Diagnosis and treatment of digestive cryptosporidiosis in allogeneic haematopoietic stem cell transplant recipients: A prospective single centre study. *Bone Marrow Transplantation*, *46*(6), 858–862. https://doi.org/10.1038/bmt.2010.200

31. Bonatti, H., Barroso, L. F., Sawyer, R. G., Kotton, C. N., & Sifri, C. D. (2012). Cryptosporidium enteritis in solid organ transplant recipients: Multicenter retrospective evaluation of 10 cases reveals an association with elevated tacrolimus concentrations. *Transplant Infectious Disease*, *14*(6), 635–648. https://doi.org/10.1111/j.1399-3062.2012.00719.x

32. Krause, I., Amir, J., Cleper, R., Dagan, A., Behor, J., Samra, Z., & Davidovits, M. (2012). Cryptosporidiosis in children following solid organ transplantation. *Pediatric Infectious Disease Journal*, *31*(11), 1135–1138. https://doi.org/10.1097/INF.0b013e31826780f7

33. Chen, L., Wilson, D. J., Xu, Y., Aldrich, C. C., Felczak, K., Sham, Y. Y., & Pankiewicz, K. W. (2010). Triazole-linked inhibitors of inosine monophosphate dehydrogenase from human and mycobacterium tuberculosis. *Journal of Medicinal Chemistry*, *53*(12), 4768–4778. https://doi.org/10.1021/jm100424m

34. Gollapalli, D. R., MacPherson, I. S., Liechti, G., Gorla, S. K., Goldberg, J. B., & Hedstrom, L. (2010). Structural determinants of inhibitor selectivity in prokaryotic IMP dehydrogenases. *Chemistry and Biology*, *17*(10), 1084–1091. https://doi.org/10.1016/j.chembiol.2010.07.014

35. Hedstrom, L., Liechti, G., Goldberg, J. B., & Gollapalli, D. R. (2011). The antibiotic potential of prokaryotic IMP dehydrogenase inhibitors. *Current medicinal chemistry*, *18*(13), 1909–18. https://doi.org/10.2174/092986711795590129

36. Usha, V., Hobrath, J. V., Gurcha, S. S., Reynolds, R. C., & Besra, G. S. (2012). Identification of novel Mt-Guab2 inhibitor series active against M. tuberculosis. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033886

37. Wang, W., & Hedstrom, L. (1997). Kinetic mechanism of human inosine 5'-monophosphate dehydrogenase type II: random addition of substrates and ordered release of products. *Biochemistry*, *36*(28), 8479–83. https://doi.org/10.1021/bi970226n

38. Zimmermann, A., Gu, J. J., Spychala, J., & Mitchell, B. S. (1996). Inosine monophosphate dehydrogenase expression: Transcriptional regulation of the type I and type II genes. In *Advances in Enzyme Regulation* (Vol. 36, pp. 75–84). Elsevier Ltd. https://doi.org/10.1016/0065-2571(95)00012-7

39. Striepen, B., Pruijssers, A. J. P., Huang, J., Li, C., Gubbels, M. J., Umejiego, N. N., … Kissinger, J. C. (2004). Gene transfer in the evolution of parasite nucleotide biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 3154–3159. https://doi.org/10.1073/pnas.0304686101

40. Zhang, R. G., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., … Collart, F. R. (1999). Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry*, *38*(15), 4691–4700. https://doi.org/10.1021/bi982858v

41. Bhagavan, N. V., & Ha, C.-E. (2015). Nucleotide Metabolism. In *Essentials of Medical*

Biochemistry (pp. 465–487). Elsevier. https://doi.org/10.1016/b978-0-12-416687-5.00025-7

42. Kim, Y., Makowska-Grzyska, M., Gorla, S. K., Gollapalli, D. R., Cuny, G. D., Joachimiak, A., & Hedstrom, L. (2015). Structure of Cryptosporidium IMP dehydrogenase bound to an inhibitor with in vivo antiparasitic activity. *Acta Crystallographica Section F:Structural Biology Communications*, *71*, 531–538. https://doi.org/10.1107/S2053230X15000187

43. Makowska-Grzyska, M., Kim, Y., Maltseva, N., Osipiuk, J., Gu, M., Zhang, M., … Joachimiak, A. (2015). A novel cofactor-binding mode in bacterial IMP dehydrogenases explains inhibitor selectivity. *Journal of Biological Chemistry*, *290*(9), 5893–5911. https://doi.org/10.1074/jbc.M114.619767

44. Hedstrom, L. (2009). IMP dehydrogenase: Structure, mechanism, and inhibition. *Chemical Reviews*, *109*(7), 2903–2928. https://doi.org/10.1021/cr900021w

45. Felczak, K., Chen, L., Wilson, D., Williams, J., Vince, R., Petrelli, R., … Pankiewicz, K. W. (2011). Cofactor-type inhibitors of inosine monophosphate dehydrogenase via modular approach: Targeting the pyrophosphate binding sub-domain. *Bioorganic and Medicinal Chemistry*, *19*(5), 1594–1605. https://doi.org/10.1016/j.bmc.2011.01.042

46. Colby, T. D., Vanderveen, K., Strickler, M. D., Markham, G. D., & Goldstein, B. M. (1999). Crystal structure of human type II inosine monophosphate dehydrogenase: Implications for ligand binding and drug design. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(7), 3531–3536. https://doi.org/10.1073/pnas.96.7.3531

47. Gorla, S. K., Kavitha, M., Zhang, M., Liu, X., Sharling, L., Gollapalli, D. R., … Cuny, G. D. (2012). Selective and potent urea inhibitors of cryptosporidium parvum inosine 5′-monophosphate dehydrogenase. *Journal of Medicinal Chemistry*, *55*(17), 7759–7771. https://doi.org/10.1021/jm3007917

48. Macpherson, I. S., Kirubakaran, S., Gorla, S. K., Riera, T. V, D'Aquino, J. A., Zhang, M., … Hedstrom, L. (2010). The structural basis of Cryptosporidium -specific IMP

dehydrogenase inhibitor selectivity. *Journal of the American Chemical Society*, *132*(4), 1230–1. https://doi.org/10.1021/ja909947a

49. Gorla, S. K., McNair, N. N., Yang, G., Gao, S., Hu, M., Jala, V. R., … Hedstrom, L. (2014). Validation of IMP dehydrogenase inhibitors in a mouse model of cryptosporidiosis. *Antimicrobial Agents and Chemotherapy*, *58*(3), 1603–1614. https://doi.org/10.1128/AAC.02075-13

50. Johnson, C. R., Gorla, S. K., Kavitha, M., Zhang, M., Liu, X., Striepen, B., … Hedstrom, L. (2013). Phthalazinone inhibitors of inosine-5'-monophosphate dehydrogenase from Cryptosporidium parvum. *Bioorganic & medicinal chemistry letters*, *23*(4), 1004–7. https://doi.org/10.1016/j.bmcl.2012.12.037

51. Kirubakaran, S., Gorla, S. K., Sharling, L., Zhang, M., Liu, X., Ray, S. S., … Cuny, G. D. (2012). Structure-activity relationship study of selective benzimidazole-based inhibitors of Cryptosporidium parvum IMPDH. *Bioorganic & medicinal chemistry letters*, *22*(5), 1985–8. https://doi.org/10.1016/j.bmcl.2012.01.029

52. Maurya, S. K., Gollapalli, D. R., Kirubakaran, S., Zhang, M., Johnson, C. R., Benjamin, N. N., … Cuny, G. D. (2009). Triazole inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. *Journal of medicinal chemistry*, *52*(15), 4623–30. https://doi.org/10.1021/jm900410u

53. Sharling, L., Liu, X., Gollapalli, D. R., Maurya, S. K., Hedstrom, L., & Striepen, B. (2010). A screening pipeline for antiparasitic agents targeting cryptosporidium inosine monophosphate dehydrogenase. *PLoS neglected tropical diseases*, *4*(8), e794. https://doi.org/10.1371/journal.pntd.0000794

54. Umejiego, N. N., Gollapalli, D., Sharling, L., Volftsun, A., Lu, J., Benjamin, N. N., … Hedstrom, L. (2008). Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against cryptosporidiosis. *Chemistry & biology*, *15*(1), 70–7. https://doi.org/10.1016/j.chembiol.2007.12.010

55. Sun, Z., Khan, J., Makowska-Grzyska, M., Zhang, M., Cho, J. H., Suebsuwong, C., … Cuny, G. D. (2014). Synthesis, in vitro evaluation and cocrystal structure of 4-oxo-[1]benzopyrano[4,3-c]pyrazole Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase (CpIMPDH) inhibitors. *Journal of medicinal chemistry*, *57*(24), 10544–50. https://doi.org/10.1021/jm501527z

56. Gorla, S. K., Kavitha, M., Zhang, M., Chin, J. E. W., Liu, X., Striepen, B., … Cuny, G. D. (2013). Optimization of benzoxazole-based inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. *Journal of medicinal chemistry*, *56*(10), 4028–43. https://doi.org/10.1021/jm400241j

57. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*. https://doi.org/10.1038/msb.2011.75

58. Abrahamsen, M. S., Templeton, T. J., Enomoto, S., Abrahante, J. E., Zhu, G., Lancto, C. A., … Kapur, V. (2004). Complete Genome Sequence of the Apicomplexan, Cryptosporidium parvum. *Science*, *304*(5669), 441–445. https://doi.org/10.1126/science.1094786

59. Tiedeman, A. A., & Smith, J. M. (1991). Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase. *Gene*, *97*(2), 289–93. https://doi.org/10.1016/0378-1119(91)90065-j

60. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., … Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

61. Sintchak, M. D., Fleming, M. A., Futer, O., Raybuck, S. A., Chambers, S. P., Caron, P. R., … Wilson, K. P. (1996). Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. *Cell*, *85*(6), 921–930. https://doi.org/10.1016/S0092-8674(00)81275-1

62. Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server), W407–W410. https://doi.org/10.1093/nar/gkm290

63. Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in enzymology*, *277*, 396–404.

64. K. Gopalakrishnan, G. Sowmiya, S. S. Sheik, & K. Sekar. (2007). Ramachandran Plot on The Web (2.0). *Protein & Peptide Letters*, *14*(7), 669–671. https://doi.org/10.2174/092986607781483912

65. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

66. Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., … Sali, A. (2006). Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics*, *Chapter 5*, Unit-5.6. https://doi.org/10.1002/0471250953.bi0506s15

67. Trott, O., & Olson, A. J. (2010). Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc.21334

68. Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., … Hou, T. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: The prediction accuracy of sampling power and scoring power. *Physical Chemistry Chemical Physics*, *18*(18), 12964–12975. https://doi.org/10.1039/c6cp01555g

69. Machaba, K. E., Mhlongo, N. N., & Soliman, M. E. S. (2018). Induced Mutation Proves a Potential Target for TB Therapy: A Molecular Dynamics Study on LprG. *Cell Biochemistry*

and Biophysics, *76*(3), 345–356. https://doi.org/10.1007/s12013-018-0852-7

70. Oguntade, S., Ramharack, P., & Soliman, M. E. (2017). Characterizing the ligand-binding landscape of Zika NS3 helicase-promising lead compounds as potential inhibitors. *Future Virology*, *12*(6), 261–273. https://doi.org/10.2217/fvl-2017-0014

71. Olotu, F. A., & Soliman, M. E. S. (2018). From mutational inactivation to aberrant gain-of-function: Unraveling the structural basis of mutant p53 oncogenic transition. *Journal of Cellular Biochemistry*, *119*(3), 2646–2652. https://doi.org/10.1002/jcb.26430

72. Agoni, C., Munsamy, G., Ramhrack, P., & Soliman, M. (2020). Human Rhinovirus Inhibition Through Capsid "Canyon" Perturbation: Structural Insights into The Role of a Novel Benzothiophene Derivative. *Cell Biochemistry and Biophysics*, *78*, 3–13.

73. Agoni, C., Salifu, E. Y., Munsamy, G., Olotu, F. A., & Soliman, M. (2019). CF3-pyridinyl substitution on anti-malarial therapeutics: Probing differential ligand binding and dynamical inhibitory effects of a novel triazolopyrimidine-based inhibitor on Plasmodium falciparum Dihydroorotate dehydrogenase. *Chemistry & Biodiversity*. https://doi.org/10.1002/cbdv.201900365

74. Olotu, F. A., & Soliman, M. E. S. (2019). Dynamic perspectives into the mechanisms of mutation-induced p53-DNA binding loss and inactivation using active perturbation theory: Structural and molecular insights toward the design of potent reactivators in cancer therapy. *Journal of cellular biochemistry*, *120*(1), 951–966. https://doi.org/10.1002/jcb.27458

75. Nair, P. C., & Miners, J. O. (2014). Molecular dynamics simulations: from structure function relationships to drug discovery. *In Silico Pharmacology*, *2*(1). https://doi.org/10.1186/s40203-014-0004-8

76. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of computational chemistry*, *25*(9), 1157–74. https://doi.org/10.1002/jcc.20035

77. Grest, G. S., & Kremer, K. (1986). Molecular dynamics simulation for polymers in the

presence of a heat bath. *Physical Review A*, *33*(5), 3628–3631. https://doi.org/10.1103/PhysRevA.33.3628

78. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

79. Ryckaert, J.-P., Ciccotti+, G., & Berendsen, H. J. C. (1977). *Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes*. *JOURNAL OF COMPUTATIONAL. PHYSICS* (Vol. 23).

80. Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, *9*(7), 3084–3095. https://doi.org/10.1021/ct400341p

81. Seifert, E. (2014). OriginPro 9.1: scientific data analysis and graphing software-software review. *Journal of chemical information and modeling*, *54*(5), 1552. https://doi.org/10.1021/ci500161d

82. Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, *10*(5), 449–61. https://doi.org/10.1517/17460441.2015.1032936

83. Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, *51*(1), 69–82. https://doi.org/10.1021/ci100275a

84. Chaudhary, N., & Aparoy, P. (2017). Deciphering the mechanism behind the varied binding activities of COXIBs through Molecular Dynamic Simulations, MM-PBSA binding energy calculations and per-residue energy decomposition studies. *Journal of biomolecular structure & dynamics*, *35*(4), 868–882. https://doi.org/10.1080/07391102.2016.1165736

85. Gupta, A., Chaudhary, N., & Aparoy, P. (2018). MM-PBSA and per-residue decomposition

energy studies on 7-Phenyl-imidazoquinolin-4(5H)-one derivatives: Identification of crucial site points at microsomal prostaglandin E synthase-1 (mPGES-1) active site. *International journal of biological macromolecules*, *119*, 352–359. https://doi.org/10.1016/j.ijbiomac.2018.07.050

86. Beg, A., Khan, F., Lobb, K., Islam, A., Ahmad, F., & Hassan, M. (2019). High throughput screening, docking, and molecular dynamics studies to identify potential inhibitors of human calcium/calmodulin-dependent protein kinase IV. *J. Biomol. Struct. Dyn.*, *37*, 2179–2192.

87. Dahiya, R., Mohammad, T., Roy, S., Anwar, S., Gupta, P., Haque, A., … Ahmad, F. (2019). Investigation of inhibitory potential of quercetin to the pyruvate dehydrogenase kinase 3:Towards implications in anticancer therapy. *Int. J. Biol. Macromol.*, *136*, 1076–1085.

88. Fatima, S., Mohammad, T., Jairajpuri, D., Rehman, M., Hussain, A., Samim, M., … Hassan, M. (2019). Identification and evaluation of glutathione conjugate gamma-l-glutamyl-l-cysteine for improved drug delivery to the brain. *J. Biomol. Struct. Dyn.*, 1–11.

89. Gulzar, M., Ali, S., Khan, F., Khan, P., Taneja, P., & Hassan, M. (2019). Binding mechanism of ca eic acid and simvastatin to the integrin linked kinase for therapeutic implications: A comparative docking and MD simulation studies. *J. Biomol. Struct. Dyn.*, *37*, 4327–4337.

90. Kuzmanic, A., & Zagrovic, B. (2010). Determination of ensemble-average pairwise root mean square deviation from experimental B-factors. *Biophys. J.*, *98*, 861–871.

91. Naz, F., Shahbaaz, M., Bisetty, K., Islam, A., Ahmad, F., & Hassan, M. (2015). Designing New Kinase Inhibitor Derivatives as Therapeutics Against Common Complex Diseases: Structural Basis of Microtubule Affinity-Regulating Kinase 4 (MARK4) Inhibition. *OMICS*, *19*, 700–711.

92. Naz, F., Shahbaaz, M., Khan, S., Bisetty, K., Islam, A., Ahmad, F., & Hassan, M. (2015). PKR-inhibitor binds efficiently with human microtubule affinity-regulating kinase 4. *J.*

Mol. Graph. Model., *62*, 245–252.

93. Naz, H., Shahbaaz, M., Haque, M., Bisetty, K., Islam, A., Ahmad, F., & Hassan, M. (2017). Urea-induced denaturation of human calcium/calmodulin-dependent protein kinase IV: A combined spectroscopic and MD simulation studies. *J. Biomol. Struct. Dyn.*, *35*, 463–475.

94. Menendez, C., Mazola, Y., Guirola, O., Palomares, S., Chinea, G., Hernandez, L., & Musacchio, A. (2015). A comparative molecular dynamics study of thermophilic and mesophilic beta-fructosidase enzymes. *J. Mol. Model.*, *21*, 2772.

95. Ali, S., Hassan, M., Islam, A., & Ahmad, F. (2014). A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. *Curr. Protein Pept. Sci.*, *15*, 456–476.

96. Ausaf Ali, S., Imtaiyaz Hassan, M., Islam, A., & Ahmad, F. (2014). *Send Orders for Reprints to reprints@benthamscience.net A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States*.

97. Durham, E., Dorr, B., Woetzel, N., Staritzbichler, R., & Meiler, J. (2009). Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *Journal of Molecular Modeling*, *15*(9), 1093–1108. https://doi.org/10.1007/s00894-009-0454-9

98. Ali, S., Khan, F., Mohammad, T., Lan, D., Hassan, M., & Wang, Y. (2019). Identification and Evaluation of Inhibitors of Lipase from Malassezia restricta using Virtual High-Throughput Screening and Molecular Dynamics Studies. *Int. J. Mol. Sci.*, *20*, 884.

99. Khan, F., Shahbaaz, M., Bisetty, K., Waheed, A., Sly, W., Ahmad, F., & Hassan, M. (2016). Large scale analysis of the mutational landscape in beta-glucuronidase: A major player of mucopolysaccharidosis type VII. Gene, *576*, 36–44.

CHAPTER 6

A Probable Means to an End: Exploring P131 Pharmacophoric Scaffold to Identify Potential Inhibitors of *Cryptosporidium parvum* Inosine Monophosphate Dehydrogenase

Kehinde F. Omolabi[a], Emmanuel A. Iwuchukwu[a], Clement Agoni[a], Fisayo A. Olotu[a] and Mahmoud E.S. Soliman[a*]

[a]Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Lab webpage: http://soliman.ukzn.ac.za/

ORCID No

Kehinde F Omolabi: https://orcid.org/0000-0001-9611-7147

Emmanuel A. Iwuchukwu: https://orcid.org/0000-0002-9554-016X

Clement Agoni: https://orcid.org/0000-0001-6130-8031

Fisayo A. Olotu: https://orcid.org/0000-0003-3604-5983

Mahmoud E.S Soliman: https://orcid.org/0000-0002-8711-7783

6.1 Abstract

Compound P131 has been established to inhibit *Cryptosporidium parvum*'s inosine monophosphate dehydrogenase (*Cp*IMPDH). Its inhibitory activity supersedes that of paromomycin, which is extensively used in treating cryptosporidiosis. Through the per residue energy decomposition approach, crucial moieties of P131 were identified and subsequently adopted to create a pharmacophore model for virtual screening in the ZINC database. This search generated eight ADMET-compliant hits that were examined thoroughly to fit into the active site of *Cp*IMPDH via molecular docking. Three compounds ZINC46542062, ZINC58646829, and ZINC89780094, with favorable docking scores of -8.3kcal/mol, -8.2 kcal/mol, and -7.5kcal/mol, were selected. The potential inhibitory mechanism of these compounds was probed using molecular dynamics simulation and Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA) analyses. Results revealed that one of the hits (ZINC46542062) exhibited a higher binding free energy of -39.52kcal/mol than P131, which had -34.6 kcal/mol. Conformational perturbation induced by the binding of the identified hits to *Cp*IMPDH was similar to P131, suggesting a similarity in inhibitory mechanisms. Also, *in silico* investigation of the properties of the hit compounds implied superior physicochemical properties with regards to their synthetic accessibility, lipophilicity, number of hydrogen bond donors and acceptors in comparison with P131. ZINC46542062 was identified as a promising hit compound with the highest binding affinity to the target protein and favorable physicochemical and pharmacokinetic properties relative to P131. The identified compounds can serve as a basis for conducting further experimental investigations towards the development of anticryptosporidials, which can overcome the challenges of existing therapeutic options.

Keywords: P131, *Cryptosporidium parvum* IMPDH, Per residue energy decomposition, ADMET, Virtual Screening, MM/PBSA.

6.2 Introduction

Cryptosporidiosis is an opportunistic parasitic disease and a significant cause of diarrhea in humans [1, 2]. Immunocompromised populations such as infants, HIV, and T-cell immunodeficient patients are primarily at risk [3–5]. Clinical manifestation of the infection includes abdominal pain, nausea, watery diarrhea, and low-grade fever. If not managed properly, the symptoms can be life-threatening [6–9]. Cryptosporidiosis takes the second position after rotavirus as the leading cause of diarrheal related death in children less than five years old [10]. In sub–Saharan Africa, where the disease is highly endemic, it accounts for around 2.4 million deaths in children young of >24months [11]. Most human cryptosporidiosis is caused by *Cryptosporidium hominis* and *Cryptosporidium parvum* [12, 13].

Presently there is a therapeutic challenge in treating cryptosporidiosis [14]. The only FDA-approved drug for the disease is nitazoxanide, a compound that contains both a thiazole ring and a benzamidine ring [4]. Although nitazoxanide is well-tolerated and does not present with a significant adverse drug reaction, it is not effective in immunodeficient individuals [15]. Other non-FDA-approved drugs administered in treating cryptosporidiosis but with limited success include paromomycin, azithromycin, and rifaximin [14, 16–18]. The side effects of these drugs include nephrotoxicity, hepatotoxicity, hearing loss, seizures, etc.[19,20]. In light of the inadequacies of the current anticryptosporidials, it is therefore urgent to discover therapeutic compounds that have the propensity to overcome the deficiencies of these drugs.

Hedstrom *et al*. research group tackled this by synthesizing several thiazole-based ligands inhibiting *Cryptosporidium* growth [21, 22, 31, 23–30]. These efforts produced a potential drug molecule- P131, which demonstrated excellent anticryptosporidial activity *in vivo.* In the *in vivo* experiment using a murine model, paromomycin was used as a control as nitazoxanide is not bioavailable. P131 at a single dose had the same therapeutic outcome when administered at 250mg/kg body weight compared to the control group, which was treated with paromomycin at 2000mg/kg body weight. With a thrice-daily administration of both drugs at the same concentration stated above, P131 elicited a superior parasiticidal activity when compared to paromomycin [23].

P131 potentiates its therapeutic action by inhibiting the *Cryptosporidium parvum*'s inosine monophosphate dehydrogenase (*Cp*IMPDH). It does this by binding to the NAD^+ site (co-factor binding site) of the enzyme [32–35]. Inosine monophosphate dehydrogenase (IMPDH) is an essential enzyme for almost all organisms [34]. It catalyzes the conversion of inosine monophosphate (IMP) to xanthine monophosphate (XMP), which is the first and the rate-limiting step in the biosynthesis of guanine nucleotides [34]. Guanine nucleotides function as precursors for glycosylation, RNA, DNA, and tetrahydrobiopterin synthesis [36, 37]. Therefore the inhibition of IMPDH hinders proliferation and eventuates in cell death.

Pharmacophore modeling is gaining ground as one of the paths to discovering novel drug candidates with favorable pharmacokinetics and pharmacodynamics properties [38]. These properties will altogether function to overcome the drawbacks limiting the efficacy of existing drug molecules. Among many others, these downsides include cross-resistance, toxicity, and adverse drug reaction [39]. The primary application of pharmacophore modeling is virtual screening, which is a shorter but efficient route to otherwise capital-intensive and time-consuming processes of drug discovery [40].

Although an abstract concept, pharmacophore modeling is an approach that elucidates and identifies the common chemical moieties of a set of ligands, which are crucial in eliciting a biological function [41–48]. In this study, the per residue energy decomposition (PRED) method was employed in the pharmacophore modeling of P131 coupled with a virtual screening of the ZINC database. The screening is to generate compounds with the potential to treat cryptosporidiosis. We examined the physicochemical and therapeutic suitability of the compounds as anticryptosporidials. Finally, to understand the molecular mechanism of these new drugs' action, molecular dynamics simulation of the new compounds was carried out using the parent compound P131 as a comparative reference. The conclusions from this might give the experimental research groups potential drug leads that probably would overcome the limited efficacy of nitazoxanide.

6.3 Computational Methods

6.3.1 Pharmacophore model generation and virtual screening

At the initial step, We obtained *Cryptosporidium parvum* IMPDH (*Cp*IMPDH) co-crystallized with P131 at the co-factor site and inosine monophosphate at the active site from Protein Data Bank (PDB code:4RV8) [32]. Although a tetramer, only chains A and D with 652 residues were used to reduce computational resources. The target protein structure (*Cp*IMPDH) was prepared using the Graphical User Interface (GUI) of UCSF Chimera [49], which involved the removal of ions, crystal waters, and non-standard residues. Missing residues were added using MODELLER [50]. A molecular dynamics simulation run of 20ns was carried out to stabilize the pose of P131 in the co-factor pocket of *Cp*IMPDH and serve as the cornerstone for the generation of a P131 pharmacophoric model. The total binding free energy of P131 to *Cp*IMPDH was computed. After that, the per residue energy decomposition (PRED) was analyzed by manually adding the binding residues of P131 to the MMPB/SA.py script integrated with Amber18Tools. PRED estimates the binding free energy each binding site residue contributes to the ligand's overall binding free energy to the protein. Extrapolating from the PRED, the residues which contributed most to the binding of P131 and the moieties these residues interacted with in P131 were used to build a pharmacophore model (Figure 6.1A-D). The pharmacophore query was uploaded in ZINCPharmer [51] to screen the ZINC database [52] for hits that possess these pharmacophoric moieties. The inclusion criteria were as follows: molecular weight of<500g/mol, hydrogen bond acceptors<10, hydrogen bond donors<5, and rotatable bonds<5. The output was further filtered down to exclude non-druglike hits by further screening with Lipinski's rule of five and ADMET properties [53, 54].

6.3.2 Molecular docking

With the prior definition of the P131 binding site in *Cp*IMPDH, we performed molecular docking of the hits generated through virtual screening. Molecular docking was done to ascertain the complementarity of the hits to the co-factor binding site in *Cp*IMPDH [55]. With Autodock Vina, the ligand docking estimations were carried out [56]. Employing the Autodock Graphical user interface by MGL tools, Gasteiger partial chargers were assigned, and atom types were defined [57]. The docked conformations were produced using the Lamarckian Genetic Algorithm [58]. In converting the ligand SDF file to the mol2 format, Maestro software was used [59], and to pdbqt

format, we employed Raccoon integrated with AutoDock suite [57]. Autodock Vina was used in defining the dimensions and coordinates of the grid box. The grid box was defined as center (X=2.649, Y=21.997, Z=77.399) and size (X=11.165, Y=9.462, Z=7.323. During the docking process, a maximum of 8 conformers was considered for each compound. After screening, molecular docking, and filtering, we selected the top three ligands with the best binding affinities towards *Cp*IMPDH. The docked conformation of the compounds complexed with *Cp*IMPDH was generated and visualized in the ViewDock plugin-integrated Chimera [49]. After that, MD simulation and post-MD analysis were carried out on the docked complexes.

6.3.3 Physicochemical properties and bioactivity screening of identified compounds

The pharmacokinetic (ADMET) properties of the identified compounds and the parent compound P131 were evaluated using validated online prediction tools, which have also been used extensively in other studies [60–62]. These include SWISSADME [53], which was used to predict/compute molecular weight, synthetic accessibility, lipophilicity, hydrogen bond acceptor and donor, bioavailability score, water-solubility, gastrointestinal absorption and, blood-brain barrier permeability. Molinspiration Cheminformatics [63] was employed in predicting the topological surface area (TPSA) and the number of rotatable bonds. Likewise, oral toxicities and LD_{50} of the identified compounds and P131 were predicted using the ProTox web server [64]. This presents a faster approach for determining doses that are or have the potential to be toxic. Finally, DataWarrior [65] was used to predict ligand efficiency (LE), ligand lipophilic efficiency (LLE), and ligand-efficiency dependent lipophilicity (LELP). All these web servers were used to determine the pharmacokinetic (ADMET) properties of the ligands and to evaluate how well they adhered to Lipinski's rule of five, which are a set of rules that predict the drug-likeness of a molecule [54].

6.3.4 Molecular dynamics (MD) simulation

Sequel to the docking of the identified hits in the co-factor site of *Cp*IMPDH, we prepared the docked complexes for MD simulation. Each system was subdivided into apo (IMPDH bound to IMP in its active site) and complexes, comprised of apo bound by P131 and apo bound by the hits in the co-factor site. Afterward, these were set up for a 200ns-MD simulation according to previously reported protocols [66–68]. MD simulation was performed using the Graphical

Processor Unit (GPU) version of the Particle Mesh Ewald Molecular Dynamics (PMEMD) engine in the AMBER18 suite coupled with integrated modules [69]. FF14SB forcefield was used in defining protein parameters. P131 parameterization was done using the ANTECHAMBER module [70]. Afterward, the topology and parameter files for P131 and the identified compounds complexes were generated with the LEAP module, which neutralizes the complexes by adding counter ions and solvates them in a 10 Å TIP3P water box. The systems were minimized partially at 2500 steps, with the restraint potential set at 500 kcal/mol$Å^2$. A full minimization followed this for 5000 steps without energy restraints. Heating of the systems took place for 50 ps from 0 to 300 k in an NVT canonical ensemble using a Langevin thermostat [71] and at a harmonic restraint of 5 kcal/mol $Å^2$. The systems were then equilibrated at 300 k for 1000 ps without energy restraints with Berendsen barostat, keeping the atmospheric pressure at 1 bar [72]. The MD production run was after that carried out for 200 ns [73]. The trajectories obtained were analyzed by the integrated *C*PPTRAJ and PTRAJ modules [74]. Origin data analytical tool was used in creating the needed plots [75]. 3D visualization of the structures and corresponding analyses were carried out on the GUI of UCSF Chimera.

6.3.5 Thermodynamics calculations

This calculation estimates the binding free energy (BFE) of P131 and the identified compounds to *Cp*IMPDH. BFE gives information about the stability and the binding free energy of a ligand bound to a protein [76]. In our study, Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA), which evaluates ligand interaction in biological macromolecules, was employed in estimating the BFE of the compounds [76, 77]. It is a very robust, widely used, and reliable analytical tool [78, 79]. MM/PBSA was calculated via AmberTools18-integrated MM/PBSA.py python script. This script employs continuum solvent models to automatically analyze binding free energies of snapshots from MD simulation [80].

Mathematically, binding free energy is depicted by the following equation;

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (1)$$

$$\Delta G_{sol} = \Delta G_{pol\ (PB)} + \Delta G_{np} \quad (2)$$

Therefore, $\Delta G_{bind} = \Delta E_{MM} + \Delta G_{pol\ (PB)} + \Delta G_{np}$ (3)

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (4)$$

ΔE_{int} is given by the summation of $\Delta E_{angle,}$ $\Delta E_{bond,}$ and $\Delta E_{torsion.}$ Molecular dynamics simulation was run on the complex only, and a single trajectory was used. This approach minimizes error and noise; therefore, ΔE_{int} was canceled between receptor, ligand, and complex.[77]

$$\Delta E_{MM} = \Delta E_{ele} + \Delta E_{vdW} \quad (5)$$

In the equation presented above, ΔE_{ele} and ΔE_{vdw} represent electrostatic and Vander Waal energy contributions, respectively. ΔE_{MM} is gas-phase energy. The solvation energy contribution is $\Delta G_{sol,}$ which is constituted by polar solvation energy contribution ($\Delta G_{pol\ (PB)}$) and non-polar solvation energy contribution (ΔG_{np}). ΔE_{int} represents the internal energy contribution, and $T\Delta S$, the conformational entropy change. To estimate the different energy each binding site residues contributes to the stabilization and affinity of the ligand, we analyzed the per residue decomposition (PRED), which, as earlier stated, was investigated by manually adding the binding residues of P131 and the hits generated to MMPB/SA.py script integrated with Amber18Tools

6.4 Results and Discussion

6.4.1 Per-residue-based energy decomposition upon P131 binding leads to the creation of a pharmacophore model for hit search.

In constructing our PRED-based pharmacophore model, a 20ns MD of P131 bound to *Cp*IMPDH was performed, and the total binding free energy was computed. The overall binding free energy was subsequently decomposed on a per residue basis to determine the energy contributed by each binding site residue to the stability of the protein-ligand complex [81]. The residues contributing considerably higher energy to the binding and stability of P131 were analyzed *vis-a-vis* the structural features of P131 they interacted with. These residues were deemed crucial to the interactions of P131 in *Cp*IMPDH. These residues and the respective binding free energies they contributed were ASP252 (-9.82 kcal/mol), SER164 (-2.62 kcal/mol) ALA165 (-1.65 kcal/mol) MET302 (-1.73 kcal/mol) and in the second chain (chain D), PRO26 (-1.28 kcal/mol) and TYR358 (-1.7 kcal/mol) . The interaction between these residues and P131 is shown in Figure 6.1. The moieties interacting with the residues contributing these energies made up the chemical scaffold

that was used in screening the ZINC database for potential compounds that might portend a better therapeutic outcome than P131 (Figure 6.1b).

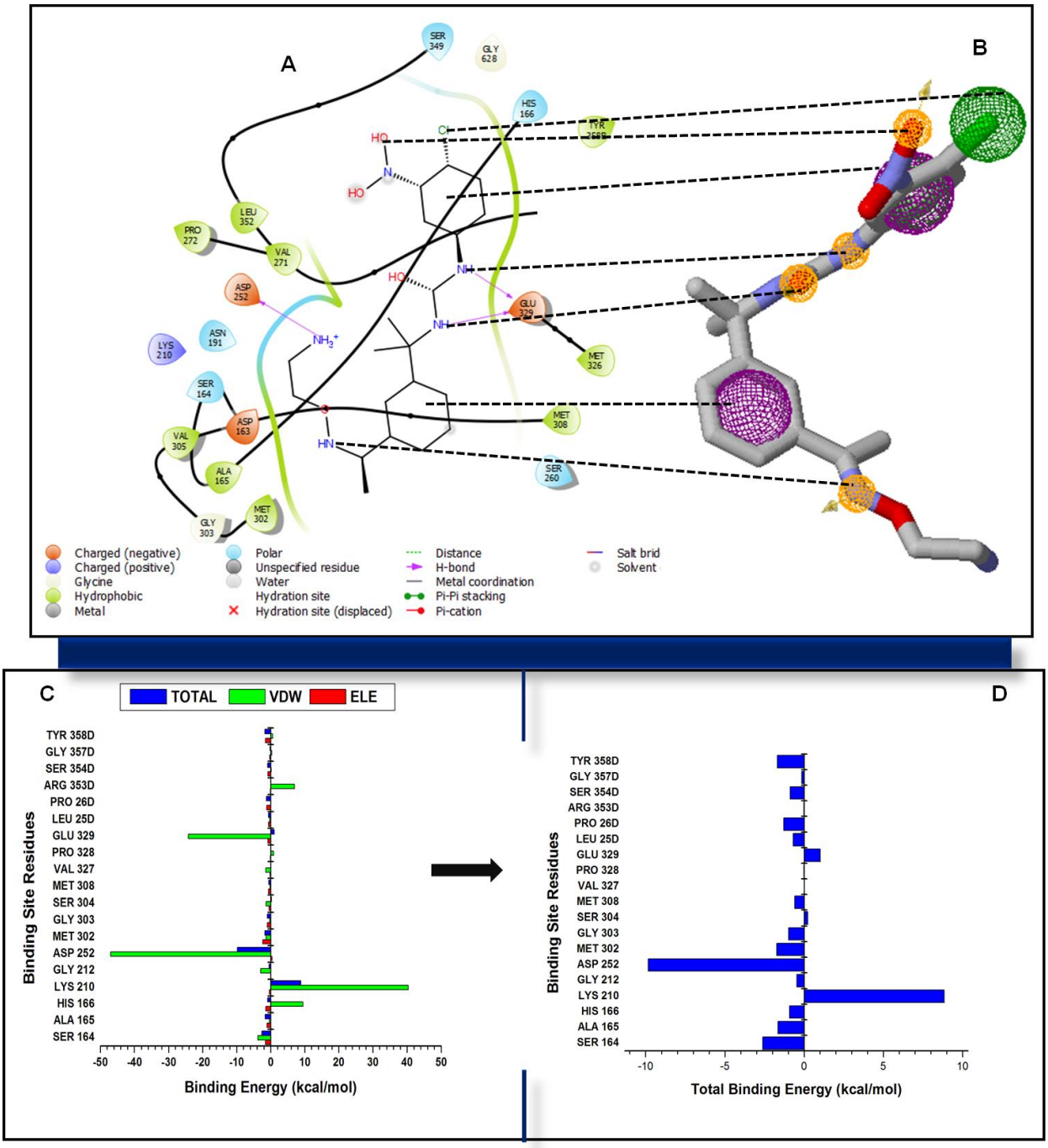
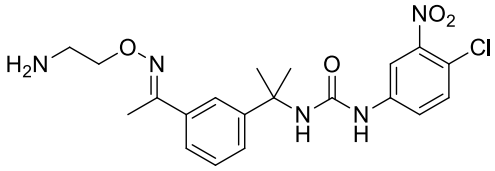
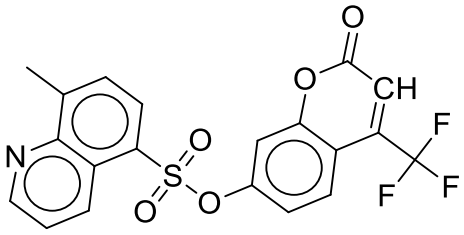
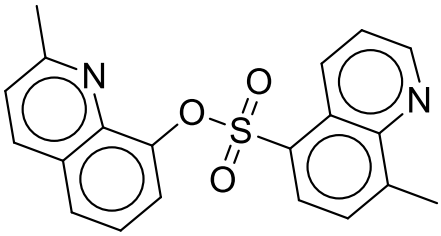
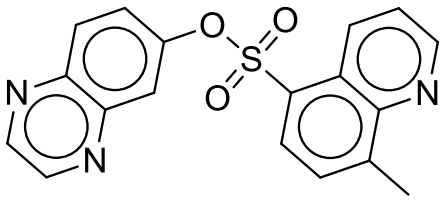


Figure 6.1 P131 pharmacophore model creation. A) 2D layout of binding site residue interaction with P131 after 20ns MD simulation. B) The pharmacophoric moieties selected in P131, which interacted with the highest energy contributing binding site residues. C) Per residue energy decomposition showing different energy components, each binding site residue contributed to the total binding free energy of P131. D) Per residue plot showing only the total binding free energy contributed by each binding site residue.

6.4.2 Pharmacophore-based virtual screening

Virtual screening has evolved as a rapid and dynamic approach to produce potential drug-like compounds that may be optimized to interact efficiently with its therapeutic target [82]. It provides an important starting point in drug discovery called lead identification [83, 84]. The moieties in P131 interacting with the critical amino acid residues were selected on the ZINCPharmer platform. These included two aromatic rings, four hydrogen bond acceptors, and three hydrophobic rings. This pharmacophore scaffold was submitted to ZINCPharmer for pharmacophore pattern matching. The parameters of the hits to be generated were set to have the following- rotatable bond of ≤ 5 molecular weight of $\leq$ 500Da. This was necessary to filter out non-drug-like hits that deviated from Lipinski's rule of 5. With the pharmacophoric moieties selected as well as the filtering parameters, a total of eight ZINC- compounds were outputted. Subsequently, these compounds were docked into the NAD^+ binding site of *Cp*IMPDH, and their binding affinity was scored. The top three of the identified compounds having the best docking scores were selected for further analysis. These compounds and their respective docking scores were ZINC46542062 (-8.3 kcal/mol), ZINC58646829 (-8.2 kcal/mol), and ZINC89780094 (-7.5 kcal/mol). ZINC46542062 and ZINC58646829 docking scores were better than P131, which was -7.9 kcal/mol. The 2D representation of the identified hits and their respective docking scores are presented in Table 1.

Table 6.1 2D representation of P131 and the identified compounds with their docking scores

Compound	2D representation	Docking scores (kcal/mol)
P131		-7.9
ZINC46542062		-8.3
ZINC58646829		-8.2
ZINC89780094		-7.5

6.4.3 The differential binding free energy profiles of P131 and the identified compounds in *Cp*IMPDH

Binding free energy calculations have evolved as a dynamic, inexpensive, and extensively used tool in estimating at the atomic level, the interactions between a therapeutic compound and its biological target. It elucidates the role of each binding site residues in the ligand's stability and binding. The binding free energies of the identified compounds were calculated after molecular dynamics simulation. We excluded the conformational entropy effect as it is computationally expensive to calculate by normal mode analysis [85]. Snapshots from 100ns-200ns exhibited relative stabilities and were sampled for the energy calculation to minimize conformational entropy. All the compounds presented with favorable binding free energies (Table 6.2). However, ZINC46542062 had the highest binding free energy of -39.52 kcal/mol, which was better than the reference compound P131 (-34.61 kcal/mol). This was followed by ZINC89780094 (-32.68 kcal/mol) and ZINC58646829 (-23.15 kcal/mol). Per-residue energy contribution to the overall total binding free energies of the identified compounds and prominent energy contributing residues for each complex are presented in Suppl Figure 6.1 – 6.3 and Suppl Table 6.1.

Table 6.2 Differential MM/PBSA binding free energies of *Cryptosporidium parvum* IMPDH in complex with P131 and the identified compounds

System	Energy components (kcal/mol)				
	ΔE_{vdw}	ΔE_{ele}	ΔE_{MM}	ΔG_{sol}	ΔG_{bind}
P131	-51.36 ±0.19	-74.65 ± 0.57	-126.± 0.48	91.4±0.45	-34.61±0.17
ZINC46542062	-45.75±0.09	-200.86±0.26	-246.61±0.2	207.08±0.23	-39.52±0.01
ZINC58646829	-38.45±0.27	-17.16±0.25	-55.6±0.41	32.22±0.26	-23.15±0.26
ZINC89780094	-38.74±0.22	-22.84±0.29	-61.59±0.41	28.91±0.20	-32.68±0.28

6.4.4 Exploring the interaction dynamics of P131 and the identified compounds across the simulation period

To further understand what could have accounted for the favorable interactions of P131 and the identified compounds evidenced by the binding free energies (Table 2), we visualized the molecular interactions taking place at different time points of the MD simulation. The snapshots were selected at 50ns, 100ns, 150ns, and 200ns, representing the initial, intermediate, and final time-points. In P131, at the initial stage of simulation (50ns), three conventional hydrogen bonds were formed by Glu329, Ser164, and Met302. This number thinned out as the simulation went on, leaving only Glu329 as the only consistent hydrogen-bond forming residue (Figure 6.2). Asp252 and Asp163 also were consistent in forming an attractive charge interaction during the period of simulation (Figure 6.2). Asp252 and Ser164 were among the highest energy contributing residues to the binding and stabilization of P131 in the NAD^+ binding site in *Cp*IMPDH (Suppl Table 6.1).

The binding sites' multiple interactions with the three fluorine atom of ZINC46542062 could have afforded it the highest binding free energy value than P131 and other ZINC compounds. At each time point, the fluorine atoms were engaged in forming a conventional hydrogen bond, fluorine, pi-alkyl, and pi-sigma interactions at almost the same time. TYR358D, ASP163, LYS73, and CYS219 formed a steady conventional hydrogen bond throughout the simulation while MET302 maintained a constant pi-sulfur interaction.

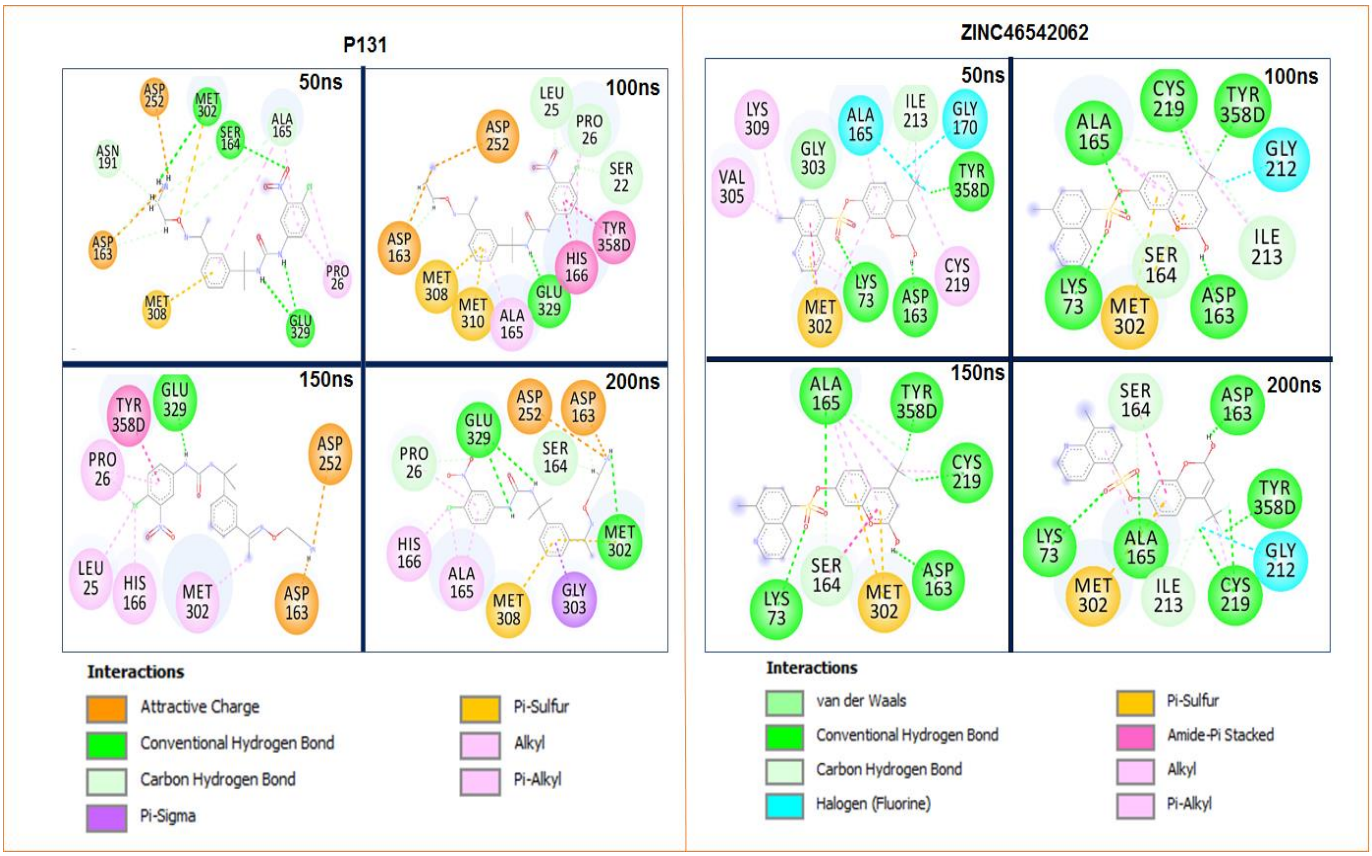


Figure 6.2 A comparative time-based interaction dynamics of P131 and ZINC46542062 at the NAD^+ binding site of *Cp*IMPDH. Time-points were selected at 50ns, 100ns, 150ns, and 200ns for both compounds.

The interaction dynamics of ZINC58646829 and ZINC89780094 are presented in Figure 6.3. In ZINC89780094, consistent conventional hydrogen bonds were formed by SER164, SER169, ASN171 in the entire course of the simulation. A favorable pi-sigma bond mediated by HIS 166 was also observed all through. On the converse, in ZINC58646829, weaker interactions were

observed when compared to other screened compounds. Alkyl and pi-alkyl interactions dominated the simulation period. The foregoing could have accounted for ZINC58646829 having the lowest binding free energy than other compounds examined (Table 2).

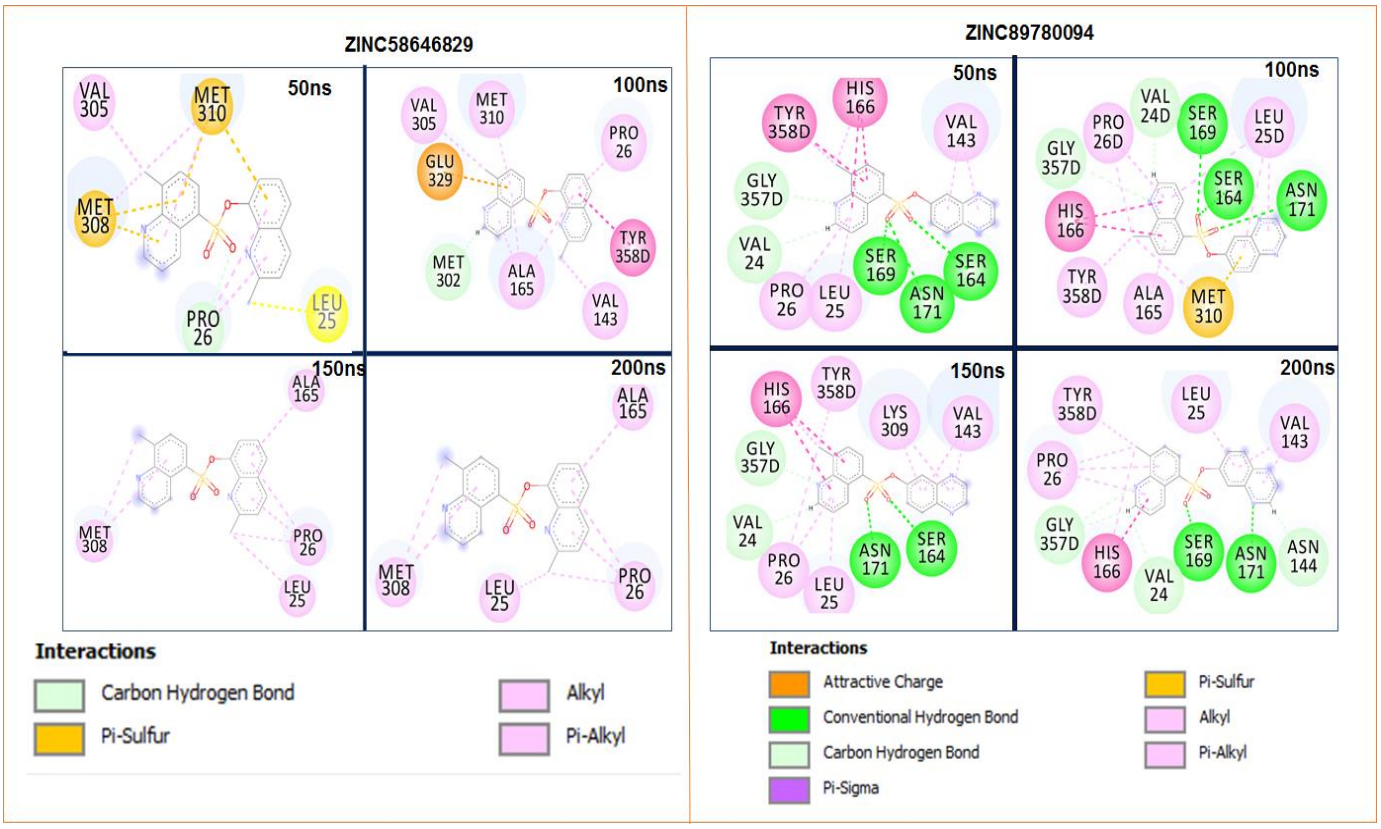


Figure 6.3 A comparative time-based interaction dynamics of ZINC58646829 and ZINC89780094 at the NAD^+ binding site of *Cp*IMPDH. Time points were selected at 50ns, 100ns, 150ns, and 200ns for both compounds.

It is also important to note that *Cryptosporidium parvum* though a eukaryotic organism, however through lateral gene transfer, it possesses a prokaryotic IMPDH [21, 89]. This permits anticryptosporidial selective inhibition of *Cp*IMPDH, leaving the eukaryotic hosts' IMPDH unaffected. There are two residues in the NAD^+ site of *Cp*IMPDH and prokaryotic IMPDH, responsible for anticryptosporidial drugs' selectivity. These are Ala165 in chain A, and Tyr 358 in chain D. These residues have been substituted in eukaryotic IMPDH [90]. We examined the consistency of these residues' interaction with P131 and the identified compounds at different time-points selected. Despite their constant motion, the ligands interacted with at least one of these critical residues (Figures 6.2 & 6.3). This interaction further underpins their importance in the potentiation of the therapeutic action of potential *Cp*IMPDH inhibitors.

6.4.5 Evaluation of the drug-likeness and pharmacokinetics of P131 and the identified compounds

Early in drug discovery stages, drug lead pharmacokinetics properties must be determined to funnel down to compounds that can be optimized to a favorable range of optimal absorption, distribution, metabolism, excretion, and toxicity (ADMET). 60% of drug failures have been attributed to ADMET challenges [91]. The conventional rules of 5 by Lipinski summarily posits that, for a drug to be orally active, it must violate not more than two criteria from the following: not more than 10 hydrogen bond acceptors and 5 hydrogen bond donors, octanol-water partition effect of not greater than 5, lipophilicity not greater than 5 and a molecular weight of less than 500g/mol [54, 92]. Going by the rule above, the molecular weight of P131 and the identified compounds were within the range of acceptable limits (≤500). ZINC46542062 had the highest (435.37 g/mol), while ZINC89780094 had the lowest (351.38 g/mol).

One of the vital indices in determining the pharmacokinetic properties of drugs is lipophilicity (logP). For a drug to potentiate its action, it must interact with its specific target protein. In doing this, it must cross cell membranes, which are almost always lipophilic; therefore, the drug must be soluble in non- polar solvents, fats, and lipids [93]. Lipophilicity impacts drug potency, solubility, permeability, selectivity, and toxicology [94–98]. Lipinski places the logP benchmark as 5 [92]. Compounds with logP higher than 5 have higher toxicity, poorer absorption, lower solubility, and excretion [99, 100]. Also, a very low logP value of compounds presents with ADMET challenges [93]. Generally, drug leads having a logP ranging from 2 to 3 are favorable.[101] ZINC89780094 had the lowest value of LogP of 2.45; however, the parent drug P131 and the other two drug leads had >3 value of LogP (Table 3). This implies that ZINC89780094 has a higher propensity for cell membrane permeability and bioavailability. Beyond Lipinski's rules, other parameters have been established to play a role in estimating the drug-likeness of compounds. These parameters include ligand efficiency (LE), ligand lipophilic efficiency (LLE), and ligand-efficiency dependent lipophilicity (LELP)[102]. Both LLE and LELP take logP into account while LE does not [93]. LE is used in estimating the affinity of a ligand to its therapeutic target [103]. It also calculates and compares the potency of compounds *vis-a-vis* their molecular size [104, 105].

Table 6.3 Comparative evaluation of the drug-likeness of P131 and the identified compounds

Parameters	P131	ZINC46542062	ZINC58646829	ZINC89780094	Acceptable Limit
Molecular Formular	C19H24CIN503	C20H12F3NO5S	C20H16N2O3S	C18H13N3O3S	
Parameters predicted accurately					
Molecular weight (g/mol)	405.88	435.37	364.42	351.38	≤500
H-Bond Acceptor	8	6	5	6	≤10
H-Bond Donor	4	0	0	0	≤5
Rotatable Bonds	10	4	3	3	≤10
Parameters predicted to reasonable accuracy					
Lipophilicity (logP)	3.19	3.20	3.11	2.45	≤5
Water Solubility	Moderately soluble	Moderately soluble	Moderately soluble	Moderately soluble	
TPSA (A^2)	117.5	86.48	69.16	82.05	≤140
Parameters roughly estimated					
LE (kcal/mol/heavy atom)	0.44	0.41	0.47	0.49	> 0.3
LLE	10.19	5.4	5.09	6.36	> 5
LELP	-2.70	8.74	8.24	5.34	-10 to +10
Bioavailability Score	0.55	0.55	0.55	0.55	
LD_{50} (mg/kg)	2000	2662	10750	11500	
Synthetic Accessibility	3.85	3.34	3.06	2.89	10
GI Absorption	High	Low	High	High	
BBB Permeability	No	No	No	No	

Though it can account for the potency irrespective of their molecular sizes, however, not factoring their lipophilicity has its drawbacks as employing LE as an efficiency index alone can lead to the selection of compounds with higher potency but low ADMET characteristics [106–108]. From our results, P131 and the identified compounds are within the LE range, which is >0.3 kcal/mol/heavy atom (Table 3). This indicates that they have desired potency at the right weight, and their ADMET properties may further be optimized without losing their potencies.

LLE measures both the potency of the compounds and its lipophilicity while excluding its molecular size[95, 109]. It exploits how the potency of compounds can be improved while at the same time maintaining low lipophilicity.[95] With the optimal value set at >5 [95], P131 had the highest LLE value of 10.31, followed by ZINC89780094, ZINC46542062, and ZINC58646829, which had 6.36, 5.4, and 5.09, respectively (Table 3). The import of the result is that the compounds can serve as a suitable starting material for further optimization and development.

LELP, on the other hand, considers the molecular size, potency, and lipophilicity of compounds as a composite unit [102]. It overcomes the lipophilicity exclusion of LE and the molecular size exclusion of LLE. It has been found that drugs having a suboptimal value of LELP did not go far in the drug development process as their drug-likeness is negatively affected [102]. P131 and the identified compounds all had LELP value within the acceptable range of LELP (Table 3), with the lowest being P131 (-2.7kcal/mol) and the highest as ZINC46542062 (8.7 kcal/mol).

Synthetic accessibility score (SA score) is another parameter that evaluates how easy or difficult it is to synthesize a compound [110]. SA score must be determined early as a compound that might have been assessed to have excellent ADMET properties might be challenging to synthesize. SA score ranges from 1 (easy to synthesize) to 10 (difficult to synthesize).[110] From the four compounds assessed, P131 was the hardest to synthesize with an SA score of 3.85, followed by ZINC46542062, ZINC58646829, and ZINC89780094 with SA score of 3.34, 3.06, and 2.89, respectively (Table 3). Further works involve evaluating the topological polar surface area (TPSA) metric of the compounds that measure the contribution of oxygen, nitrogen, and hydrogens to the compound molecular surface area [111]. It gives information about the ability of compounds to permeate the cells. Therefore, it has been postulated that the lower the TPSA value, the easier will be the assimilation of the compounds into the cell [112–114]. Although all the compounds fell within the range of allowable TPSA ($140A^2$) [115], P131 will have some difficulty in permeating

the cell when compared to the hits as it has the highest TPSA ($117.5A^2$) while ZINC58646829 has the lowest ($69.16A^2$) (Table 3). The value of TPSA of P131 is not far removed from the higher number of hydrogen bond donors and acceptors it possesses compared to other identified compounds as the number of hydrogen bonding bears a correlation with TPSA [114].

Finally, the toxicological assessment of the compounds was done. This parameter is presented as LD_{50}. It has been established that the lower the LD_{50}, the more toxic the drug and otherwise [64]. Extrapolating this to P131 and the identified compounds, P131 has the lowest LD_{50} (2000mg/kg), making it the most potentially toxic of all the hits, closely followed by ZINC46542062, which has 2662mg/kg. Both ZINC58646829 and ZINC89780094 appear safer, having an LD_{50} value of 10750mg/kg and 11500mg/kg, respectively (Table 3). All the physicochemical properties that have been evaluated and their established upper limits, both P131 and the identified compounds, have favorable ADMET properties but can still benefit from some improvements.

6.4.6 Conformational perturbations induced by the binding of P131 and the identified compounds

When ligands bind to proteins, they induce conformational changes that that may consequently affect the functions of the proteins [116, 117]. The structural changes that occur when P131 and the identified compounds bind to *Cp*IMPDH were evaluated by calculating the root mean square deviation (RMSD), the root mean square fluctuation (RMSF), and the radius of gyration (RoG) of the generated MD trajectories.

RMSD has evolved as a parameter for comparing protein structure, characterizing related proteins conformation, and stability. It also assesses the quality of molecular dynamics simulation [118, 119]. It uses the alpha carbon of each amino acid residues that make up the protein to estimate the conformational changes during the simulation with reference to its initial structural conformation [118]. From the RMSD graph (Figure 6.4a), all systems attained equilibrium at around 30ns and exhibited variations in their stability throughout the simulation time. The binding of P131 and the identified hits stabilized the structure of *Cp*IMPDH when compared to the apo. The RMSD of the unliganded *Cp*IMPDH was 2.92Å; the binding of P131 to *Cp*IMPDH minimally stabilized the system with RMSD value of 2.84 Å. However, ZINC46542062, ZINC58646829, and

ZINC89780094 stabilized it more than P131 with an RMSD value of 2.47Å, 2.24Å, and 2.48Å, respectively.

The root mean square fluctuations estimates the flexibility of the protein structure [120]. This it does by measuring the fluctuation of the C-alpha atom of each of the amino acid residues that make up the protein. It is interesting to note that in the apo and complexes, the highest residue fluctuation occurred at chain B (Figure 6.4b). The unbound system had an average RMSF value of 12.57Å, while P131-*Cp*IMPDH had 8.57Å. The average RMSF values of *Cp*IMPDH complexed with the identified compounds followed that of the apo closely. *Cp*IMPDH- ZINC46542062 had 12.40Å while *Cp*IMPDH- ZINC89780094 had 12.79Å. However, *Cp*IMPDH- ZINC58646829 had the highest average residue fluctuation of 13.89Å. The discrepancy in the RMSF values of P131 relative to the identified compounds might stem from the size of P131, which is bigger than the other ligands, and the fact that it bends at the linker region, which affords it more allowance to interact with more residues in the adjacent chain of *Cp*IMPDH (chain B), thereby reducing the fluctuation of that particular chain.

Closely related to the RMSD is the RoG, which also provides insight into the protein structure's stability by determining how compact or loose the protein structure is. This is estimated by measuring the distance of the individual atom of the protein to their centroid. A high RoG value indicates that the protein structure is packed loosely while a lower value portrays the otherwise. The binding of the P131and the identified compounds made *Cp*IMPDH a bit loose when compared with the apo (27.13Å). The binding of P131 to *Cp*IMPDH increased the protein's RoG to 28.01Å. ZINC58646829 impacted the RoG of *Cp*IMPDH the most, with an average value of 28.13Å. ZINC46542062 had 28Å, while ZINC89780094 had the least among the ligands with an average value of 27.96Å (Figure 6.4c). The slight difference in the RoG values of the apo and the complexes shows an insignificant structural deviation of the complexes from the native structure. This conclusion also lends credence to the information given by e average RMSD values of the apo and the complexes. Therefore, we can safely postulate from the RMSD and RoG indices that the identified compounds have almost the same effect on CpIMPDH compared to P131, which implies that they might also potentiate their action by the same structural inhibitory mechanism.

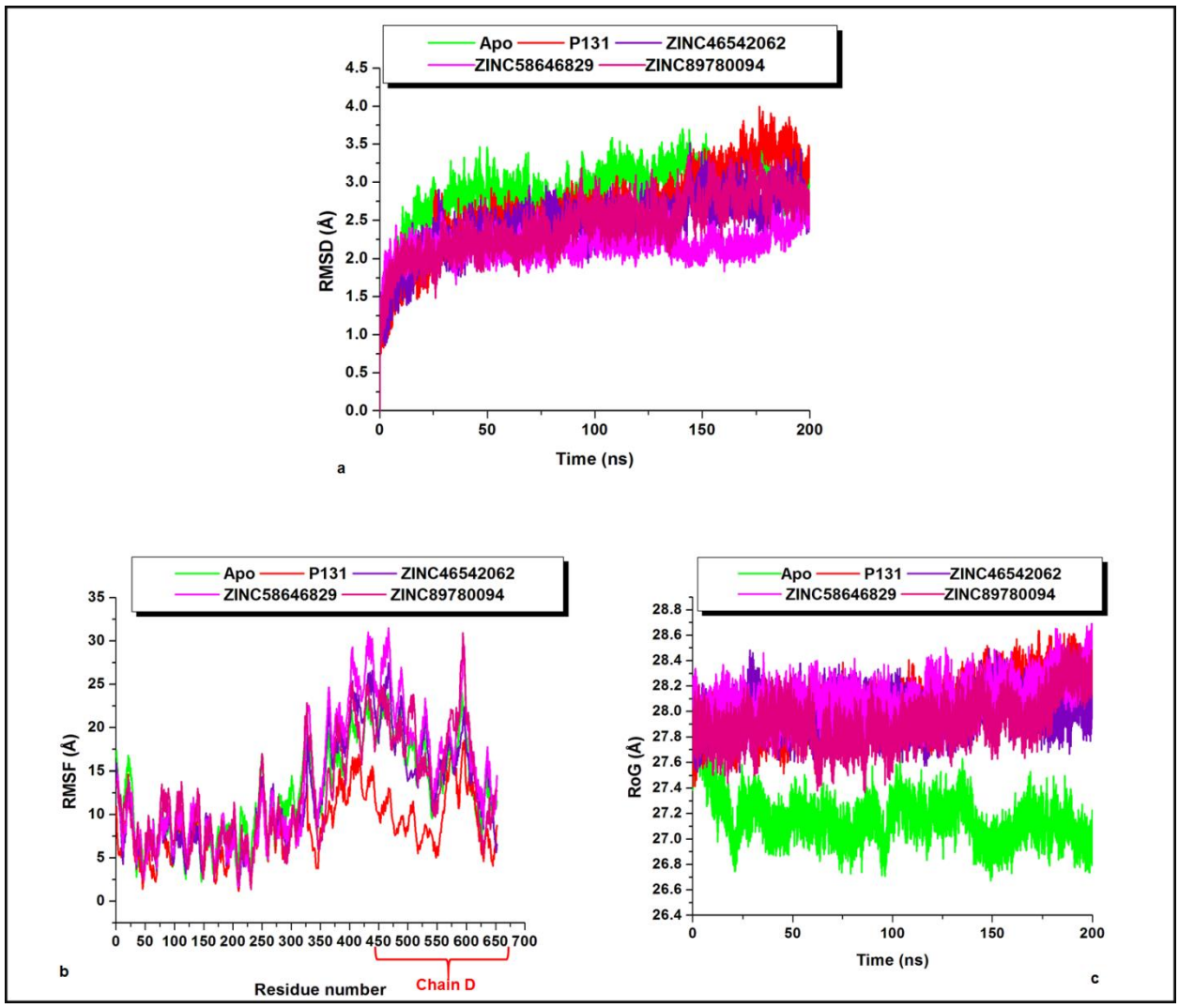


Figure 6.4 A representation of the structural alterations mediated by the binding of P131 and the identified hits compounds to *Cp*IMPDH. We estimated across the simulation period a) Root mean square deviation (RMSD). b)Root mean square of fluctuation (RMSF) and c) Radius of gyration (RoG).

6.5 Conclusions

The limited efficacy of nitazoxanide in treating cryptosporidiosis led to the synthesis of P131, which inhibits *Cp*IMPDH. Inferring from the total binding free energy of P131 and the identified compounds, ZINC46542062 had a better binding free energy to *Cp*IMPDH than P131, suggesting ZINC46542062 might have better inhibitory potential than P131. Also, from the evaluation of their pharmacokinetic parameters, the identified compounds were more therapeutically suitable than P131, especially in their toxicity (LD_{50}) index, synthetic accessibility, lipophilicity, and the number of rotatable bonds, hydrogen bonds, and acceptors, which altogether impact the permeability of compounds into the cell. On the other hand, P131 and the identified compounds influenced the conformational dynamics of *Cp*IMPDH, almost similarly with negligible variations. Insight from this study could serve as a starting point in designing new inhibitors for *Cp*IMPDH with improved therapeutic properties.

Acknowledgment

The authors appreciate Centre for High Performance Computing (CHPC) , Cape Town, South Africa for making computational resources available.

Declarations

Funding

The authors did not receive support from any organization for the submitted work.

Conflict of interest.

The authors declare no conflict of interest.

Ethics approval

N/A

Consent to participate

N/A

Consent for publication

N/A

Availability of data and material

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Code availability

N/A

Authors' contributions

Conceptualization, Methodology, Formal analysis and investigation, Writing - original draft preparation: [Kehinde F. Omolabi]; Methodology: [Emmanuel A. Iwuchukwu]; Writing - review and editing: [Clement Agoni]; Writing - review and editing: [Fisayo A. Olotu]; Supervision: [Mahmoud E.S Soliman].

6.6 References

1. Current WL, Garcia LS (1991) Cryptosporidiosis. Clin. Microbiol. Rev. 4:325–358

2. Clark DP (1999) New insights into human Cryptosporidiosis. Clin. Microbiol. Rev. 12:554–563

3. Shirley D-AT, Moonah SN, Kotloff KL (2012) Burden of disease from cryptosporidiosis. Curr Opin Infect Dis 25:555–63. https://doi.org/10.1097/QCO.0b013e328357e569

4. Checkley W, Jr. AW, Jaganath D (2015) A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. 15

5. Chappell CL, Okhuysen PC, Sterling CR, et al (1999) Infectivity of Cryptosporidium parvum in healthy adults with pre-existing anti-C. parvum serum immunoglobulin G. Am J Trop Med Hyg 60:157–64. https://doi.org/10.4269/ajtmh.1999.60.157

6. Mac Kenzie WR, Hoxie NJ, Proctor ME, et al (1994) A massive outbreak in milwaukee of

cryptosporidium infection transmitted through the public water supply. N Engl J Med 331:161–167. https://doi.org/10.1056/NEJM199407213310304

7. Molbak K, Hojlyng N, Gottschau A, et al (1993) Cryptosporidiosis in infancy and childhood mortality in Guinea Bissau, West Africa. Br Med J 307:417–420. https://doi.org/10.1136/bmj.307.6901.417

8. Kotloff KL, Nataro JP, Blackwelder WC, et al (2013) Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. Lancet 382:209–222. https://doi.org/10.1016/S0140-6736(13)60844-2

9. Newman RD, Sears CL, Moore SR, et al (1999) Longitudinal Study of Cryptosporidium Infection in Children in Northeastern Brazil . J Infect Dis 180:167–175. https://doi.org/10.1086/314820

10. Khan A, Shams S, Khan S, et al (2019) Evaluation of prevalence and risk factors associated with Cryptosporidium infection in rural population of district Buner, Pakistan. PLoS One 14:. https://doi.org/10.1371/journal.pone.0209188

11. Squire SA, Ryan U (2017) Cryptosporidium and Giardia in Africa: current and future challenges. Parasites and Vectors 10:1–32. https://doi.org/10.1186/s13071-017-2111-y

12. Vanathy K, Parija SC, Mandal J, et al (2017) Cryptosporidiosis: A mini review. Trop Parasitol 7:72. https://doi.org/10.4103/TP.TP_25_17

13. Checkley W, White AC, Jaganath D, et al (2015) A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. Lancet Infect. Dis. 15:85–94

14. Sparks H, Nair G, Castellanos-Gonzalez A, White AC (2015) Treatment of Cryptosporidium: What We Know, Gaps, and the Way Forward. Curr Trop Med Reports 2:181–187. https://doi.org/10.1007/s40475-015-0056-9

15. Amadi B, Mwiya M, Sianongo S, et al (2009) High dose prolonged treatment with nitazoxanide is not effective for cryptosporidiosis in HIV positive Zambian children: A randomised controlled trial. BMC Infect Dis 9:. https://doi.org/10.1186/1471-2334-9-195

16. Amenta M, Dalle Nogare ER, Colomba C, et al (1999) Intestinal protozoa in HIV-infected patients: effect of rifaximin in Cryptosporidium parvum and Blastocystis hominis infections. J Chemother 11:391–5. https://doi.org/10.1179/joc.1999.11.5.391

17. Hicks P, Zwiener RJ, Squires J, Savell V (1996) Azithromycin therapy for Cryptosporidium parvum infection in four children infected with human immunodeficiency virus. J Pediatr 129:297–300. https://doi.org/10.1016/S0022-3476(96)70258-5

18. Gathe JC, Mayberry C, Clemmons J, Nemecek J (2008) Resolution of severe cryptosporidial diarrhea with rifaximin in patients with AIDS. J. Acquir. Immune Defic. Syndr. 48:365–366

19. Paromomycin Oral : Uses, Side Effects, Interactions, Pictures, Warnings & Dosing - WebMD. https://www.webmd.com/drugs/2/drug-5160/paromomycin-oral/details#side-effects. Accessed 28 Jun 2020

20. White AC, Chappell CL, Valdez LM, et al Combination Drug Therapy for Cryptosporidiosis in AIDS

21. Umejiego NN, Li C, Riera T, et al (2004) Cryptosporidium parvum IMP dehydrogenase: Identification of functional, structural, and dynamic properties that can be exploited for drug design. J Biol Chem 279:40320–40327. https://doi.org/10.1074/jbc.M407121200

22. Gorla SK, Kavitha M, Zhang M, et al (2012) Selective and potent urea inhibitors of cryptosporidium parvum inosine 5′-monophosphate dehydrogenase. J Med Chem 55:7759–7771. https://doi.org/10.1021/jm3007917

23. Gorla SK, McNair NN, Yang G, et al (2014) Validation of IMP dehydrogenase inhibitors in a mouse model of cryptosporidiosis. Antimicrob Agents Chemother 58:1603–1614. https://doi.org/10.1128/AAC.02075-13

24. Gorla SK, Kavitha M, Zhang M, et al (2013) Optimization of benzoxazole-based inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. J Med Chem 56:4028–43. https://doi.org/10.1021/jm400241j

25. Johnson CR, Gorla SK, Kavitha M, et al (2013) Phthalazinone inhibitors of inosine-5'-monophosphate dehydrogenase from Cryptosporidium parvum. Bioorg Med Chem Lett 23:1004–7. https://doi.org/10.1016/j.bmcl.2012.12.037

26. Kirubakaran S, Gorla SK, Sharling L, et al (2012) Structure-activity relationship study of selective benzimidazole-based inhibitors of Cryptosporidium parvum IMPDH. Bioorg Med Chem Lett 22:1985–8. https://doi.org/10.1016/j.bmcl.2012.01.029

27. Macpherson IS, Kirubakaran S, Gorla SK, et al (2010) The structural basis of Cryptosporidium -specific IMP dehydrogenase inhibitor selectivity. J Am Chem Soc 132:1230–1. https://doi.org/10.1021/ja909947a

28. Maurya SK, Gollapalli DR, Kirubakaran S, et al (2009) Triazole inhibitors of Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase. J Med Chem 52:4623–30. https://doi.org/10.1021/jm900410u

29. Sharling L, Liu X, Gollapalli DR, et al (2010) A screening pipeline for antiparasitic agents targeting cryptosporidium inosine monophosphate dehydrogenase. PLoS Negl Trop Dis 4:e794. https://doi.org/10.1371/journal.pntd.0000794

30. Sun Z, Khan J, Makowska-Grzyska M, et al (2014) Synthesis, in vitro evaluation and cocrystal structure of 4-oxo-[1]benzopyrano[4,3-c]pyrazole Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase (CpIMPDH) inhibitors. J Med Chem 57:10544–50. https://doi.org/10.1021/jm501527z

31. Umejiego NN, Gollapalli D, Sharling L, et al (2008) Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against cryptosporidiosis. Chem Biol 15:70–7. https://doi.org/10.1016/j.chembiol.2007.12.010

32. Kim Y, Makowska-Grzyska M, Gorla SK, et al (2015) Structure of Cryptosporidium IMP

dehydrogenase bound to an inhibitor with in vivo antiparasitic activity. Acta Crystallogr Sect FStructural Biol Commun 71:531–538. https://doi.org/10.1107/S2053230X15000187

33. Makowska-Grzyska M, Kim Y, Maltseva N, et al (2015) A novel cofactor-binding mode in bacterial IMP dehydrogenases explains inhibitor selectivity. J Biol Chem 290:5893–5911. https://doi.org/10.1074/jbc.M114.619767

34. Hedstrom L (2009) IMP dehydrogenase: Structure, mechanism, and inhibition. Chem Rev 109:2903–2928. https://doi.org/10.1021/cr900021w

35. Felczak K, Chen L, Wilson D, et al (2011) Cofactor-type inhibitors of inosine monophosphate dehydrogenase via modular approach: Targeting the pyrophosphate binding sub-domain. Bioorganic Med Chem 19:1594–1605. https://doi.org/10.1016/j.bmc.2011.01.042

36. ALLISON AC, KOWALSKI WJ, MULLER CD, EUGUI EM (1993) Mechanisms of Action of Mycophenolic Acid. Ann N Y Acad Sci 696:63–87. https://doi.org/10.1111/j.1749-6632.1993.tb17143.x

37. Allison AC, Eugui EM (2000) Mycophenolate mofetil and its mechanisms of action. Immunopharmacology 47:85–118. https://doi.org/10.1016/S0162-3109(00)00188-0

38. Hassan Baig M, Ahmad K, Roy S, et al (2016) Computer Aided Drug Design: Success and Limitations. Curr Pharm Des 22:572–581. https://doi.org/10.2174/1381612822666151125000550

39. Kaalia R, Kumar A, Srinivasan A, Ghosh I (2015) An Ab Initio Method for Designing Multi-Target Specific Pharmacophores using Complementary Interaction Field of Aspartic Proteases. Mol Inform 34:380–393. https://doi.org/10.1002/minf.201400157

40. Qing X, Lee XY, De Raeymaeker J, et al (2014) Pharmacophore modeling: Advances, Limitations, And current utility in drug discovery. J Receptor Ligand Channel Res 7:81–92. https://doi.org/10.2147/JRLCR.S46843

41. Kaserer T, Beck KR, Akram M, et al (2015) Pharmacophore models and pharmacophore-based virtual screening: Concepts and applications exemplified on hydroxysteroid dehydrogenases. Molecules 20:22799–22832. https://doi.org/10.3390/molecules201219880

42. Dror O, Schneidman-Duhovny D, Inbar Y, et al (2009) Novel approach for efficient pharmacophore-based virtual screening: Method and applications. J Chem Inf Model 49:2333–2343. https://doi.org/10.1021/ci900263d

43. Richmond NJ, Abrams CA, Wolohan PRN, et al (2006) GALAHAD: 1. Pharmacophore identification by hypermolecular alignment of ligands in 3D. J Comput Aided Mol Des 20:567–587. https://doi.org/10.1007/s10822-006-9082-y

44. Jones G, Willett P, Glen RC (1995) A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. J Comput Aided Mol Des 9:532–549. https://doi.org/10.1007/BF00124324

45. Leach AR, Gillet VJ, Lewis RA, Taylor R (2010) Three-dimensional pharmacophore methods in drug discovery. J. Med. Chem. 53:539–558

46. Simulations B (2019) Learning the Footprints and Fingerprints : Pharmacophore Modeling in the Discovery of Potential Drug Candidates. https://doi.org/10.3844/jobsp.2019.

47. Kumalo HM, Soliman ME (2016) Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel β-Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. Cell Mol Bioeng 9:175–189. https://doi.org/10.1007/s12195-015-0421-8

48. Cele FN, Ramesh M, Soliman MES (2016) Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: A study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. Drug Des Devel Ther 10:1365–1377. https://doi.org/10.2147/DDDT.S95533

49. Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera - A visualization system

for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084

50. Eswar N, Webb B, Marti-Renom MA, et al (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinforma Chapter 5:Unit-5.6. https://doi.org/10.1002/0471250953.bi0506s15

51. ZINCPharmer: pharmacophore search of the ZINC database. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3394271/. Accessed 29 Apr 2020

52. Irwin JJ, Shoichet BK (2005) ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. J Chem Inf Model 45:177. https://doi.org/10.1021/CI049714

53. Daina A, Michielin O, Zoete V (2017) SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7:1–13. https://doi.org/10.1038/srep42717

54. Lipinski CA (2000) Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 44:235–249. https://doi.org/10.1016/S1056-8719(00)00107-6

55. Ripphausen P, Nisius B, Peltason L, Bajorath J (2010) Quo vadis, virtual screening? A comprehensive survey of prospective applications. J Med Chem 53:8461–8467. https://doi.org/10.1021/jm101020z

56. Trott O, Olson A (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455–461. https://doi.org/10.1002/jcc.21334.AutoDock

57. Forli S, Huey R, Pique ME, et al (2016) Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nat Protoc 11:905–919. https://doi.org/10.1038/nprot.2016.051

58. Morris GM, Ruth H, Lindstrom W, et al (2009) Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30:2785–2791. https://doi.org/10.1002/jcc.21256

59. Schrödinger Release 2020-3: Maestro, Schrödinger, LLC, New York, NY 2020. No Title

60. Karthick V, Nagasundaram N, Doss CGP, et al (2016) Virtual screening of the inhibitors targeting at the viral protein 40 of Ebola virus. Infect Dis Poverty 5:. https://doi.org/10.1186/s40249-016-0105-1

61. Lawal M, Olotu FA, Soliman MES (2018) Across the blood-brain barrier: Neurotherapeutic screening and characterization of naringenin as a novel CRMP-2 inhibitor in the treatment of Alzheimer's disease using bioinformatics and computational tools. Comput Biol Med 98:168–177. https://doi.org/10.1016/J.COMPBIOMED.2018.05.012

62. Liao C, Sitzmann M, Pugliese A, Nicklaus MC (2011) Software and resources for computational medicinal chemistry. Future Med. Chem. 3:1057–1085

63. Molinspiration Cheminformatics. https://www.molinspiration.com/. Accessed 1 May 2020

64. Drwal MN, Banerjee P, Dunkel M, et al (2014) ProTox: A web server for the in silico prediction of rodent oral toxicity. Nucleic Acids Res 42:3–8. https://doi.org/10.1093/nar/gku401

65. Sander T, Freyss J, Von Korff M, Rufener C (2015) DataWarrior: An open-source program for chemistry aware data visualization and analysis. J Chem Inf Model 55:460–473. https://doi.org/10.1021/ci500588j

66. Agoni C, Munsamy G, Ramhrack P, Soliman M (2020) Human Rhinovirus Inhibition Through Capsid "Canyon" Perturbation: Structural Insights into The Role of a Novel Benzothiophene Derivative. Cell Biochem Biophys 78:3–13

67. Agoni C, Salifu EY, Munsamy G, et al (2019) CF3-pyridinyl substitution on anti-malarial therapeutics: Probing differential ligand binding and dynamical inhibitory effects of a novel

triazolopyrimidine-based inhibitor on Plasmodium falciparum Dihydroorotate dehydrogenase. Chem Biodivers. https://doi.org/10.1002/cbdv.201900365

68. Olotu FA, Soliman MES (2019) Dynamic perspectives into the mechanisms of mutation-induced p53-DNA binding loss and inactivation using active perturbation theory: Structural and molecular insights toward the design of potent reactivators in cancer therapy. J Cell Biochem 120:951–966. https://doi.org/10.1002/jcb.27458

69. Nair PC, Miners JO (2014) Molecular dynamics simulations: from structure function relationships to drug discovery. Silico Pharmacol 2:. https://doi.org/10.1186/s40203-014-0004-8

70. Wang J, Wolf RM, Caldwell JW, et al (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–74. https://doi.org/10.1002/jcc.20035

71. Grest GS, Kremer K (1986) Molecular dynamics simulation for polymers in the presence of a heat bath. Phys Rev A 33:3628–3631. https://doi.org/10.1103/PhysRevA.33.3628

72. Berendsen HJC, Postma JPM, Van Gunsteren WF, et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690. https://doi.org/10.1063/1.448118

73. Ryckaert J-P, Ciccotti+ G, Berendsen HJC (1977) Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes

74. Roe DR, Cheatham TE (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 9:3084–3095. https://doi.org/10.1021/ct400341p

75. Seifert E (2014) OriginPro 9.1: scientific data analysis and graphing software-software review. J Chem Inf Model 54:1552. https://doi.org/10.1021/ci500161d

76. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10:449–61. https://doi.org/10.1517/17460441.2015.1032936

77. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82. https://doi.org/10.1021/ci100275a

78. Chaudhary N, Aparoy P (2017) Deciphering the mechanism behind the varied binding activities of COXIBs through Molecular Dynamic Simulations, MM-PBSA binding energy calculations and per-residue energy decomposition studies. J Biomol Struct Dyn 35:868–882. https://doi.org/10.1080/07391102.2016.1165736

79. Gupta A, Chaudhary N, Aparoy P (2018) MM-PBSA and per-residue decomposition energy studies on 7-Phenyl-imidazoquinolin-4(5H)-one derivatives: Identification of crucial site points at microsomal prostaglandin E synthase-1 (mPGES-1) active site. Int J Biol Macromol 119:352–359. https://doi.org/10.1016/j.ijbiomac.2018.07.050

80. Case DA (2018) Amber 18. Univ California, San Fr

81. Woods CJ, Malaisree M, Michel J, et al (2014) Rapid decomposition and visualisation of protein-ligand binding free energies by residue and by water. Faraday Discuss 169:477–499. https://doi.org/10.1039/c3fd00125c

82. Shoichet BK (2004) Virtual screening of chemical libraries. Nature 432:862–865

83. Kenakin T (2003) Predicting therapeutic value in the lead optimization phase of drug discovery. Nat. Rev. Drug Discov. 2:429–438

84. Huber W (2005) A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions. J Mol Recognit 18:273–281. https://doi.org/10.1002/jmr.744

85. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R et al (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688

86. Yang T, Wu JC, Yan C, et al (2011) Virtual screening using molecular simulations. Proteins

Struct Funct Bioinforma 79:1940–1951. https://doi.org/10.1002/prot.23018

87. Wallnoefer HG, Liedl KR, Fox T (2011) A challenging system: Free energy prediction for factor Xa. J Comput Chem 32:1743–1752. https://doi.org/10.1002/jcc.21758

88. Weis A, Katebzadeh K, Söderhjelm P, et al (2006) Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field. J Med Chem 49:6596–6606. https://doi.org/10.1021/jm0608210

89. Striepen B, Pruijssers AJP, Huang J, et al (2004) Gene transfer in the evolution of parasite nucleotide biosynthesis. Proc Natl Acad Sci U S A 101:3154–3159. https://doi.org/10.1073/pnas.0304686101

90. Hedstrom L, Liechti G, Goldberg JB, Gollapalli DR (2011) The antibiotic potential of prokaryotic IMP dehydrogenase inhibitors. Curr Med Chem 18:1909–18. https://doi.org/10.2174/092986711795590129

91. Tetko I V (2002) ADMETox data, models and challenges

92. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26. https://doi.org/10.1016/S0169-409X(00)00129-0

93. Arnott JA, Planey SL (2012) The influence of lipophilicity in drug discovery and design. Expert Opin Drug Discov 7:863–875. https://doi.org/10.1517/17460441.2012.714363

94. Liu X, Testa B, Fahr A (2011) Lipophilicity and its relationship with passive drug permeation. Pharm. Res. 28:962–977

95. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 6:881–890. https://doi.org/10.1038/nrd2445

96. Waring MJ (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower log D limits based on permeability.

Bioorganic Med Chem Lett 19:2844–2851. https://doi.org/10.1016/j.bmcl.2009.03.109

97. Gleeson MP, Hersey A, Montanari D, Overington J (2011) Probing the links between in vitro potency, ADMET and physicochemical parameters. Nat Rev Drug Discov 10:197–208. https://doi.org/10.1038/nrd3367

98. Hughes JD, Blagg J, Price DA, et al (2008) Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorganic Med Chem Lett 18:4872–4875. https://doi.org/10.1016/j.bmcl.2008.07.071

99. Greene N, Aleo MD, Louise-May S, et al (2010) Using an in vitro cytotoxicity assay to aid in compound selection for in vivo safety studies. Bioorganic Med Chem Lett 20:5308–5312. https://doi.org/10.1016/j.bmcl.2010.06.129

100. Price DA, Blagg J, Jones L, et al (2009) Physicochemical drug properties associated with in vivo toxicological outcomes: A review. Expert Opin. Drug Metab. Toxicol. 5:921–931

101. Ahmed SSSJ, Ramakrishnan V (2012) Systems Biological Approach of Molecular Descriptors Connectivity: Optimal Descriptors for Oral Bioavailability Prediction. PLoS One 7:e40654. https://doi.org/10.1371/journal.pone.0040654

102. Hopkins AL, Keserü GM, Leeson PD, et al (2014) The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov 13:105–121. https://doi.org/10.1038/nrd4163

103. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: A useful metric for lead selection. Drug Discov. Today 9:430–431

104. Carr RAE, Congreve M, Murray CW, Rees DC (2005) Fragment-based lead discovery: Leads by design. Drug Discov. Today 10:987–992

105. Mortenson PN, Murray CW (2011) Assessing the lipophilicity of fragments and early hits. J Comput Aided Mol Des 25:663–667. https://doi.org/10.1007/s10822-011-9435-z

106. Bembenek SD, Tounge BA, Reynolds CH (2009) Ligand efficiency and fragment-based

drug discovery. Drug Discov. Today 14:278–283

107. Nissink JWM (2009) Simple size-independent measure of ligand efficiency. J Chem Inf Model 49:1617–1622. https://doi.org/10.1021/ci900094m

108. Reynolds CH, Tounge BA, Bembenek SD (2008) Ligand binding efficiency: Trends, physical basis, and implications. J Med Chem 51:2432–2438. https://doi.org/10.1021/jm701255b

109. Ryckmans T, Edwards MP, Horne VA, et al (2009) Rapid assessment of a novel series of selective CB2 agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis. Bioorganic Med Chem Lett 19:4406–4409. https://doi.org/10.1016/j.bmcl.2009.05.062

110. Ertl P, Schuffenhauer A (2009) Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J Cheminform 1:8. https://doi.org/10.1186/1758-2946-1-8

111. Bahmani A, Saaidpour S, Rostami A (2017) A Simple, Robust and Efficient Computational Method for n-Octanol/Water Partition Coefficients of Substituted Aromatic Drugs. Sci Rep 7:1–14. https://doi.org/10.1038/s41598-017-05964-z

112. Ertl P, Rohde B, Selzer P (2000) Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J Med Chem 43:3714–3717. https://doi.org/10.1021/jm000942e

113. Fernandes J, Gattass CR (2009) Topological polar surface area defines substrate transport by multidrug resistance associated protein 1 (MRP1/ABCC1). J Med Chem 52:1214–1218. https://doi.org/10.1021/jm801389m

114. Prasanna S, Doerksen R (2008) Topological Polar Surface Area: A Useful Descriptor in 2D-QSAR. Curr Med Chem 16:21–41. https://doi.org/10.2174/092986709787002817

115. Congreve M, Carr R, Murray C, Jhoti H (2003) A “Rule of Three” for fragment-based lead

discovery? Drug Discov. Today 8:876–877

116. Gulzar M, Ali S, Khan F, et al (2019) Binding mechanism of ca eic acid and simvastatin to the integrin linked kinase for therapeutic implications: A comparative docking and MD simulation studies. J Biomol Struct Dyn 37:4327–4337

117. Machaba KE, Mhlongo NN, Soliman MES (2018) Induced Mutation Proves a Potential Target for TB Therapy: A Molecular Dynamics Study on LprG. Cell Biochem Biophys 76:345–356. https://doi.org/10.1007/s12013-018-0852-7

118. Pitera JW (2014) Expected distributions of root-mean-square positional deviations in proteins. J Phys Chem B 118:6526–6530. https://doi.org/10.1021/jp412776d

119. Brüschweiler R (2002) Efficient RMSD measures for the comparison of two molecular ensembles. Proteins Struct Funct Bioinforma 50:26–34. https://doi.org/10.1002/prot.10250

120. Król M, Roterman I, Piekarska B, et al (2005) Analysis of correlated domain motions in IgG light chain reveals possible mechanisms of immunological signal transduction. Proteins Struct Funct Genet 59:545–554. https://doi.org/10.1002/prot.20434

CHAPTER 7

"Finding the Needle in the Haystack"- Will Natural Products Fit for Purpose in the Treatment of Cryptosporidiosis? - A Theoretical Perspective.

Kehinde F. Omolabi[a], Clement Agoni[a], Fisayo A. Olotu[a], and Mahmoud E.S. Soliman[a*]

[a]Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Lab webpage: http://soliman.ukzn.ac.za/

Running Head: Alternative anticryptosporidial therapy

7.1 Abstract

The limitations that have inundated mainstream anticryptosporidials have necessitated the exploration of alternative approaches towards the design of novel anticryptosporidials with improved therapeutic activity. We virtually screened a dedicated library of 107,000 natural compounds available in the ZINC database against *Cryptosporidium parvum* inosine monophosphate dehydrogenase-NAD^+ binding site. The top three compounds identified with the best complementarity to the *Cp*IMPDH-NAD^+ binding site included ZINC5225833, ZINC4258873, and ZINC3841381. Estimation of the total binding free energy using Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA) method had the three compounds eliciting favorable binding free energies. ZINC3841381 had the best binding free energy of -58.43kcal/mol followed by ZINC4258873 with -51.1kcal/mol and ZINC5225833, -38.04kcal/mol. The binding of the natural compounds to *Cp*IMPDH induced structural perturbations in the protein compared to the apo. We evaluated the compliance of the natural compounds to Lipinski's rule of 5 and other pharmacokinetic parameters. From this evaluation, the natural compounds are drug-likely but can still make do with some modifications to improve their therapeutic potentials. The results obtained from this study can serve as a preliminary background in the further experimental exploration of ZINC5225833, ZINC4258873, and ZINC3841381 as potential anticryptosporidials.

Keywords: Virtual screening; Molecular docking; Molecular dynamics simulation; ADMET compliance; Binding free energy.

7.2 Introduction

Cryptosporidium infection remains the leading cause of diarrhea-related morbidity and mortality in children younger than five years old, especially in developing countries[1–3]. Adult patients who are immune-compromised by HIV-infection, organ-transplants, etc., are not spared from this menace [4]. Presently, there is still the need to discover newer chemotherapeutic compounds that can overcome the challenges mounted against nitazoxanide, which is the only FDA- approved drug for cryptosporidiosis. Early clinical trials revealed that nitazoxanide was able to clear only 75% of the parasitic load in immune-competent children, 52% in immune-compromised children, and with prolonged use in HIV co-infected patients, it was almost ineffective [5–7]. Other drugs, though with limited efficacy, have been repurposed as anticryptosporidials. Examples of this include spiramycin, azithromycin, paromomycin, rifamycins, clarithromycin, etc [8–10].

In approaching rational drug design for cryptosporidiosis, *Cryptosporidium parvum* inosine monophosphate dehydrogenase (*Cp*IMPDH) has been extensively explored as a promising druggable target [11–15]. Generally, inosine 5′-monophosphate dehydrogenase (IMPDH) catalyzes the first and rate-limiting step in the synthesis of guanosine monophosphate (GMP), which is the nicotinamide-adenine dinucleotide (NAD^+)-dependent oxidation of inosine 5′-monophosphate (IMP) to xanthosine 5′-monophosphate (XMP) [16]. IMPDH is integral to organisms' survival as its inhibition leads to the depletion of the building blocks of DNA and RNA, thereby hampering cell proliferation and growth [17]. A unique attraction point of *Cp*IMPDH chemotherapeutic targeting is the selectivity it affords relative to its eukaryotic hosts' IMPDH. This is because *Cryptosporidium* harbors a prokaryotic IMPDH despite being a eukaryote [18,19]. The differing structural properties of prokaryotic and eukaryotic IMPDHs ensures the selectivity observed in the mechanism of action of anticryptosporidials [20,21].

The limitations of conventional drugs in treating cryptosporidiosis spurred the exploration of natural products as an alternative treatment for this disease. Natural products are substances synthesized and produced by living organisms, and these substances almost always have therapeutic activities [22]. The propensity of natural products in ameliorating disease state is being explored in treating parasitic diseases. For instance, in Northwest Cameroon, the hexane extracts of *Salicaceae* leaves, *Euphorbiaceae* roots, and ethanolic extracts of the bark of *Combretaceae* are

used in the treatment of Onchocerciasis [23,24]. Gentistic, ellagic and gallic acids have also been reported to inhibit the growth of *C. elegans*. Likewise, *Aceraceae, Fagaceae, and Rosaceae* families [25]. In treating Schistosomiasis, artemisinin has been used in bringing down the parasitemia level in mice infected with *S.japonicum* [26,27]. Curcumin, vernodalin, avocado, and soybean oil have been used to disrupt mating in *Schistosoma spp* to mitigate egg production [28–30]. Trypanocidal activities have also been observed with active compounds derived from plant *spp* belonging to *Piperaceae Asteraceae and Lauraceae*'s families [31–33].

In circumventing the time and financial resources invested in the traditional drug discovery processes, high-throughput screening and combinatorial chemistry technologies have evolved to accelerate these processes by allowing large libraries of compounds to be screened in an instant [34]. This is made possible partly by docking the ligands in the pre-determined binding site of the biological target's 3D-structure [35,36]. The complementarity of the molecules to the binding site is thereafter scored [37,38]. This concept is known as structure-based virtual screening. The compounds that displayed good complementarity to the protein's binding site are further analyzed by integrating molecular dynamics simulation and *in silico* determination of drug- likeness as carried out in similar studies [39,40]. This will lay a strong ground-work for the lead compounds' experimental efficacy tests. Virtual screening presents with some advantages: the shortened cycle of potential drug discovery and the possibility of generating a greater number of hits than experimental high throughput screening [41,42].

In this study, the NAD^+ site of *Cp*IMPDH, which has been validated and repeatedly targeted in cryptosporidiosis chemotherapy, was employed in screening a dedicated library of natural compounds. Compounds having the best complementarity to the NAD^+ binding site were subjected further to molecular dynamics simulation. This was necessary to understand the possible inhibitory mechanism of the selected ligands. We went further to analyze *in silico,* the physicochemical properties of these ligands to understand the extent of their drug-likeness. We believe findings from this study would further give the experimentalists possible drug leads from natural sources that hold great potential better than nitazoxanide in treating cryptosporidiosis.

7.2 Computational methodologies

7.2.1 Retrieval and preparation of systems

The 3D- crystallographic structure of *Cp*IMPDH with P131 bound at the NAD^+ binding site and IMP bound at the active site was retrieved from RCSB Protein Data Bank (Entry code: 4RV8) [12]. 4RV8 is a tetramer with a resolution of 2.05Å. However, we used a dimer (chain A and D) principally because it reflected [12] and to minimize computational resources. Co-crystallized molecules reported not to be relevant to the protein's function (ions, crystallographic water) were removed, and the 68 (34 in each chain) missing residues were filled with MODELLER algorithm on the UCSF Chimera Graphic User Interface [43,44].

7.2.2 Virtual screening of natural compounds against NAD^+ binding site of *Cp*IMPDH

The coordinates of the NAD^+ binding site of *Cp*IMPDH were determined using the co-crystallized P131 in 4RV8 as reference. The grid-box dimensions were center (X=2.649, Y=21.997, Z=77.399) and size (X=11.165, Y=9.462, Z=7.323. The mapped NAD^+ site on the target protein was virtually screened against 107,000 natural compounds curated in the ZINC repository of natural compounds (https://zinc.docking.org/substances/subsets/natural-products/). Autodock Vina, which was integrated with High-performance computing [45] was used to execute the screening, and a maximum of 8 conformers, was considered for each compound. Thereafter, the docking score of each compound was computed. The top three compounds with the best binding affinity were selected. Their docked conformation in the NAD^+ site in *Cp*IMPDH was visualized in the ViewDock plugin-integrated Chimera [46]. These complexes were prepared for molecular dynamics simulation studies to provide structural and dynamical perspective into the inhibitory potential of natural compounds against cryptosporidiosis.

7.2.3 Molecular dynamics (MD) simulations

The systems were divided into apo- IMPDH bound to IMP in its active site and complexes, which had apo and the natural compounds bound in the NAD^+ binding site. They were then subjected to a long-timescale MD simulation run (300ns) on AMBER18 Graphical Processing Unit (GPU) [47] Protein parameters were determined via FF14SB force-field. The ANTECHAMBER module was used in the ligands' parameterization [48]. The LEAP module was used in the solvation of the

system in 10 Å TIP3P water box and the generation of topology files. Partial minimization and full minimization was executed in 2500 steps and 5000 steps, respectively. Heating the systems was for 50 ps from 0 to 300 k in an NVT canonical ensemble using a Langevin thermostat [49]. The systems were subjected to equilibration without energy restraint for 1000ps. We carried out the MD production run for 300 ns while constricting the hydrogen bond with the SHAKE algorithm. CPPTRAJ and PTRAJ modules were used in analyzing the resulting trajectories [50]. The data were plotted using Origin data analytical tool [51].

7.2.4 Post- Molecular Dynamics Simulation Analysis

Post-MD analyses carried out included C-α Root mean square deviation (RMSD), Radius of gyration (RoG), Root mean square fluctuations (RMSF), Principal component analysis, Solvent accessible surface area, Dynamics cross-correlation matrix, and binding free energy analysis via Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA) method.

7.2.5 Thermodynamics Calculations

Molecular Mechanics Generalized Poisson Boltzmann Surface Area (MM/PBSA) method was used to estimate the predicted hits' binding free energy to the NAD^+ binding site of *Cp*IMPDH. MM/PBSA is an efficient tool used to evaluate ligands' interaction with biological macromolecules [52,53]. MM/PBSA was calculated via AmberTools18-integrated MM/PBSA.py python script. This script employs continuum solvent models to automatically analyze binding free energies of MD simulation's snapshots [47]. In calculating the binding free energies, only frames generated after each protein had achieved convergence was used. These frames represented trajectories that have completely stabilized, as shown in the RMSD plot (Figure 7.5). Therefore, any computational or entropic artifacts that could interfere with the calculation of the total binding free energy were minimized. In this study, stable frames from 270ns -300ns (30,000 frames) were considered.

Mathematically binding free energy is calculated as follows:

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{inhibitor}) \quad (1)$$

$$\Delta G_{bind} = \Delta G_{gas} + \Delta G_{sol} - T\Delta S \quad (2)$$

$$\Delta G_{gas} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (3)$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{np,sol} \quad (4)$$

From the above, the summation of the internal (ΔE_{int}), electrostatic (ΔE_{ele}), and van der Waals (ΔE_{vdW}) energies gives the gas-phase energy (ΔG_{gas}). In contrast, the solvation free energy (ΔG_{sol}) is estimated by the polar solvation (ΔG_{PB}) and non-polar ($\Delta G_{np,sol}$) contributions. The estimated ΔG_{bind} was decomposed into individual energies contributed by each binding site residue through the per-residue energy decomposition analysis (PRED). The PRED was analyzed by manually adding the binding residues of the selected natural compounds to MMPB/SA.py script incorporated in Amber18Tools. This was imperative to identify specific residues that contributed to the ligands' affinity and stability in *Cp*IMPDH.

7.2.6 Screening of the physicochemical properties of the natural compounds

The natural compounds' comparative physicochemical properties were evaluated using online prediction tools, which use have also been validated in other studies.[54–56] These include SWISSADME [57], DataWarrior [58], Molsoft program (http://molsoft.com/mprop/), Molinspiration Chemoinformatics [59]. The evaluation of the pharmacokinetic properties of the ligands was determined by these tools and also assessed their adherence to Lipinski's rule of five [60].

7.3 Results and Discussion

7.3.1 Sampling the complementarity of the natural compounds in NAD^+ binding site of *Cp*IMPDH

Molecular docking models at an atomic level evaluate the complementarity, conformation, position, and orientation of ligand in the binding site of the protein of interest [61]. Monte Carlo methods employed in AutoDock Vina was used in this study [62]. Here, different poses of ligands were generated via bond rotation and were selected based on their potential energies [62]. With the defined coordinates of the NAD^+ binding site of *Cp*IMPDH as mentioned in section 2.2, the top three compounds with the best docking scores were ZINC5225833 (-9.8 kcal/mol), ZINC4258873 (-9.1 kcal/mol), and ZINC3841381 (-9.0 kcal/mol). The top 50 natural compounds with the highest binding affinity scores are presented in Supplementary Table 7.1. The 2D

structures of the top three natural compounds with the best binding scores are shown in Figure 7.1.

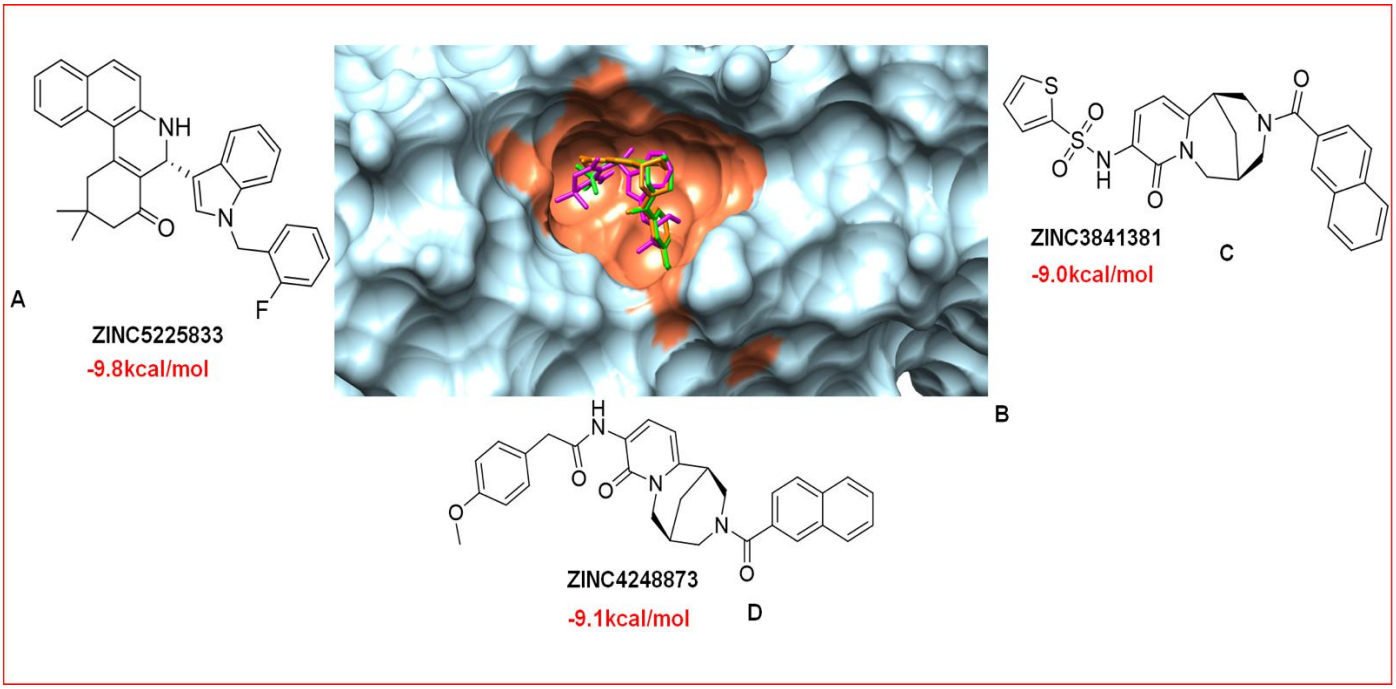


Figure 7.1 A) 2D structure and docking score of ZINC5225833. **B**) Surface representation of the docking of the natural compounds in NAD^+ binding site (coral) of *Cp*IMPDH (light blue). **C**) 2D structure and docking score of ZINC4248873. **D)** 2D structure and docking score of ZINC3481381.

7.3.2 Quantification of the total binding free energy of ZINC5225833, ZINC4258873, and ZINC3841381 in *Cp*IMPDH

Binding free energies of the natural compounds were computed *via* MM/PBSA tools, which, aside from estimating the total binding energy, also calculate various energetic contributions of the binding site residues to the overall total binding free energy. In all the energy components estimated for the ligand-protein interaction, Vander Waals force was the dominant contributory factor to the favorable total binding energies observed in all (Table 1). *Cp*IMPDH-ZINC3841381 had the best binding energy (-58.43 kcal/mol) despite having the lowest docking affinity of the three (-9.0 kcal/mol). This buttresses the importance of following up molecular docking with molecular simulation as the latter gives a more accurate representation of what obtains in the biological system. *Cp*IMPDH-ZINC4258873 had binding free energy of -51.10 kcal/mol, and

*Cp*IMPDH-ZINC5225833 had the least binding energy (-38.04 kcal/mol) though it had the best docking score (-9.8 kcal/mol).

Table 7.1. Estimation of the total binding free energies of the natural compounds in *Cp*IMPDH

Binding free energy analysis			
Energy Components (kcal/mol)	***Cp*IMPDH-ZINC5225833**	***Cp*IMPDH-ZINC4258873**	***Cp*IMPDH-ZINC3841381**
ΔE_{vdW}	-48.99±0.12	-64.86±0.14	-60.84±0.12
ΔE_{ele}	-16.40±0.17	-16.45±0.15	-13.00±0.11
ΔG_{gas}	-65.39±0.23	-81.31±0.19	-73.84±0.15
$\Delta G_{ele,sol(PB)}$	32.86±0.16	37.69±0.11	21.73±0.10
$\Delta G_{np,sol}$	-5.51±0.11	-7.48±0.10	-6.32±0.10
ΔG_{sol}	27.35±0.15	30.21±0.11	15.41±0.10
ΔH	-38.04±0.14	-51.10±0.12	-58.43±0.09
ΔG_{bind}	-38.04±0.13	-51.10±0.14	-58.43±0.13

*Cp*IMPDH-ZINC3841381 having the best binding energy, could not be disassociated from the lower solvation energy (**$\Delta G_{(PB)}$**) invested in solvating this complex compared to other complexes (Table 1). In *Cp*IMPDH-ZINC3841381, the solvation energy was 21.73kcal/mol compared to *Cp*IMPDH-ZINC5225833 which was 32.86kcal/mol and *Cp*IMPDH-ZINC4258873 was 37.69kcal/mol. The orientation of the ZINC compounds in the NAD^+ site of *Cp*IMPDH deduced from the non-polar solvation energy ($\Delta G_{np,sol)}$ could have also influenced the total binding energy of the compounds to the protein. ZINC5225833 had the least non-polar solvation energy of -5.51 kcal/mol, which meant that the ligand was oriented in a manner that made it more surface exposed, thereby interacting more with the polar environment and reducing its total binding free energy than ZINC4258873 (-7.48 kcal/mol) and ZINC3841381(-6.32 kcal/mol).

7.3.3 Energy contributions of binding site residues to the affinity and stability of the natural compounds.

The energy contributions of each residue that interacted with the ligands at the NAD^+ binding site of *Cp*IMPDH were computed using the MM/PBSA tool. This was necessary to understand the group of residues that significantly affect the binding and stability of the ligand. As presented in Figure 7.2A, the first four residues that contributed the best in *Cp*IMPDH-ZINC5225833 were Met308 (-3.11 kcal/mol), Met302 (-1.78 kcal/mol), Ala165 (-1.65 kcal/mol) and Lys 309 (-1.48 kcal/mol). In *Cp*IMPDH-ZINC4258873, the most favorable energy contributing residues were Met326 (-2.37 kcal/mol), Ala165 (-2.08 kcal/mol), Pro26D (-1.91 kcal/mol) and Leu25D (-1.90 kcal/mol) (Figure 7.2C). *Cp*IMPDH-ZINC3841381 which had the leading total binding free energy had Ser164 (-4.37 kcal/mol), Ala165 (-3.15kcal/mol), His166 (-2.0 kcal/mol) and Met302 (-1.8 kcal/mol) (Figure 7.3A).

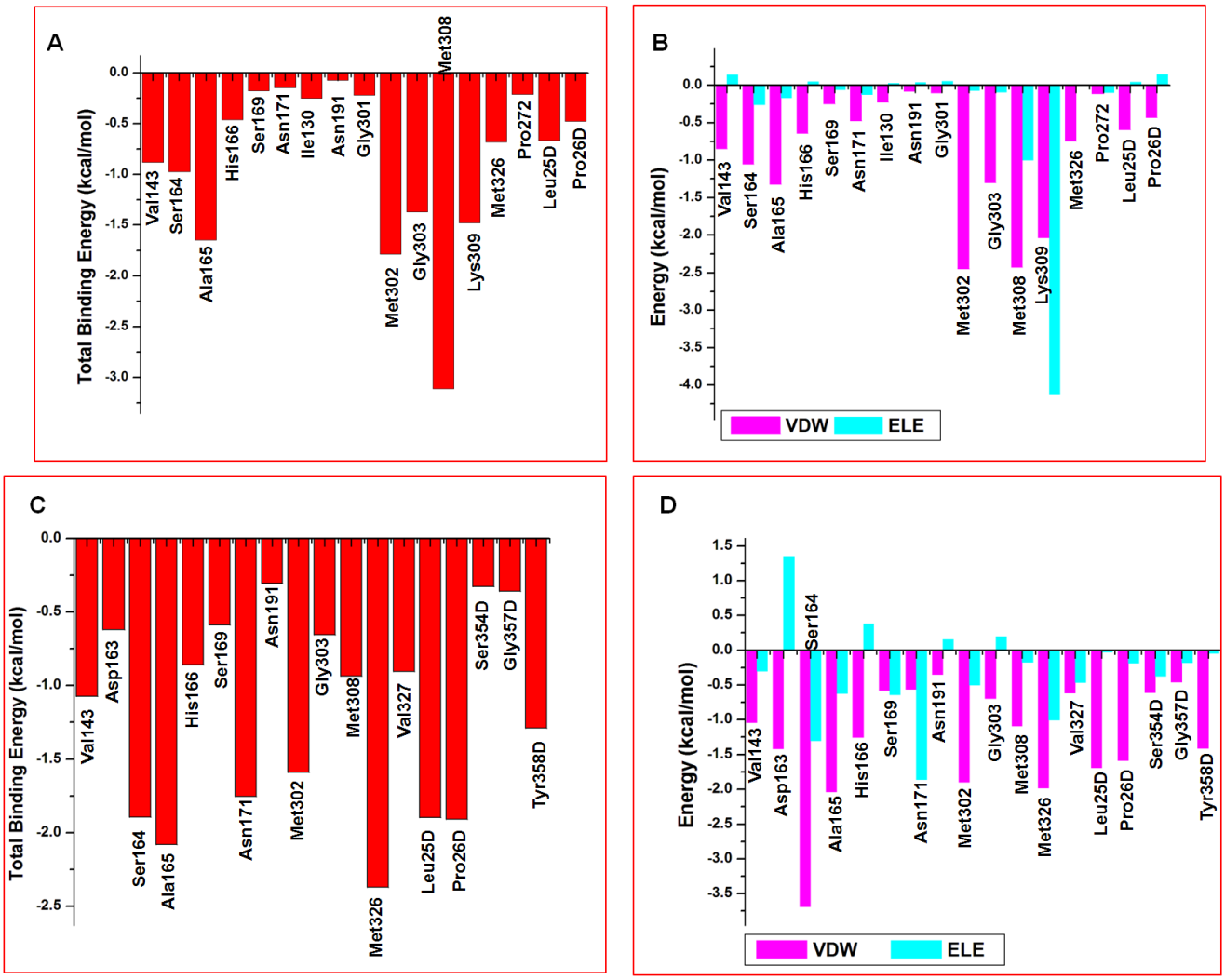


Figure 7.2 A) Energy contributions of binding sites residues to the total binding energy of ZINC5225833. **B)** Vander Waal and electrostatic energy contributions of each binding residues to the total binding energy of ZINC5225833. **C)** Energetic contributions of each binding sites residues to the total binding energy of ZINC4258873. **D)** Vander Waal and electrostatic energy contributions of each binding residues to the total binding energy of ZINC4258873.

We further decomposed the total energy contributed by each binding site residues to Vander Waal and electrostatic energies. Zeroing in on only the electrostatic energies contribution which is reputed to be the strongest interaction, the residual contributions are enumerated as follow. In *Cp*IMPDH-ZINC5225833, the top four residues were Lys309 (-4.12 kcal/mol), Met308 (-1.0 kcal/mol), Ser164 (-0.26kcal/mol) and Ile172 (-0.25 kcal/mol) (Figure 7.2B). In *Cp*IMPDH-ZINC4258873, the best electrostatic energy contributing residues were Asn171 (-1.89 kcal/mol),

Ser164 (-1.3 kcal/mol), Ala165 (-0.62 kcal/mol) and Met302 (-0.5 kcal/mol) (Figure 7.2D). Finally, in *Cp*IMPDH-ZINC3841381, the residues were Ser164 (-5.39 kcal/mol), Ala165 (-1.16 kcal/mol), Ser354D (-0.66 kcal/mol) and Asn171 (-0.54 kcal/mol) (Figure 7.3B).

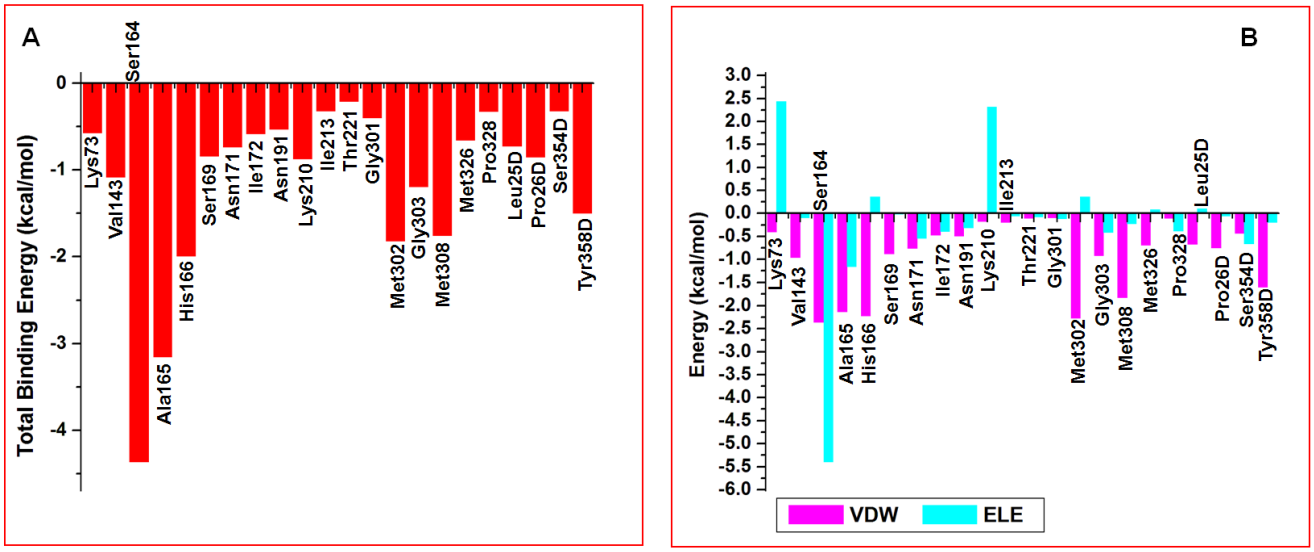


Figure 7.3 A) Energy contributions of binding sites residues to the total binding energy ZINC5225833. **B)** Vander Waal and electrostatic energy contributions of each binding residues to the total binding energy of ZINC3841381.

7.3.4 Molecular Insights into the non-covalent interactions between *Cp*IMPDH and the natural compounds

We evaluated the non-covalent interactions of the binding site residues of *Cp*IMPDH with the natural compounds. This could lend insight into the factors affecting the compounds' stability in the binding pockets and the observed disparity in the compounds' total binding free energy. We took the representative interactions at 50ns intervals, from post-equilibrated time frames (200ns) to the end of the simulation (300ns). Only residues having direct interactions were captured.

The relatively weaker but favorable binding free energy of ZINC5225833 could have been influenced by the domination of weak interactions - alkyl and π -alkyl interactions in its binding. The large numbers of this interaction though compensated by their consistent short distance (>6Å) between the residues and moieties of ZINC5225833 over the simulation period sampled, could have contributed in no small way to the favorable binding free energy (Figure 7.4). The establishment of a consistent conventional (OH--O) hydrogen bond interaction between Lys309

(200ns and 300ns) and the oxygen group of ZINC5225833 at 1.92Å and 2.35Å, respectively, explained why Lys309 was the most electrostatic energy-contributing residue (-4.12 kcal/mol) to the binding of ZINC5225833 (Figure 7.2B). Met 302 and 326 formed SH-π bond with the aromatic rings of the ligand. It has been reported that SH-π interactions confer ligand stability even when the interaction is long-distance [63]. This can partly interpret why the Met302 and Met308 of all the binding site residues were the most Vander Waal energy-contributing residues (-2.45kcal/mol and -2.43 kcal/mol, respectively) to the binding of ZINC5225833 (Figure 7.2B). Almost all the residues were involved in multiple bonding with the ligand.

In ZINC4258873 at 200ns and 300ns π- π stacking existed between the aromatic rings of Tyr358D and that of the ligand at a distance of 4.81Å and 4.78Å, respectively (Figure 7.4). This interaction constituted one of the favorable protein-ligand interactions. π- π stacking plays an essential role in forming recognition motifs and structural stability of proteins [64,65]. ASN171 prominently formed a short distance OH-O with the ligand's oxygen group, and this binding donated the most electrostatic energy (-1.86 kcal/mol) (Figure 7.2D) to the total binding free energy of the ligand. The other interactions were alkyl, pi-alkyl, and pi-sulfur. These interactions were all short distanced (>6Å), and like ZINC5225833, all the interacting residues were involved in multiple interactions with the ligand.

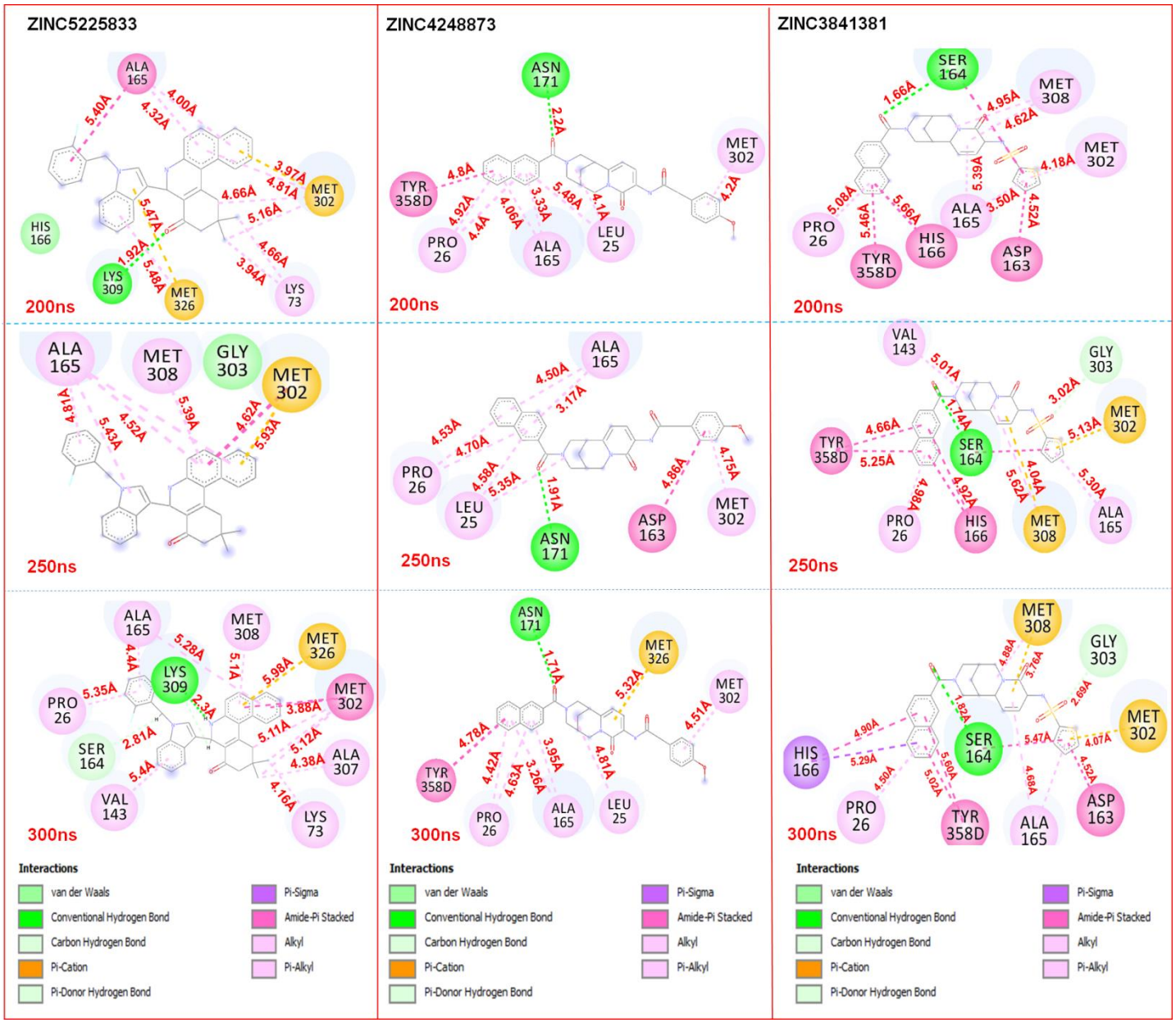


Figure 7.4 Non-covalent interactions between the binding site residues and the natural compounds at different time-points

In ZINC3841381, which had the best binding free energy could have been occasioned by it having the most diverse interactions of all the three natural compounds. Multiple π- π stacking was formed between the aromatic rings of Tyr358D and His166 in all the time-frame sampled (Figure 7.4). The conventional hydrogen bond formed by Ser164 and the simultaneous amide- π stacking interactions were pivotal to the ligand's affinity and stability in the binding pocket. These interactions had the most favorable Vander Waal and electrostatic energy contributions (-

2.36kcal/mol and -5.4kcal/mol, respectively) (Figure 7.3D). The σ-π and SH-π bonds formed by His166, Met302, and Met308 also contributed mostly to the Vander Waal energy and little to electrostatic energy.

7.3.5 Structural stability patterns induced by the binding of the natural compounds to *Cp*IMPDH

One of the many applications of the atom-positional root-mean-square deviation (RMSD) is to evaluate structural perturbation that has occurred during MD simulations [66,67]. In this context, RMSD measures by sampling the entire protein structure, the mean of the movement of backbone atoms of the complexes with reference to the apo. The RMSD analysis of the apo and complexes' global structure shows that the natural compounds' binding stabilized the protein structure compared to the apo, although not significantly. This implies that the complexes did not significantly diverge from the initial coordinates of the protein. Compared to the apo (3.34Å), *Cp*IMPDH-ZINC4258873 (2.61Å) stabilized the structure the most, followed by *Cp*IMPDH-ZINC5225833 (2.86Å) and *Cp*IMPDH-ZINC3841381 (3.15Å). In figure 7.5A, the systems equilibrated around 140ns, and this stability was maintained till the end of the simulation to a large extent. This stability gives a good premise for the reliability and confidence in subsequent structural analyses and inferences.

For efficient enzymatic catalysis of a substrate to product, the enzyme's active site undergoes some transformation and reorganization to complement the ligand's shape. This transformation can make the active site more rigid or flexible [68–70]. The active site RMSD values in Table 2 show that the active site of *Cp*IMPDH-ZINC5225833 (2.06Å) and *Cp*IMPDH-ZINC3841381 (2.10Å) during catalysis requires almost no conformational alterations when compared to apo (2.01Å). We can also put forward that on its own, the active site of the apo is also catalytically competent as regards the binding of ZINC5225833 and ZINC3841381. However, in *Cp*IMPDH-ZINC4258873 (3.2Å), the ligand binding induced quite a large geometric and conformational reorganization in the active site compared to apo. This is clearly shown in Figure 7.5C.

Table 7.2 The mean values of the structural indices of the apo and complexes.

Structural Components	*Cp*IMPDH	*Cp*IMPDH-ZINC5225833	*Cp*IMPDH-ZINC4258873	*Cp*IMPDH-ZINC3841381
		Mean RMSD Values (Å)		
Whole Protein	3.34	2.86	2.61	3.15Å
Active Site Residues	2.01	2.06	3.20	2.10
Ligand		2.29	0.84	1.82
		Mean RoG Values (Å)		
Whole Protein	28.22	28.22	28.15	28.37
Active Site Residues	10.43	10.45	10.54	10.44
		Mean RMSF Values (Å)		
Whole Protein	1.21	1.24	1.22	1.18
Binding Site Residues	0.84	0.76	0.97	0.96

The stability of the ligand in the binding pocket of *Cp*IMPDH was also examined. This is important as the (in) stability of the ligand in the binding pocket can also give insight into the structural mechanism of action of the ligands. ZINC4258873 was the most stable in the binding pocket (0.84Å). It was observed that during the entire simulation period, the movement of the ligand in the binding pocket was minimal, most especially after 50ns (Figure 7.5D). This pattern was also observed for ZINC3841381 (1.82Å). These two ligands had the most favorable binding energy. Their stability could imply that they were in continuous contact with residues crucial to their binding and stability hence the probable reason for their better binding free energies. The converse

was observed in ZINC5225833 (2.29Å) (Figure 7.5D). It having the least binding free energy might not be too far from the fact that the ligand continually fluctuated in the binding site during the entire simulation period. This fluctuation might have occasioned brief and discontinuous contact with the residues important for its stability and binding.

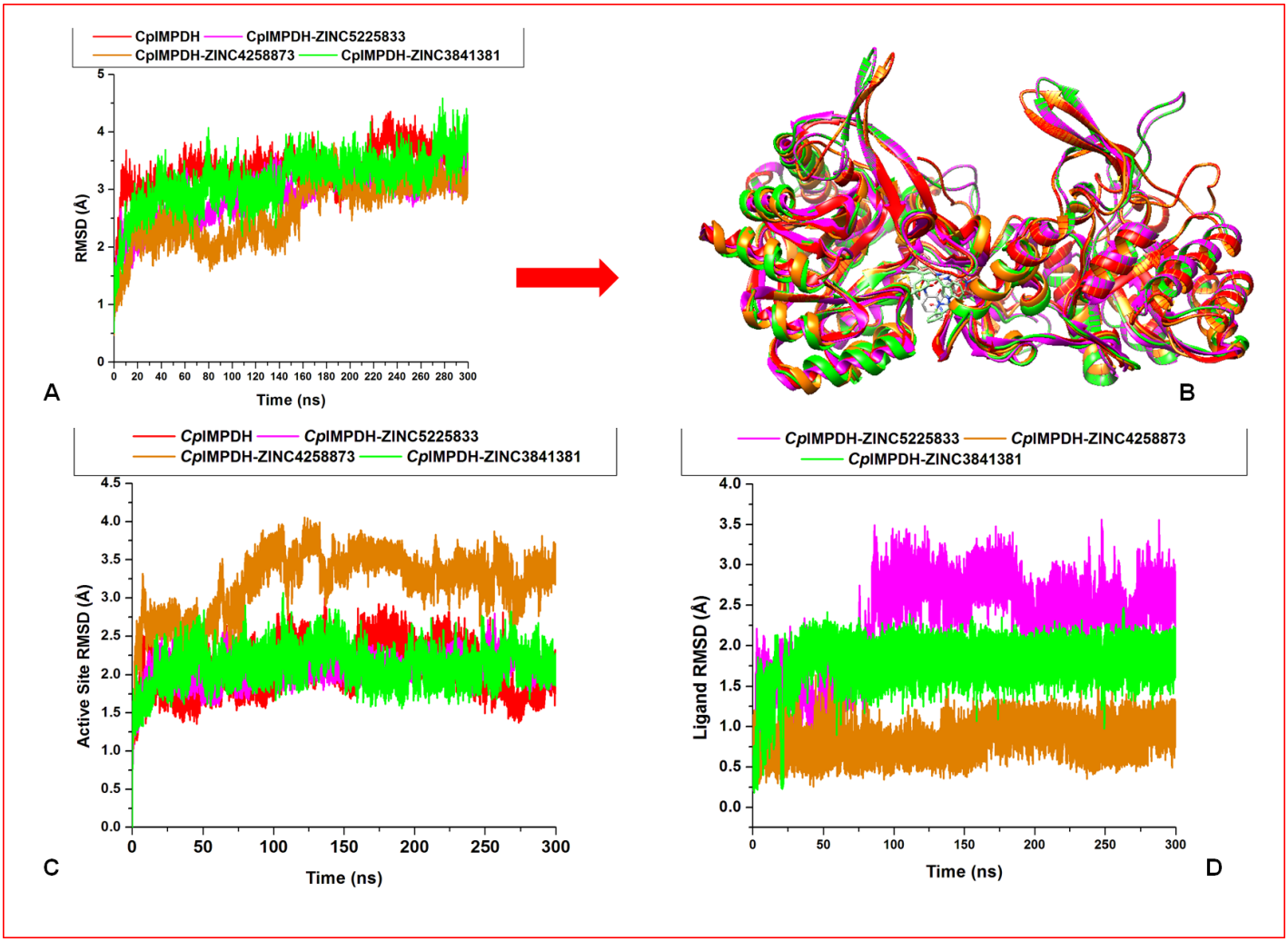


Figure 7.5 A) The RMSD trend observed in the global structure of apo and the complexes during the simulation period. **B)** Superposition of the global structure of apo and complexes at 60ns post-MD simulation. **C)** The RMSD trend observed in the active site of apo and the complexes during the simulation period. **D)** The RMSD trend observed in the ligands during the simulation period.

The radius of the C-α atoms' gyration gives insight into the compactness, shape, and folding of the protein before and after binding [71]. In the whole structure and the active site residues of the apo and the complexes, there was an induction of similar structural compactness in both systems

(Figure 7.6A and 7.6B). Inferring from the RoG analysis alone (Table 2), it detected non-significant variations in the compactness of the protein.

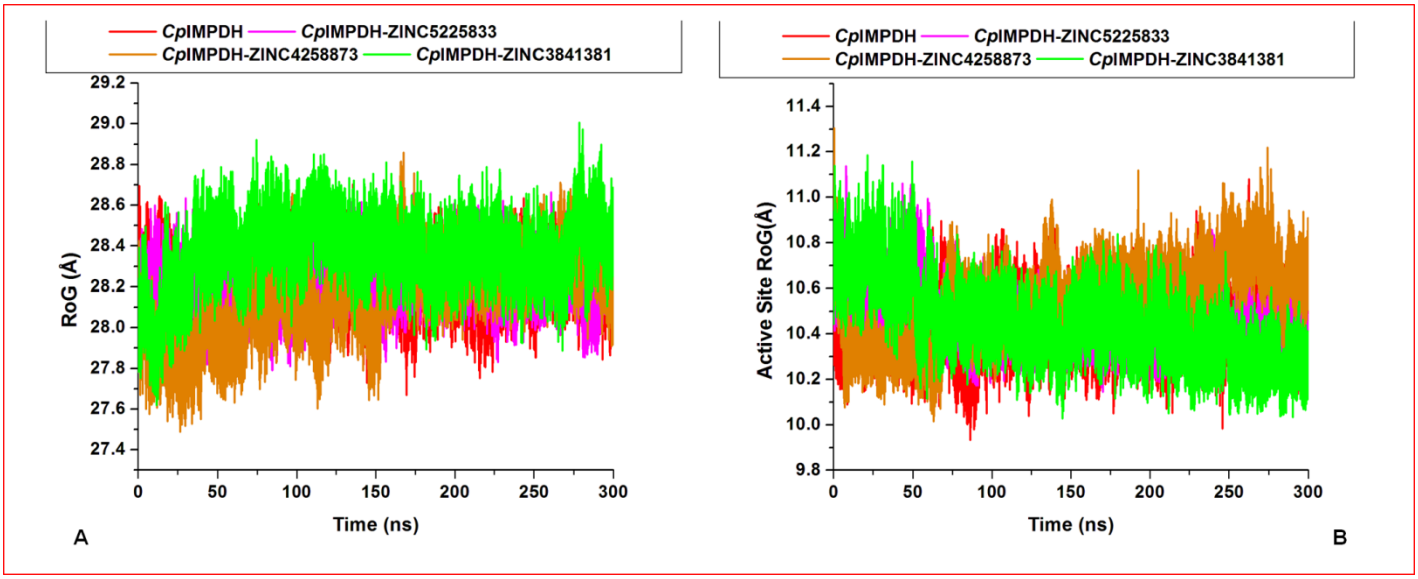


Figure 7. 6 A) The pattern observed in the RoG of the global structure of apo and the complexes during the simulation period. **B)** The pattern observed in the RoG of the active site of apo and the complexes during the simulation period.

7.3.6 Evaluation of ligand-induced conformational flexibility of *Cp*IMPDH

The flexibility of a protein can be influenced by the movement of a group of residues to bind a ligand or when the protein folds on binding a ligand [72]. Herein we analyzed two distinct classes of the flexibility of the backbone atoms of *Cp*IMPDH when interacting with ligands. First does the motion of the protein's global structure, which considers wider and inter-correlated conformational changes in the different domains constitute the protein. The second is the motion strictly restricted to the active site residues. In the global structures, there was only a subtle but not a significant shift in structural flexibility of the complexes compared to apo, which is similar to RMSD result. Apo had an RMSF mean value of 1.21Å *Cp*IMPDH-ZINC5225833, and *Cp*IMPDH-ZINC4258873 were higher 1.24Å and 1.22Å, respectively (Table 2). *Cp*IMPDH-ZINC3841381, which had the best binding energy, had a lower residual fluctuation (more rigidity) in its global structure- 1.81Å (Table 2). A conformationally rigid protein is indicative of an inhibited protein [72], while protein flexibility indicates a fully functional protein. It might follow that the binding of ZINC3841381 to *Cp*IMPDH inhibited the protein by making it more rigid than the apo. It is essential to note from

Figure 7.7A, that in both apo and the complexes regions in chain D had the most flexibility compared to chain A. The residues that fluctuated highly in chain D were residues 310-340, 499-522, 590-610, and 650-652 (Figure 7.7A).

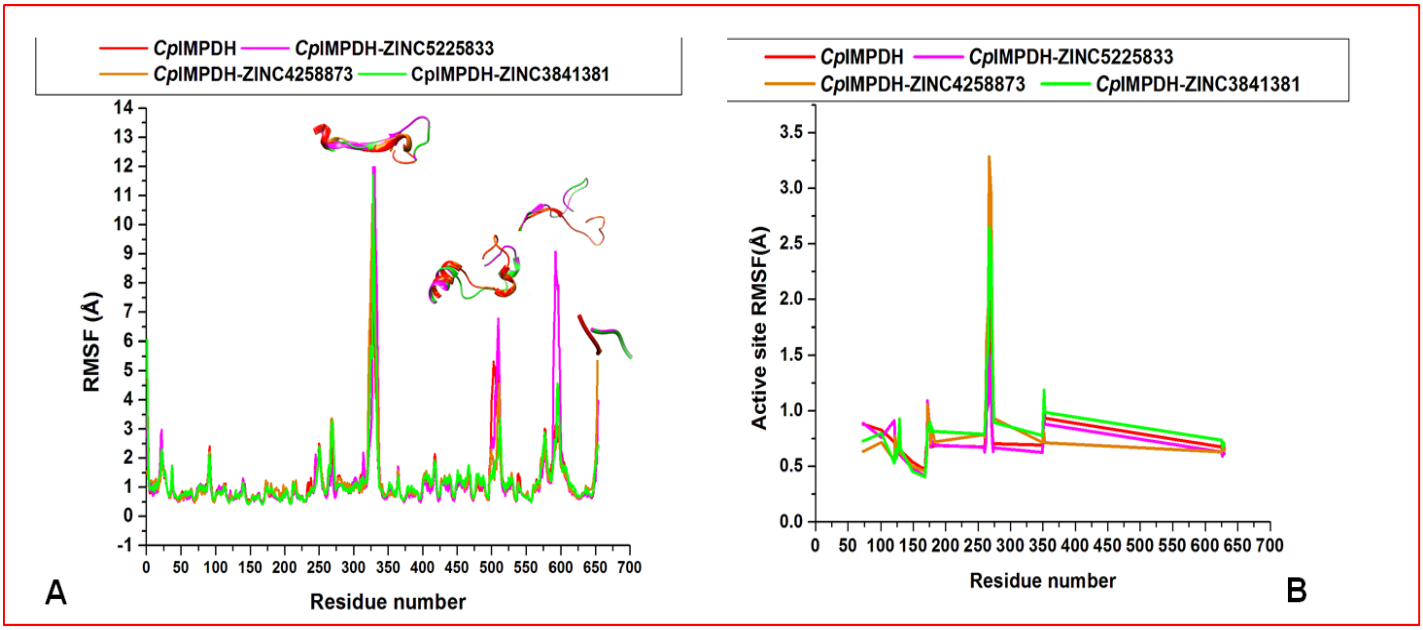


Figure 7.7 A) Root mean square of fluctuation of the global structure of the apo and complexes, the insets depict the region that fluctuated highly in the protein. **B)** Root mean square of fluctuation of the active site of the apo and complexes.

When bound to the ligand, the active sites' fluctuation was not also spectacular compared to apo. *Cp*IMPDH-ZINC4258873 and *Cp*IMPDH-ZINC3841381 had the highest fluctuation at their active site (0.97Å and 0.96Å respectively) when compared to *Cp*IMPDH-ZINC5225833 (0.76Å). Overall, ligand binding did not induce significant conformational flexibility in the complexes than apo (Figure 7.7B).

7.3.7 Solvent accessible surface exposure and its effect on the conformational stability of ligand-bound *Cp*IMPDH

The functional shape of a protein is its 3D (tertiary) structure. The assumption of this shape enables the presentation of its functional group on the outside to interact with other molecules to mediates its function. Protein folding is critical to proteins' stability and the maintenance of the 3D shape [73]. Solvent accessible surface area evaluates the extent to which the surface atoms of proteins are in contact with solvent (water) in the biological system [74]. SASA principally determines

protein (un)folding and the functional relationship between a protein and its residues [75]. Proteins have hydrophobic cores, which are made up of non-polar amino acid residues interacting with each other. The burial and folding of the hydrophobic core away from the aqueous region determine the stability of the 3D structure of the protein. If the hydrophobic interaction among the non-polar residues is weak, hydration, denaturation etc., can perturb the hydrophobic core, causing protein unfolding, which will eventuate a loss in the integrity of the protein's 3D structure and functionality [76–78].

Table 7.3 The average SASA values of the components of apo and the complexes

Structural Components	**Mean SASA Values (^{A2})**			
	***Cp*IMPDH**	***Cp*IMPDH-ZINC5225833**	***Cp*IMPDH-ZINC4258873**	***Cp*IMPDH-ZINC3841381**
Whole Protein	25295.36	25179.85	24843.15	25090.10
Active Site Residues	1046.90	777.78	781.80	759.38
Ligand Bound to Receptor		218.24	124.85	131.48
Free Ligand		579.40	618.29	571.65
% of Exposed Ligand		37.7%	20.1%	23%

In the global structure, when compared to apo, the binding of the ligands non-significantly induced an increased folding of the residues in the hydrophobic core, thereby stabilizing the protein structure (Table 3). The same pattern followed in the active site residues. The apo's active site residues were exposed to the aqueous environment making it unstable (1046.9A^2). However, the active site residues of the complexes were buried in the hydrophobic core, with *Cp*IMPDH-ZINC3841381 buried the most (759.38A^2). This stability in the active site of *Cp*IMPDH-ZINC3841381 induced by the less exposure to the aqueous region may have influenced the complex having the most favorable binding energy coupled with the fact the ligand ZINC3841381 was less surface exposed than ZINC5225833 having their ligands exposed at 23% and 37.7%, respectively. The solvent-accessible surface exposure pattern of the global structure, active site, and ligands are presented in Figure 7.8.

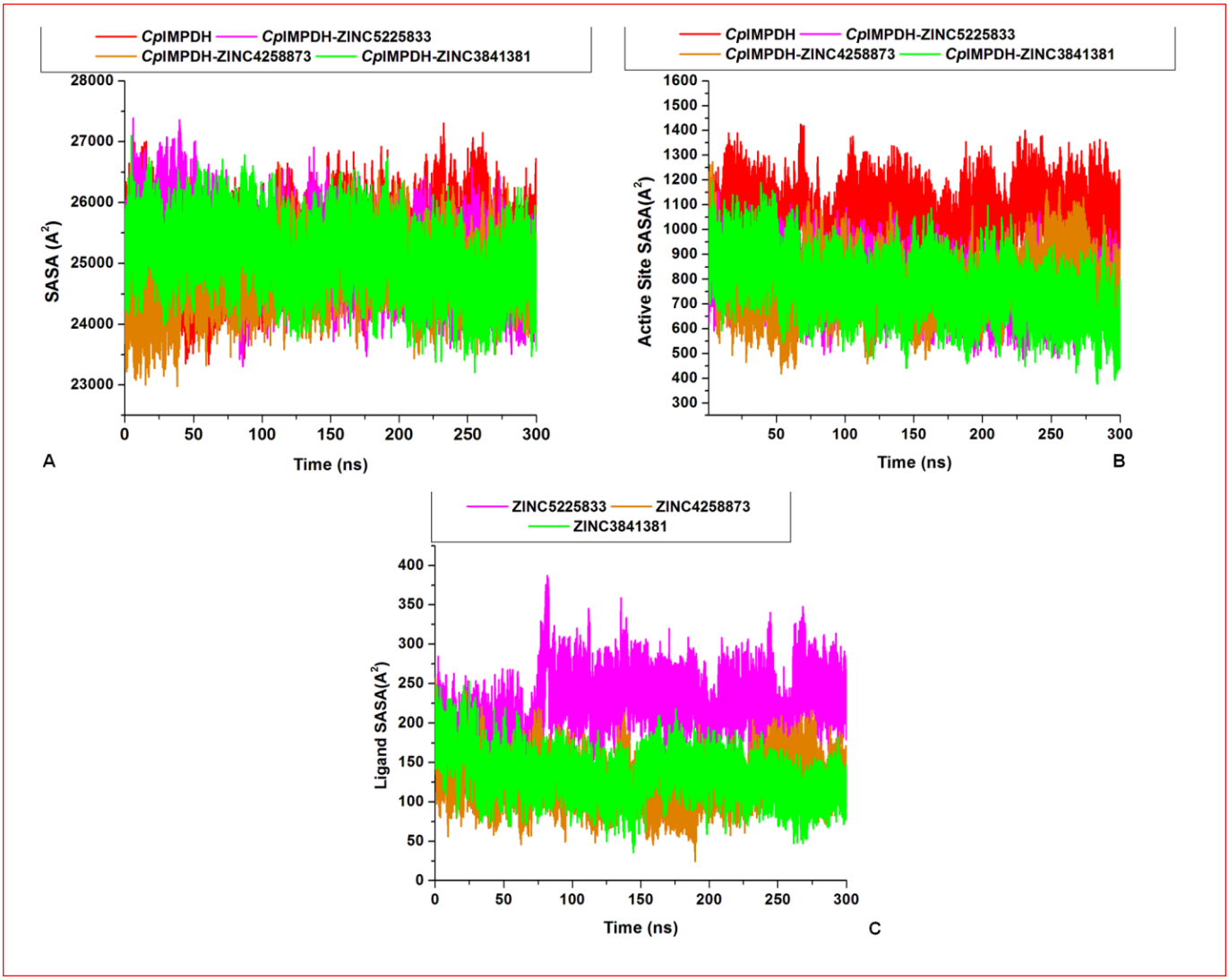


Figure 7.8 A) The solvent-accessible area of the global structure of apo and the complexes during the simulation period. **B)** The solvent-accessible area of the active site of apo and the complexes during the simulation period. **C)** The exposure of the natural compounds to the solvent accessible area over the simulation period.

7.3.8 Atomistic insights into the correlated and anti-correlated motions of *Cp*IMPDH upon ligand binding

We also analyzed the trajectories to compute the cross-correlation displacement and fluctuation between the C-α atoms of the residues making up the respective systems [79,80]. Correlated motions that are positive ranged from the color green to deep red (+1), while negatively correlated motions ranged from cyan to black (-1). Generally, the unliganded protein residual movement was

positively correlated, going by the dominant green to deep red coloration (Figure 7.9). This positive correlation trend is also not different from the bound proteins. However, a correlated movement that visibly stood out in the apo and complexes was that of residues 350-652, representing the second chain. In apo and *Cp*IMPDH-ZINC3841381, the second chain residues were less positively correlated than *Cp*IMPDH-ZINC5225833 and *Cp*IMPDH-ZINC4258873, meaning that the residues in apo and *Cp*IMPDH-ZINC3841381 fluctuated less than *Cp*IMPDH-ZINC5225833 and *Cp*IMPDH-ZINC4258873 (Figure 7.9). This is an agreement with their residual mean fluctuation value discussed earlier in section. The less positively correlated movement of *Cp*IMPDH-ZINC3841381 compared to other complexes might influenced why it had the most favorable binding free energy.

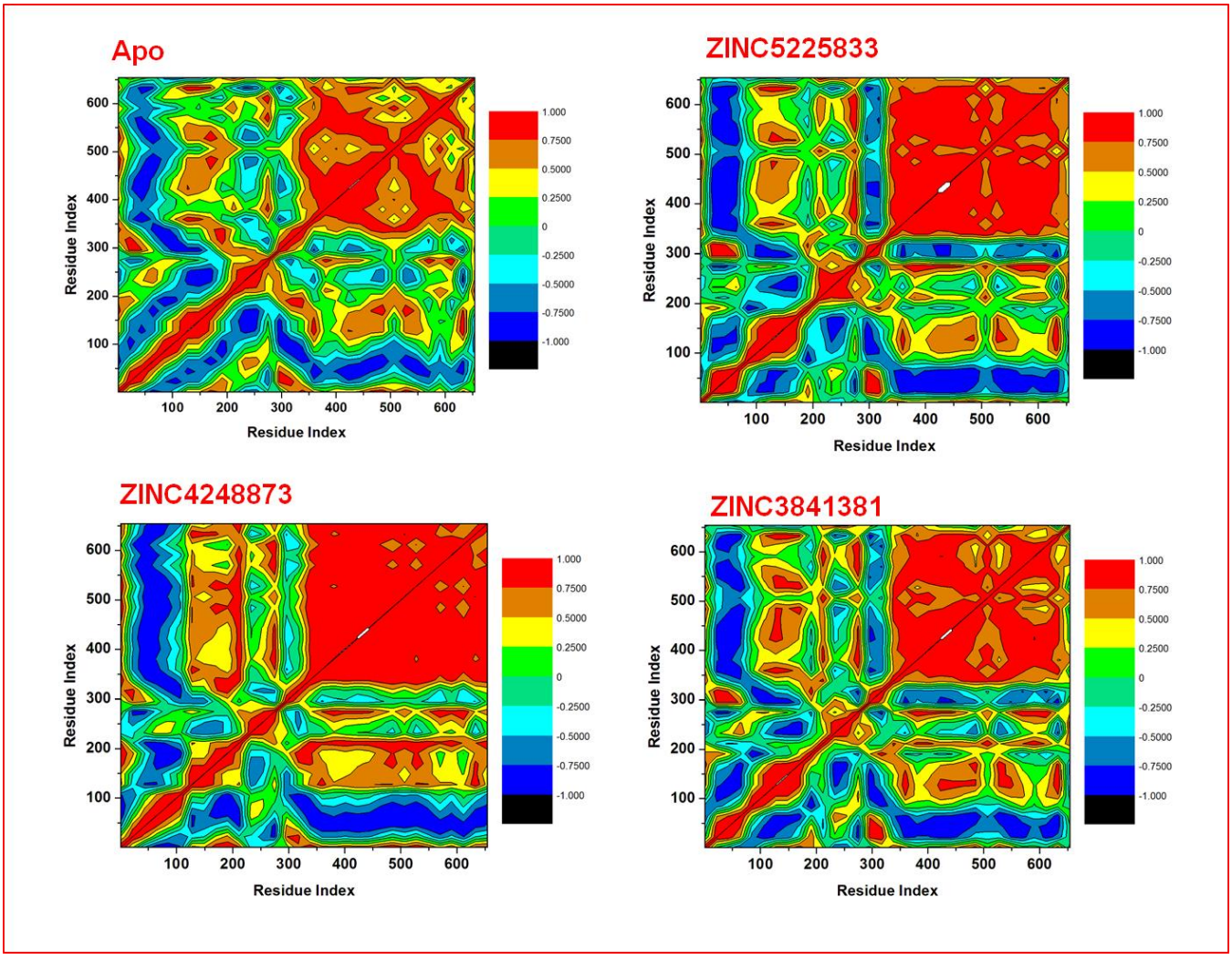


Figure 7.9 The dynamics cross-correlation matrix depicting the correlation and displacement of

residues in apo and the complexes.

In addition to the DCCM analysis, the conformational motions of apo and complexes were investigated by principal component analysis (Figure 7.10). PCA calculates the direction (eigenvector) and amplitude (eigenvalue) of protein motion along a simulation trajectory [81]. The binding of a ligand to a protein can either induce a more compact or dispersed motion of the residues of the protein. In Figure 7.10A, the binding of ligands caused a more dispersed protein movement than apo. *Cp*IMPDH-ZINC5225833, which had the least binding energy, was the most dispersed, while *Cp*IMPDH-ZINC4258873 was the least dispersed. Examining the ligands' individual motions (Figure 7.10B), we observed that ZINC3841381which had the most favorable energy induced the most compact motion, followed by ZINC4258873, and the most dispersed was ZINC5225833.

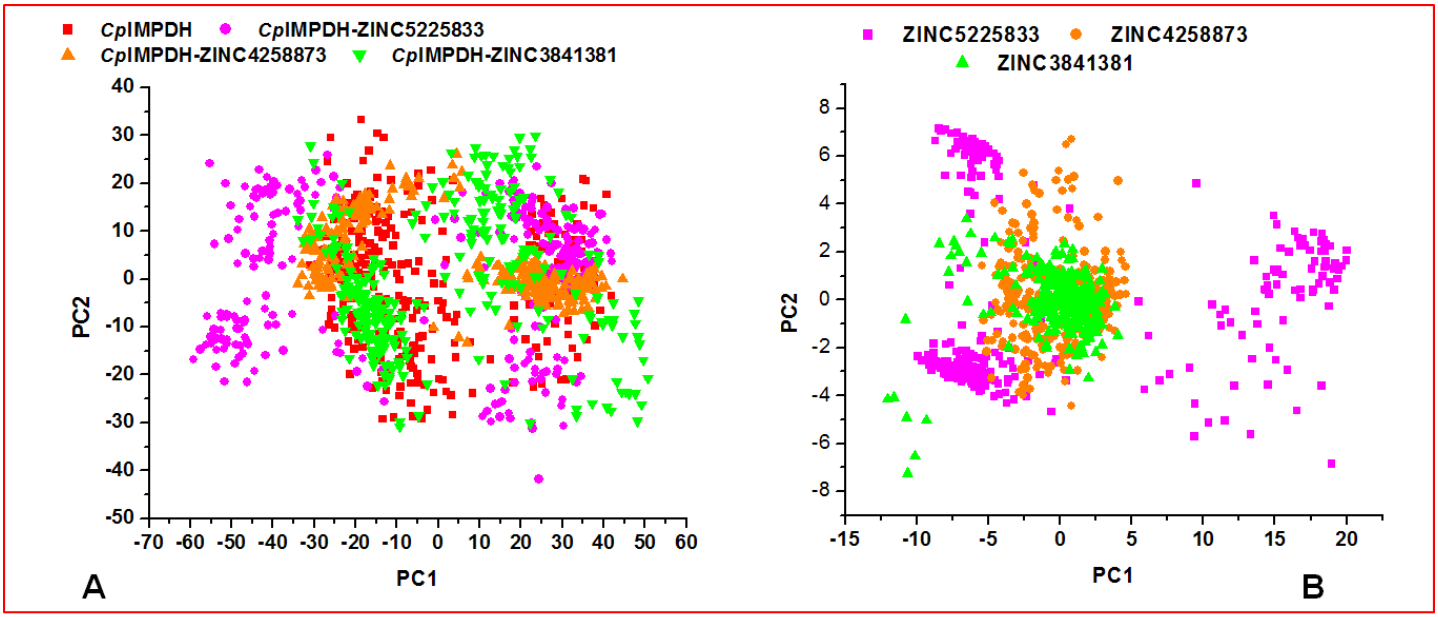


Figure 7.10 The principal component analysis showcasing the mobility of the residue in the **A)** global structure of apo and the complexes and **B)** ZINC5225833, ZINC4258873, and ZINC3841381

7.3.9 Evaluation of drug-likeness of ZINC5225833, ZINC4258873, and ZINC3841381

The effectiveness of a drug is dependent on it reaching the right biological target and executing its therapeutic function without mediating unfavorable toxicities. Therefore, it is imperative to evaluate the propensity of potential drug's absorption, distribution, metabolism, excretion, and

toxicity in the early stage of drug development. Early evaluation of these parameters saves time and cost. The natural compounds pharmacokinetics parameters were subjected to the Lipinski rule of 5, which has been extensively used for assessing drug-likeness. Lipinski posited that for a drug to be regarded as orally active, it must not run afoul of more than two criteria from the following parameters: octanol-water partition effect of not greater than 5, not more than 10 hydrogen bond acceptors and 5 hydrogen bond donors, lipophilicity not greater than 5 and a molecular weight of less than 500g/mol [60,82]. The lipophilicity index, the number of hydrogen bond donor and acceptors of the three compounds were well within the acceptable threshold described by Lipinski (Table 4). However, the molecular sizes of the natural compounds were higher than 500g/mol. Despite this, going by Lipinski's rule of drug-likeness, the compounds still have good ADMET properties. Their lipophilicity value (logP) of the ligands points out that these ligands are cell-permeable and have a high propensity to be bioavailable [83].

The synthetic accessibility is an index for assessing the ease of synthesizing a compound. It ranges between 1-10, representing graduation from the easiest to the hardest to synthesize [84]. From the ligands' SA values in Table 4, it is deducible that the ligands are relatively easier to synthesize and may not pose insurmountable challenges during this process. The topological polar surface area (TPSA) of the ligands was evaluated, the lower the TPSA value, the easier for the compound to be absorbed into the cell and vice versa [85]. The ligands' TPSA values fell within the acceptable limits ($140A^2$). However, of all the three ligands, ZINC3841381, which had the most favorable binding, may have some difficulty permeating the cell compared to other ligands as it had the highest TPSA value of $125.1A^2$.

Table 7.4 Pharmacokinetic profile of the natural compounds assessing their suitability as drug candidates.

Parameters	ZINC5225833	ZINC4258873	ZINC3841381	Acceptable Threshold
Molecular Formular	C34H29FN2O	C31H29N3O4	C26H23N3O4S2	
Molecular weight (g/mol)	500.61	507.58	505.61	≤500
Synthetic accessibility	4.76	4.99	4.98	10

Lipophilicity (logP)	4.31	3.90	3.25	≤5
TPSA (A^2)	34.03	80.64	125.1	≤140
H-Bond Acceptor	2	4	4	≤10
H-Bond Donor	1	1	1	≤5
Rotatable Bonds	3	7	5	≤10
LE (kcal/mol/heavy atom)	0.32	0.32	0.35	> 0.3
LLE	2.34	4.79	5.04	> 5
LELP	20.49	12.97	11.23	-10 to +10
Bioavailability score	0.17	0.55	0.55	
Water Solubility	Poorly soluble	Moderately soluble	Poorly Soluble	
GI Absorption	Low	High	Low	
BBB Permeability	No	No	No	
Purchasability	Yes	Yes	Yes	

Ligand efficiency (LE), ligand lipophilic efficiency (LLE), and ligand-efficiency dependent lipophilicity (LELP)[86] were evaluated in addition to the Lipinski rule of 5. LE gives information about the potency and affinity of a ligand, LLE measures both the potency and lipophilicity of a ligand, and LELP finds the association between ligand's molecular size, potency, and lipophilicity [83,87–89]. In Table 4, all the ligands are within the acceptable limits of LE, LLE but not LELP. This implies that the ligands have a favorable potency, lipophilicity index but unfavorable molecular weight, as earlier pointed out by the Lipinski rule of five. The natural compounds through in silico ADMET evaluation have shown some promise as drug-likely candidates. However, they can still be modified to improve their molecular weights, potencies, and effectiveness as therapeutic candidates.

7.4 Conclusion

The *in-silico* evaluation of the natural compounds as potential anticryptosporidials projected ZINC5225833, ZINC4258873, and ZINC3841381 as having the best complementarity to the *Cp*IMPDH-NAD^+ binding site. The estimation of these ligands' binding free energies also buttresses their propensity to inhibit the biological functions of *Cp*IMPDH in mediating their anti-

parasitic actions. The findings herein can serve as groundwork which can inform the experimental probing of these three natural compounds as promising anticryptosporidials. Further works on the qualitative structure-activity relationship of the identified potential *Cp*IMPDH inhibitor can reveal crucial functional groups of the ligands that can be modified to elicit better inhibitory activities.

Acknowledgement

We acknowledge the Centre for High Performance Computing, Cape Town for computational resources.

Funding Statement

The authors received no financial support for the study.

Disclosure Statement

The authors declare no conflict of interest.

Data Availability Statement

All data generated in this study are included in this manuscript and its supplementary file.

7.5 References

[1] Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. Lancet. 2013;382:209–222.

[2] Mor S, Tzipori S. Cryptosporidiosis in children in sub-Saharan Africa: A lingering challenge. 2008.

[3] Shirley D-AT, Moonah SN, Kotloff KL. Burden of disease from cryptosporidiosis. Curr Opin Infect Dis [Internet]. 2012 [cited 2019 Oct 26];25:555–563. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22907279.

[4] Ahmadpour E, Safarpour H, Xiao L, et al. Cryptosporidiosis in HIV-positive patients and related risk factors: A systematic review and meta-analysis [Internet]. Parasite. EDP

Sciences; 2020 [cited 2020 Oct 8]. Available from: /pmc/articles/PMC7191976/?report=abstract.

[5] Amadi B, Mwiya M, Sianongo S, et al. High dose prolonged treatment with nitazoxanide is not effective for cryptosporidiosis in HIV positive Zambian children: A randomised controlled trial. BMC Infect Dis. 2009;9.

[6] Amadi B, Mwiya M, Musuku J, et al. Effect of nitazoxanide on morbidity and mortality in Zambian children with cryptosporidiosis: A randomised controlled trial. Lancet [Internet]. 2002 [cited 2020 Oct 12];360:1375–1380. Available from: https://pubmed.ncbi.nlm.nih.gov/12423984/.

[7] Rossignol JFA, Ayoub A, Ayers MS. Treatment of diarrhea caused by Cryptosporidium parvum: A prospective randomized, double-blind, placebo-controlled study of nitazoxanide. J Infect Dis [Internet]. 2001 [cited 2020 Oct 12];184:103–106. Available from: https://pubmed.ncbi.nlm.nih.gov/11398117/.

[8] Pilla AM, Rybak MJ, Chandrasekar PH. Spiramycin in the Treatment of Cryptosporidiosis. Pharmacother J Hum Pharmacol Drug Ther [Internet]. 1987 [cited 2020 Oct 12];7:188–190. Available from: https://pubmed.ncbi.nlm.nih.gov/3432095/.

[9] Griffiths JK, Balakrishnan R, Widmer G, et al. Paromomycin and geneticin inhibit intracellular Cryptosporidium parvum without trafficking through the host cell cytoplasm: Implications for drug delivery. Infect Immun [Internet]. 1998 [cited 2020 Oct 12];66:3874–3883. Available from: /pmc/articles/PMC108441/?report=abstract.

[10] Gathe JC, Mayberry C, Clemmons J, et al. Resolution of severe cryptosporidial diarrhea with rifaximin in patients with AIDS [Internet]. J. Acquir. Immune Defic. Syndr. J Acquir Immune Defic Syndr; 2008 [cited 2020 Jun 28]. p. 365–366. Available from: https://pubmed.ncbi.nlm.nih.gov/18580340/.

[11] Maurya SK, Gollapalli DR, Kirubakaran S, et al. Triazole inhibitors of Cryptosporidium parvum inosine 5’-monophosphate dehydrogenase. J Med Chem [Internet]. 2009 [cited

2019 Oct 19];52:4623–4630. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19624136.

[12] Kim Y, Makowska-Grzyska M, Gorla SK, et al. Structure of Cryptosporidium IMP dehydrogenase bound to an inhibitor with in vivo antiparasitic activity. Acta Crystallogr Sect FStructural Biol Commun. 2015;71:531–538.

[13] Gorla SK, Kavitha M, Zhang M, et al. Selective and potent urea inhibitors of cryptosporidium parvum inosine 5′-monophosphate dehydrogenase. J Med Chem. 2012;55:7759–7771.

[14] Sun Z, Khan J, Makowska-Grzyska M, et al. Synthesis, in vitro evaluation and cocrystal structure of 4-oxo-[1]benzopyrano[4,3-c]pyrazole Cryptosporidium parvum inosine 5'-monophosphate dehydrogenase (CpIMPDH) inhibitors. J Med Chem [Internet]. 2014 [cited 2019 Oct 19];57:10544–10550. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25474504.

[15] Makowska-Grzyska M, Kim Y, Maltseva N, et al. A novel cofactor-binding mode in bacterial IMP dehydrogenases explains inhibitor selectivity. J Biol Chem. 2015;290:5893–5911.

[16] Wang W, Hedstrom L. Kinetic mechanism of human inosine 5'-monophosphate dehydrogenase type II: random addition of substrates and ordered release of products. Biochemistry [Internet]. 1997 [cited 2019 Oct 26];36:8479–8483. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9214292.

[17] Hedstrom L. IMP dehydrogenase: Structure, mechanism, and inhibition. Chem Rev. 2009;109:2903–2928.

[18] Striepen B, Pruijssers AJP, Huang J, et al. Gene transfer in the evolution of parasite nucleotide biosynthesis. Proc Natl Acad Sci U S A. 2004;101:3154–3159.

[19] Umejiego NN, Li C, Riera T, et al. Cryptosporidium parvum IMP dehydrogenase: Identification of functional, structural, and dynamic properties that can be exploited for drug

design. J Biol Chem. 2004;279:40320–40327.

[20] Zhang RG, Evans G, Rotella FJ, et al. Characteristics and crystal structure of bacterial inosine-5'- monophosphate dehydrogenase. Biochemistry. 1999;38:4691–4700.

[21] Omolabi KF, Agoni C, Olotu FA, et al. Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity — A Structural , Dynamical and Mechanistic Stance. Cell Biochem Biophys [Internet]. 2020; Available from: http://dx.doi.org/10.1007/s12013-020-00950-1.

[22] Mushtaq S, Abbasi BH, Uzair B, et al. Natural products as reservoirs of novel therapeutic agents [Internet]. EXCLI J. Leibniz Research Centre for Working Environment and Human Factors; 2018 [cited 2020 Oct 15]. p. 420–451. Available from: /pmc/articles/PMC5962900/?report=abstract.

[23] Ndjonka D, Ajonina-Ekoti I, Djafsia B, et al. Anogeissus leiocarpus extract on the parasite nematode Onchocerca ochengi and on drug resistant mutant strains of the free-living nematode Caenorhabditis elegans. Vet Parasitol [Internet]. 2012 [cited 2020 Oct 15];190:136–142. Available from: https://pubmed.ncbi.nlm.nih.gov/22727616/.

[24] Cho-Ngwa F, Abongwa M, Ngemenya MN, et al. Selective activity of extracts of Margaritaria discoidea and Homalium africanum on Onchocerca ochengi. BMC Complement Altern Med [Internet]. 2010 [cited 2020 Oct 15];10. Available from: https://pubmed.ncbi.nlm.nih.gov/21029456/.

[25] Smith RA, Pontiggia L, Waterman C, et al. Comparison of motility, recovery, and methyl-thiazolyl-tetrazolium reduction assays for use in screening plant products for anthelmintic activity. Parasitol Res [Internet]. 2009 [cited 2020 Oct 15];105:1339–1343. Available from: https://pubmed.ncbi.nlm.nih.gov/19629527/.

[26] Utzinger J, Shuhua X, N'Goran EK, et al. The potential of artemether for the control of schistosomiasis [Internet]. Int. J. Parasitol. Int J Parasitol; 2001 [cited 2020 Oct 15]. p. 1549–1562. Available from: https://pubmed.ncbi.nlm.nih.gov/11730781/.

[27] Xiao SH, Booth M, Tanner M. The prophylactic effects of artemether against schistosoma

japonicum infections. Parasitol. Today. Elsevier Ltd; 2000. p. 122–126.

[28] Magalhães LG, Machado CB, Morais ER, et al. In vitro schistosomicidal activity of curcumin against Schistosoma mansoni adult worms. Parasitol Res [Internet]. 2009 [cited 2020 Oct 15];104:1197–1201. Available from: https://pubmed.ncbi.nlm.nih.gov/19096877/.

[29] Jisaka M, Kawanaka M, Michael M, et al. Antischistosomal Activities of Sesquiterpene Lactones and Steroid Glucosides from Vernonia amygdalina, Possibly Used by Wild Chimpanzees against Parasite-related Diseases. Biosci Biotechnol Biochem [Internet]. 1992 [cited 2020 Oct 15];56:845–846. Available from: https://pubmed.ncbi.nlm.nih.gov/1368347/.

[30] Soliman MFM. Evaluation of avocado/soybean unsaponifiable alone or concurrently with praziquantel in murine schistosomiasis. Acta Trop [Internet]. 2012 [cited 2020 Oct 15];122:261–266. Available from: https://pubmed.ncbi.nlm.nih.gov/22342904/.

[31] Kato MJ, Furlan M. Chemistry and evolution of the Piperaceae. Pure Appl Chem. 2007;79:529–538.

[32] Da Silva Filho AA, Costa ES, Cunha WR, et al. In vitro antileishmanial and antimalarial activities of tetrahydrofuran lignans isolated from Nectandra megapotamica (Lauraceae). Phyther Res [Internet]. 2008 [cited 2020 Oct 15];22:1307–1310. Available from: https://pubmed.ncbi.nlm.nih.gov/18688887/.

[33] Nour AMM, Khalid SA, Kaiser M, et al. The antiprotozoal activity of sixteen asteraceae species native to sudan and bioactivity-guided isolation of xanthanolides from xanthium brasilicum. Planta Med [Internet]. 2009 [cited 2020 Oct 15];75:1363–1368. Available from: https://pubmed.ncbi.nlm.nih.gov/19431098/.

[34] Wang Z, Sun H, Shen C, et al. Combined strategies in structure-based virtual screening [Internet]. Phys. Chem. Chem. Phys. Royal Society of Chemistry; 2020 [cited 2020 Oct 15]. p. 3149–3159. Available from:

https://pubs.rsc.org/en/content/articlehtml/2020/cp/c9cp06303j.

[35] Lyne PD. Structure-based virtual screening: An overview [Internet]. Drug Discov. Today. Drug Discov Today; 2002 [cited 2020 Oct 15]. p. 1047–1055. Available from: https://pubmed.ncbi.nlm.nih.gov/12546894/.

[36] Waszkowycz B. Towards improving compound selection in structure-based virtual screening [Internet]. Drug Discov. Today. Drug Discov Today; 2008 [cited 2020 Oct 15]. p. 219–226. Available from: https://pubmed.ncbi.nlm.nih.gov/18342797/.

[37] Cavasotto C, W. Orry A. Ligand Docking and Structure-based Virtual Screening in Drug Discovery. Curr Top Med Chem [Internet]. 2007 [cited 2020 Oct 15];7:1006–1014. Available from: https://pubmed.ncbi.nlm.nih.gov/17508934/.

[38] Kitchen DB, Decornez H, Furr JR, et al. Docking and scoring in virtual screening for drug discovery: Methods and applications [Internet]. Nat. Rev. Drug Discov. Nat Rev Drug Discov; 2004 [cited 2020 Oct 15]. p. 935–949. Available from: https://pubmed.ncbi.nlm.nih.gov/15520816/.

[39] Bhardwaj VK, Singh R, Sharma J, et al. Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors. J Biomol Struct Dyn [Internet]. 2020 [cited 2021 Feb 15];1–10. Available from: https://www.tandfonline.com/doi/full/10.1080/07391102.2020.1766572.

[40] Bhardwaj VK, Singh R, Sharma J, et al. Structural based study to identify new potential inhibitors for dual specificity tyrosine-phosphorylation- regulated kinase. Comput Methods Programs Biomed [Internet]. 2020 [cited 2021 Feb 15];194:105494. Available from: https://linkinghub.elsevier.com/retrieve/pii/S016926072030167X.

[41] Schneider G. Virtual screening: An endless staircase? [Internet]. Nat. Rev. Drug Discov. Nat Rev Drug Discov; 2010 [cited 2020 Oct 15]. p. 273–276. Available from: https://pubmed.ncbi.nlm.nih.gov/20357802/.

[42] Shoichet BK. Virtual screening of chemical libraries. Nature. Nature; 2004. p. 862–865.

[43] Eswar N, Webb B, Marti-Renom MA, et al. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci. 2007;Chapter 2:Unit 2.9.

[44] Yang Z, Lasker K, Schneidman-Duhovny D, et al. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. J Struct Biol. 2012;179:269–278.

[45] Trott O, Olson AJ. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–461.

[46] Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–1612.

[47] Case DA. Amber 18. Univ California, San Fr. 2018;

[48] Wang J, Wolf RM, Caldwell JW, et al. Development and testing of a general amber force field. J Comput Chem [Internet]. 2004 [cited 2019 Oct 26];25:1157–1174. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15116359.

[49] Grest GS, Kremer K. Molecular dynamics simulation for polymers in the presence of a heat bath. Phys Rev A. 1986;33:3628–3631.

[50] Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9:3084–3095.

[51] Seifert E. OriginPro 9.1: scientific data analysis and graphing software-software review. J Chem Inf Model [Internet]. 2014 [cited 2019 Oct 26];54:1552. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24702057.

[52] Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov [Internet]. 2015 [cited 2019 Oct 26];10:449–461. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25835573.

[53] Ylilauri M, Pentikäinen OT. MMGBSA as a tool to understand the binding affinities of

filamin-peptide interactions. J Chem Inf Model. 2013;53:2626–2633.

[54] Karthick V, Nagasundaram N, Doss CGP, et al. Virtual screening of the inhibitors targeting at the viral protein 40 of Ebola virus. Infect Dis Poverty. 2016;5.

[55] Lawal M, Olotu FA, Soliman MES. Across the blood-brain barrier: Neurotherapeutic screening and characterization of naringenin as a novel CRMP-2 inhibitor in the treatment of Alzheimer's disease using bioinformatics and computational tools. Comput Biol Med. 2018;98:168–177.

[56] Liao C, Sitzmann M, Pugliese A, et al. Software and resources for computational medicinal chemistry. Future Med. Chem. Future Med Chem; 2011. p. 1057–1085.

[57] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:1–13.

[58] Sander T, Freyss J, Von Korff M, et al. DataWarrior: An open-source program for chemistry aware data visualization and analysis. J Chem Inf Model [Internet]. 2015 [cited 2020 Aug 14];55:460–473. Available from: https://pubs.acs.org/doi/abs/10.1021/ci500588j.

[59] Molinspiration Cheminformatics [Internet]. [cited 2020 May 1]. Available from: https://www.molinspiration.com/.

[60] Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods. 2000;44:235–249.

[61] McConkey B, Sobolev V, Edelman M. The performance of current methods in ligand-protein docking. undefined. 2002;

[62] Trott O, Olson A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem. 2010;31:455–461.

[63] Valley CC, Cembran A, Perlmutter JD, et al. The Methionine-aromatic motif plays a unique role in stabilizing protein structure. 2012 [cited 2020 Oct 25]; Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.M112.374504.

[64] Samanta U, Chakrabarti P. Assessing the role of tryptophan residues in the binding site. Protein Eng [Internet]. 2001 [cited 2020 Oct 26];14:7–15. Available from: https://pubmed.ncbi.nlm.nih.gov/11287674/.

[65] Macias MJ, Wiesner S, Sudol M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands [Internet]. FEBS Lett. FEBS Lett; 2002 [cited 2020 Oct 26]. p. 30–37. Available from: https://pubmed.ncbi.nlm.nih.gov/11911877/.

[66] Kuzmanic A, Zagrovic B. Determination of ensemble-average pairwise root mean square deviation from experimental B-factors. Biophys J. 2010;98:861–871.

[67] Brüschweiler R. Efficient RMSD measures for the comparison of two molecular ensembles. Proteins Struct Funct Bioinforma [Internet]. 2002 [cited 2020 Jun 1];50:26–34. Available from: http://doi.wiley.com/10.1002/prot.10250.

[68] Gunasekaran K, Nussinov R. How Different are Structurally Flexible and Rigid Binding Sites? Sequence and Structural Features Discriminating Proteins that Do and Do not Undergo Conformational Change upon Ligand Binding. J Mol Biol [Internet]. 2007 [cited 2020 Nov 8];365:257–273. Available from: https://pubmed.ncbi.nlm.nih.gov/17059826/.

[69] Gutteridge A, Thornton J. Conformational changes observed in enzyme crystal structures upon substrate binding. J Mol Biol [Internet]. 2005 [cited 2020 Nov 8];346:21–28. Available from: https://pubmed.ncbi.nlm.nih.gov/15663924/.

[70] Hammes-Schiffer S, Benkovic SJ. Relating protein motion to catalysis [Internet]. Annu. Rev. Biochem. Annu Rev Biochem; 2006 [cited 2020 Nov 8]. p. 519–541. Available from: https://pubmed.ncbi.nlm.nih.gov/16756501/.

[71] Lobanov MY, Bogatyreva NS, Galzitskaya O V. Radius of gyration as an indicator of protein structure compactness. Mol Biol. 2008;42:623–628.

[72] Teilum K, Olsen JG, Kragelund BB. Functional aspects of protein flexibility [Internet]. Cell. Mol. Life Sci. Cell Mol Life Sci; 2009 [cited 2020 Aug 18]. p. 2231–2247. Available from: https://pubmed.ncbi.nlm.nih.gov/19308324/.

[73] Murphy KP. Stabilization of protein structure. Methods Mol. Biol. 2001. p. 1–16.

[74] Lee B, Richards FM. The interpretation of protein structures: Estimation of static accessibility. J Mol Biol [Internet]. 1971 [cited 2020 Nov 12];55. Available from: https://pubmed.ncbi.nlm.nih.gov/5551392/.

[75] Gromiha M, Ahmad S. Role of Solvent Accessibility in Structure Based Drug Design. Curr Comput Aided-Drug Des [Internet]. 2005 [cited 2020 Aug 18];1:223–235. Available from: http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1573-4099&volume=1&issue=3&spage=223.

[76] Pace CN, Shirley BA, McNutt M, et al. Forces contributing to the conformational stability of proteins. FASEB J [Internet]. 1996 [cited 2020 Aug 18];10:75–83. Available from: https://pubmed.ncbi.nlm.nih.gov/8566551/.

[77] Lapidus LJ, Yao S, McGarrity KS, et al. Protein hydrophobic collapse and early folding steps observed in a microfluidic mixer. Biophys J [Internet]. 2007 [cited 2020 Aug 18];93:218–224. Available from: https://pubmed.ncbi.nlm.nih.gov/17416618/.

[78] Brun L, Isom DG, Velu P, et al. Hydration of the folding transition state ensemble of a protein. Biochemistry [Internet]. 2006 [cited 2020 Aug 18];45:3473–3480. Available from: https://pubmed.ncbi.nlm.nih.gov/16533028/.

[79] Karplus M, Ichiye T. Comment on a “Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations.” J Mol Biol [Internet]. 1995 [cited 2020 Nov 15];252:492–503. Available from: https://pubmed.ncbi.nlm.nih.gov/7563068/.

[80] Arnold GE, Ornstein RL. Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3. Biophys J. 1997;73:1147–1159.

[81] David CC, Jacobs DJ. Principal component analysis: A method for determining the essential dynamics of proteins. Methods Mol Biol [Internet]. 2014 [cited 2020 Aug 18];1084:193–226. Available from: /pmc/articles/PMC4676806/?report=abstract.

[82] Lipinski CA, Lombardo F, Dominy BW, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46:3–26.

[83] Arnott JA, Planey SL. The influence of lipophilicity in drug discovery and design. Expert Opin Drug Discov. 2012;7:863–875.

[84] Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J Cheminform. 2009;1:8.

[85] Fernandes J, Gattass CR. Topological polar surface area defines substrate transport by multidrug resistance associated protein 1 (MRP1/ABCC1). J Med Chem. 2009;52:1214–1218.

[86] Hopkins AL, Keserü GM, Leeson PD, et al. The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov. 2014;13:105–121.

[87] Bembenek SD, Tounge BA, Reynolds CH. Ligand efficiency and fragment-based drug discovery. Drug Discov. Today. Drug Discov Today; 2009. p. 278–283.

[88] Nissink JWM. Simple size-independent measure of ligand efficiency. J Chem Inf Model. 2009;49:1617–1622.

[89] Reynolds CH, Tounge BA, Bembenek SD. Ligand binding efficiency: Trends, physical basis, and implications. J Med Chem. 2008;51:2432–2438.

CHAPTER 8

8.1 Conclusion

The pooled prevalence of cryptosporidiosis is high in southern Africa, while the immune-compromised patients were observed to be disproportionally affected. The inefficacy of and emerging resistance to the current therapeutic interventions in cryptosporidiosis infection has stimulated concentrated efforts in designing newer therapeutics that hold great potential as anticryptosporidials. The progression of bioinformatics integrated tools in target-based drug discovery has made the advances in cryptosporidiosis chemotherapy possible.

P131 has been noted as a potential drug that inhibits the *Cp*IMPDH better than paromomycin. It also selectively inhibits *Cp*IMPDH and leaves the host IMPDH unaffected. This study elucidated the structural mechanism of the inhibition of *Cp*IMPDH by P131 and its mechanism of selective inhibition. To the best of our knowledge, this is the first study to explore this phenomenon. The total binding free energy evaluated also underscored the basis of the selectivity. We explored how the amino acid residues unique only to the parasitic IMPDH played a vital role in the mechanism of selection and inhibition. Molecular insights provided herein corroborate previous experimental reports and further underpin the basis of *Cp*IMPDH inhibitor selectivity. Our findings could present attractive prospects towards the design of novel anticryptosporidials with improved selectivity and binding affinity against parasitic targets.

Through the per-residue energy decomposition approach, we identified crucial moieties in P131 that are responsible for its inhibitory effect on *Cp*IMPDH. These moieties were used in the building of P131 pharmacophoric scaffold for virtual screening in the ZINC database. This is to mine for other compounds deposited in the ZINC database that will also possess a similar scaffold and may have the propensity to work to a greater degree than P131 using the total binding free energy index as a pointer. The top three compounds were evaluated via *in silico* methods and their proposed mechanism of inhibition was deduced. The identified compounds can serve as a basis for conducting further experimental investigations towards the development of anticryptosporidials, which can overcome the challenges of existing therapeutic options.

Natural compounds have been considered as a safer and easily accessible alternative to synthetic drugs. We virtually screened 107,000 natural compounds against the co-factor site of *Cp*IMPDH to select the top three compounds that displayed the best complementarity to the binding site. These compounds in complex with *Cp*IMPDH were subjected to molecular dynamics solution and analyzed. Also, their pharmacokinetic parameters were evaluated to know the degree of drug-likeness. The results obtained from evaluating the natural products as potential anticryptosporidials can serve as a preliminary background in the further experimental exploration of ZINC5225833, ZINC4258873, and ZINC3841381 as potential anticryptosporidials.

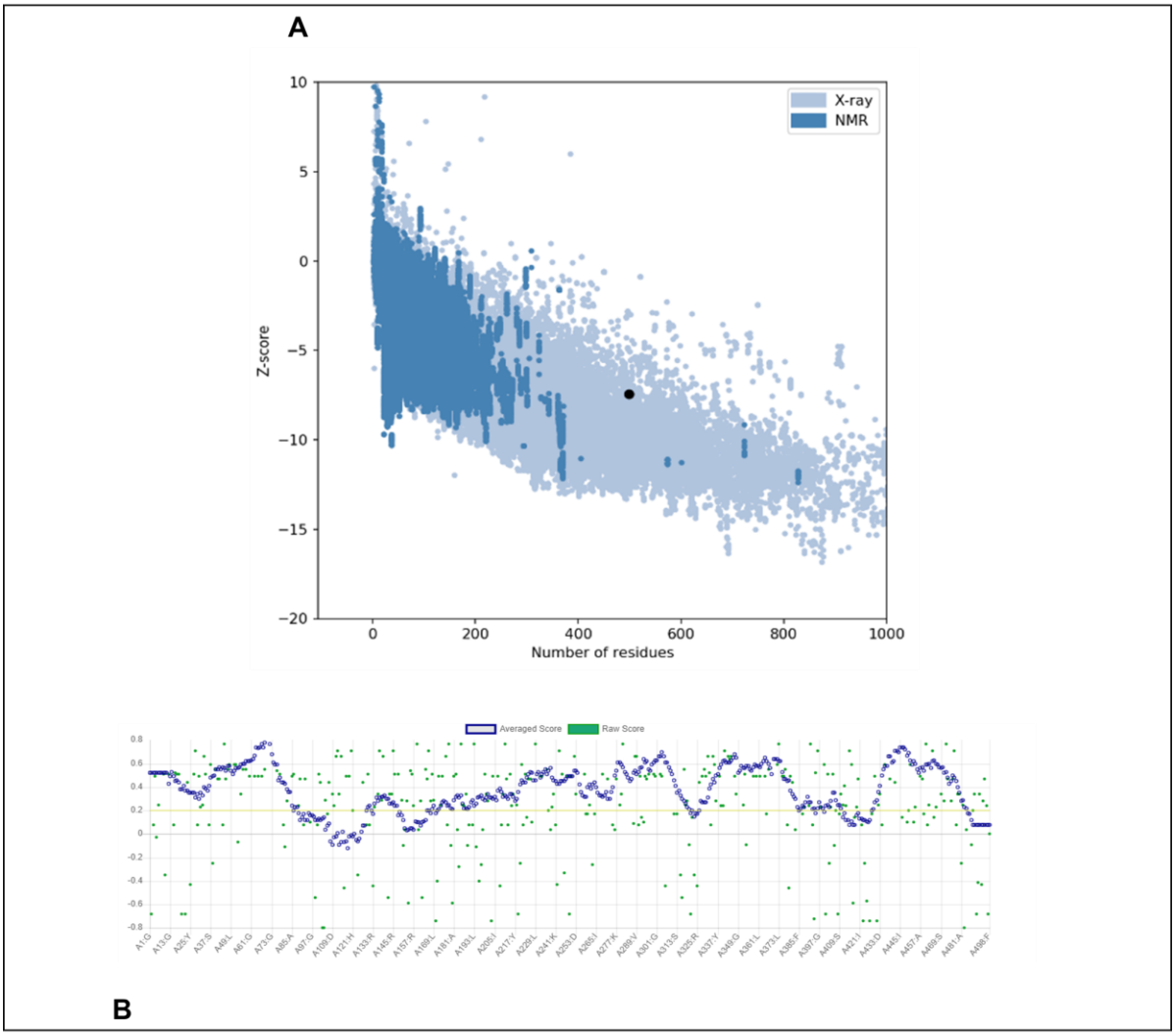
Overall, this study has successfully provided a comprehensive insight into the significance and burden of cryptosporidiosis in the southern African region. It has also highlighted potential anticryptosporidials through ligand-based and structure-based virtual screening, which can be used as starting materials for further experimental findings.

8.2 Future Perspectives

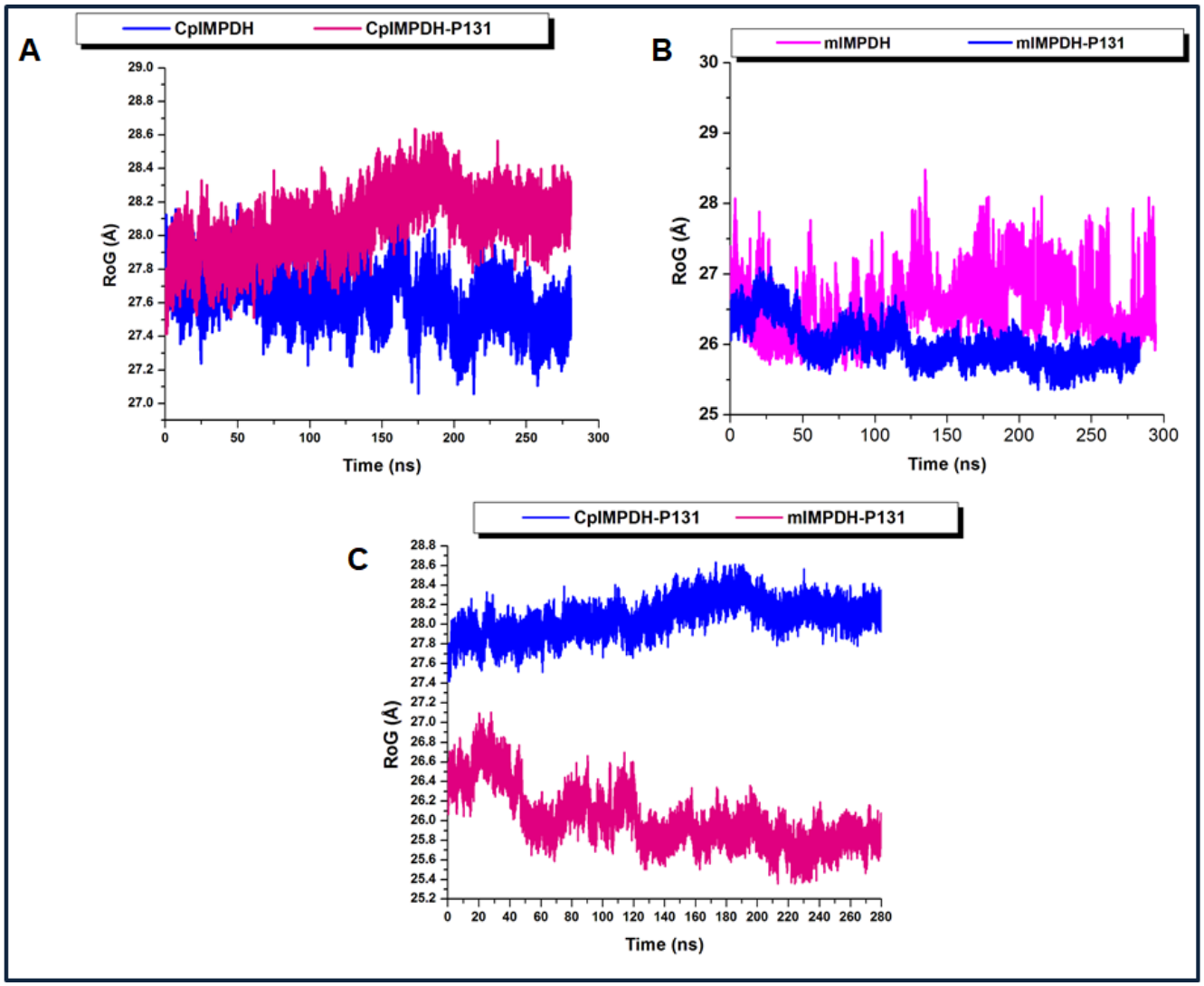
The promising anticryptosporidial compounds discovered in this study through the favorable binding free energies which they exhibited can serve as potential lead compounds. The integrated computational methods also utilized in this study can also facilitate the discovery, advanced design, and optimization of existing anticryptosporidials. However, the compounds highlighted in this study must still have to go through experimental validation to corroborate what we discovered *in silico*.

SUPPLEMENTARY INFORMATION

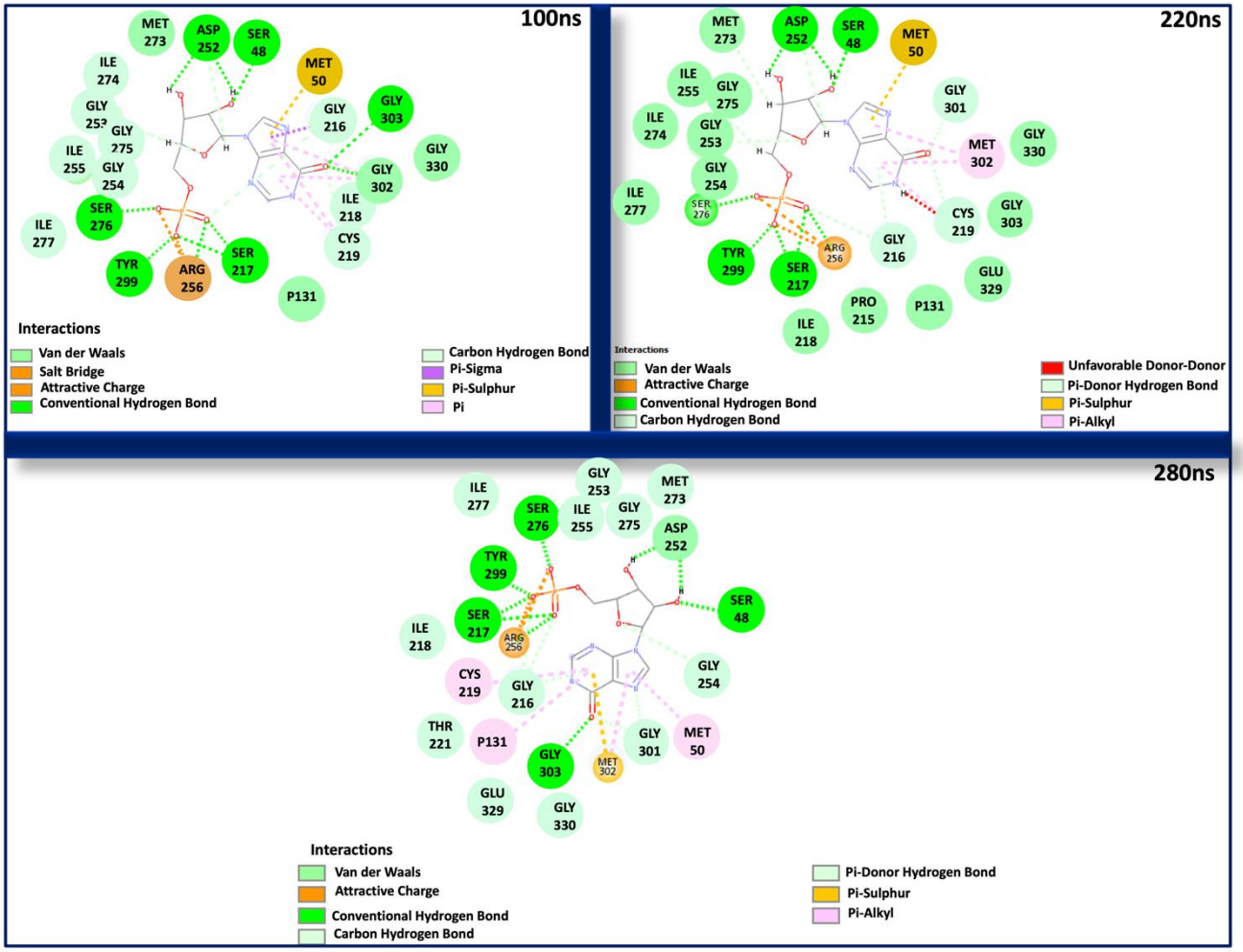
CHAPTER 5



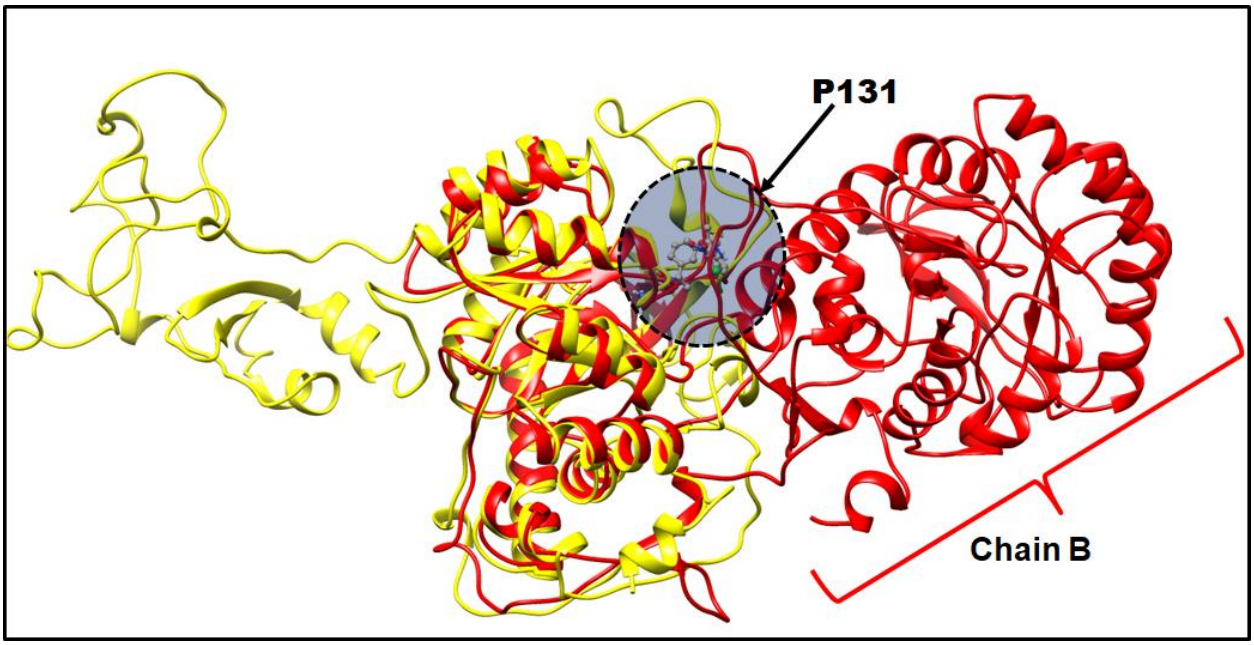
Supplementary Figure 5.1: Graphical representation of the validation of *m*IMPDH. A)PROSA webserver which returned a Z-score of -7.43 and B) VERIFY3D which estimated 80.52% of the residues to have averaged 3D-1D score >= 0.2 score.



Supplementary Figure 5.2: A) Comparative RoG plots of C-a atoms of bound and unbound conformations of *Cp*IMPDH over the simulation period. B) Comparative RoG plots of C-a atoms of bound and unbound conformations of *m*IMPDH over the simulation period. C) Comparative RoG plots of C-a atoms of bound complexes of *Cp*IMPDH and *m*IMPDH over the simulation period.



Supplementary Figure 5.3: Differential binding modes and orientations of P131 at its binding pocket in *IMP* in *Cp*IMPDH at 100ns, 220ns and 280ns. Also highlighted are associated interactions with binding pocket pockets

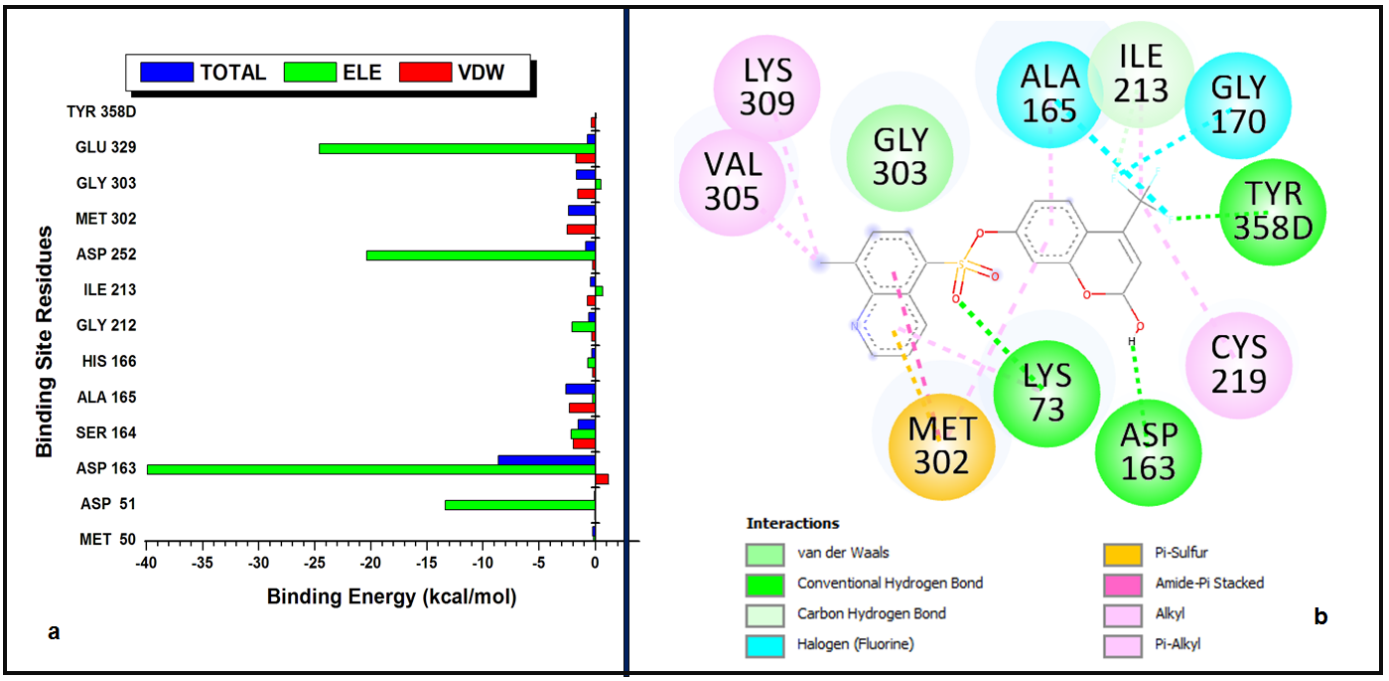


Supplementary Figure 5.4. Superimposition of P131 bound *Cp*IMPDH (red) and with unbound *m*IMPDH (yellow)

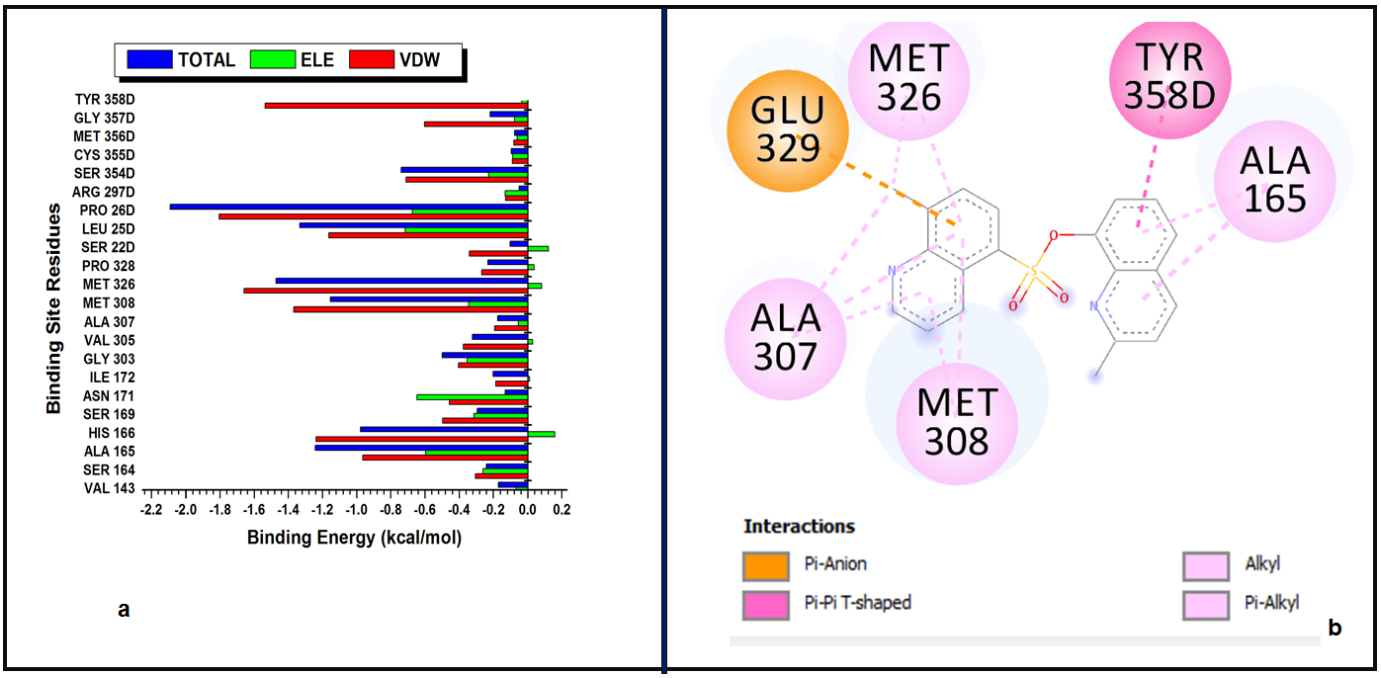
Supplementary Table 5.1: Docking scores of P131 in *m*IMPDH

Docking Output	Docking Score (kcal/mol)
1	-6.1
2	-5.3
3	-5.0
4	-4.8
5	-4.7
6	-4.7
7	-4.2
8	-4.2

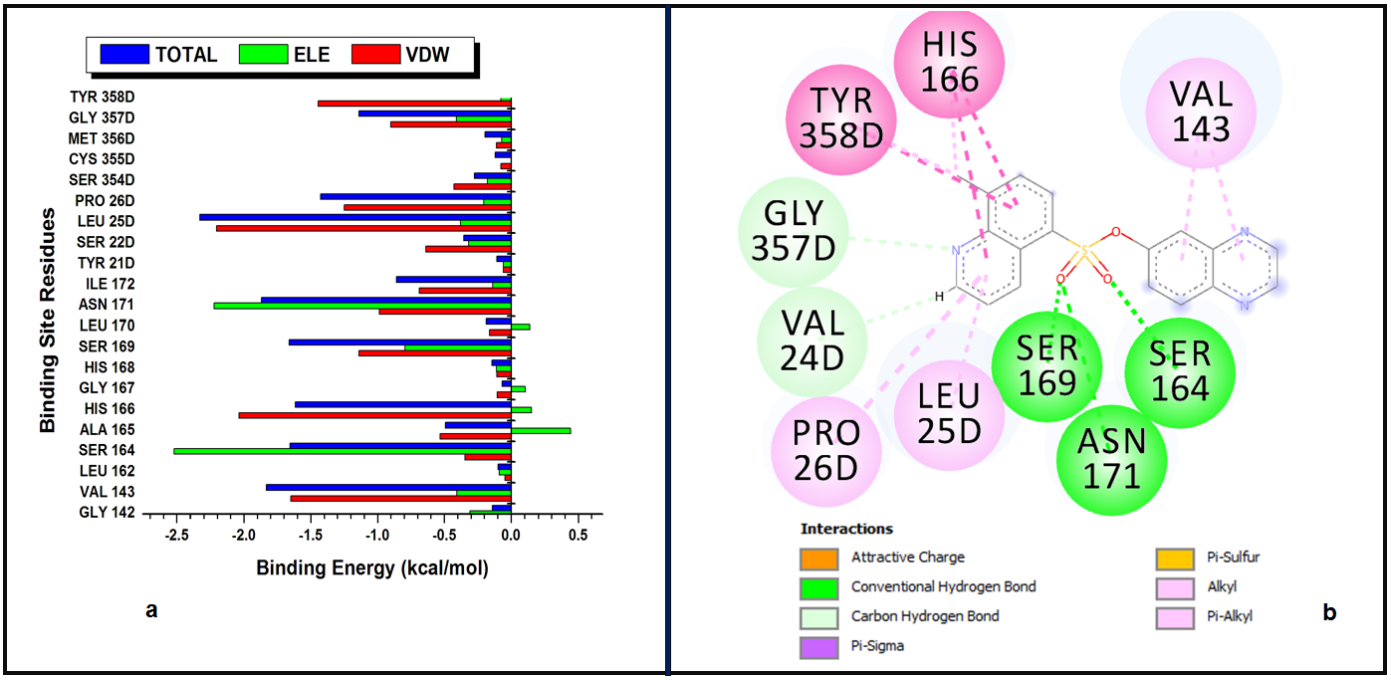
The binding pose of P131 which had the highest negative score (-6.1 kcal/mol) was used



Supplementary Figure 6.1: Overall *Cp*IMPDH-ZINC46542062 interaction and per residue energy decomposition a) Graphical representation of the energy each binding site residue contributed to the total binding free energy of ZINC46542062 in *Cp*IMPDH. b) 2D representation of the binding site residues with ZINC46542062.



Supplementary Figure 6.2: Overall *Cp*IMPDH-ZINC58646829 interaction and per residue energy decomposition. a) Graphical representation of the energy each binding site residue contributed to the total binding free energy of ZINC58646829 in *Cp*IMPDH. b) 2D representation of the binding site residues with ZINC58646829.



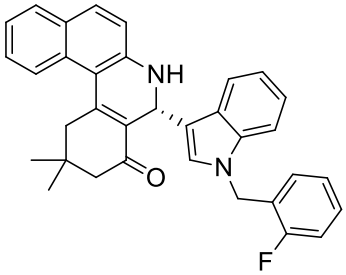
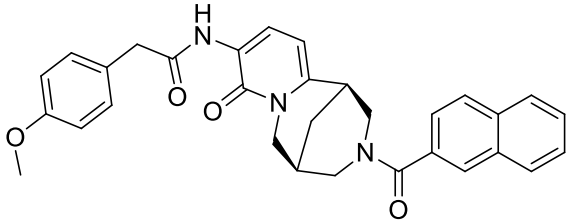
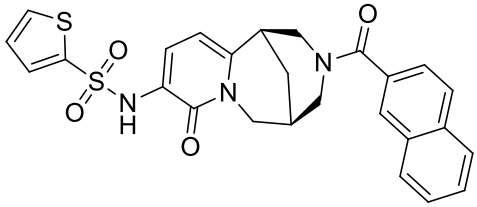
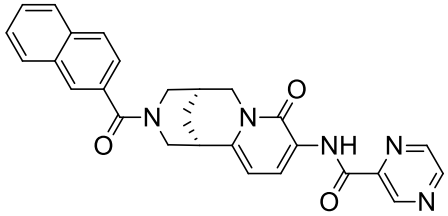
Supplementary Figure 6.3: Overall *Cp*IMPDH-ZINC89780094 interaction and per residue energy decomposition. a) Graphical representation of the energy each binding site residue contributed to the total binding free energy of ZINC89780094 in *Cp*IMPDH. b) 2D representation of the binding site residues with ZINC89780094

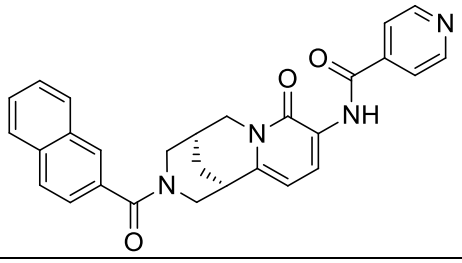
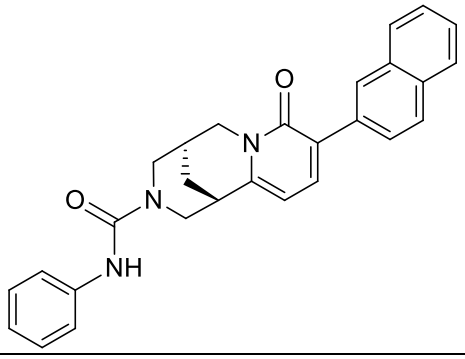
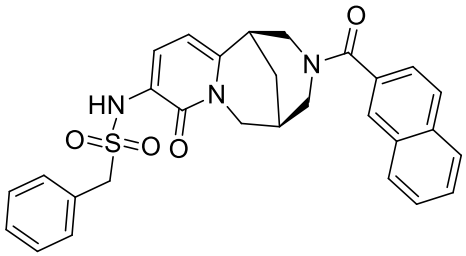
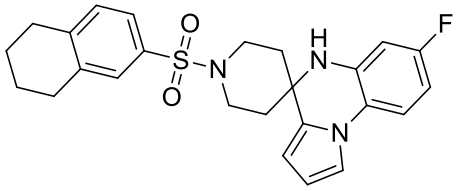
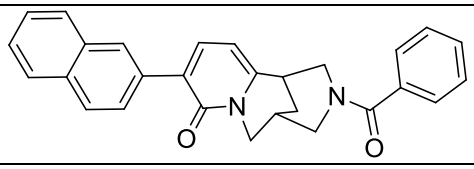
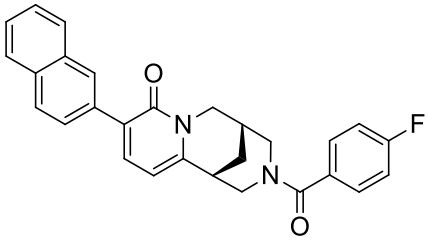
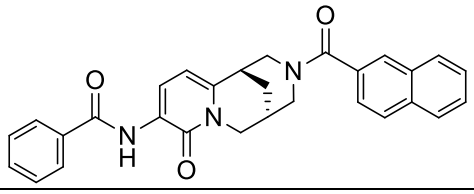
Supplementary Table 6.1: Low energy – contributing residues crucial to the binding of p131 and the identified compounds at the co-factor binding site of *Cryptosporidium parvum* IMPDH

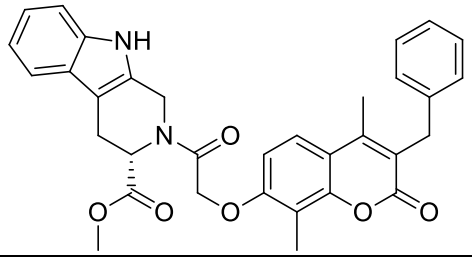
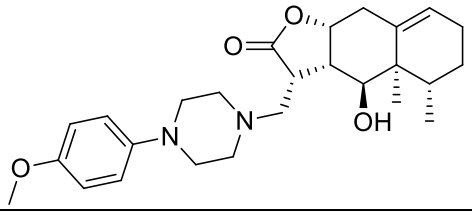
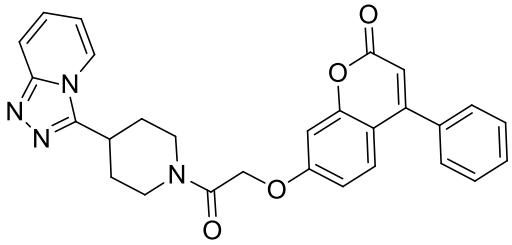
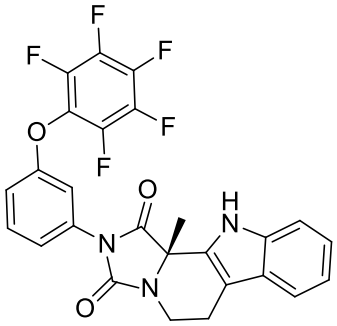
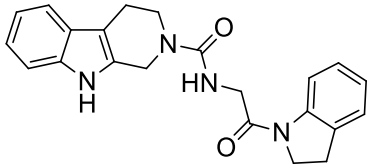
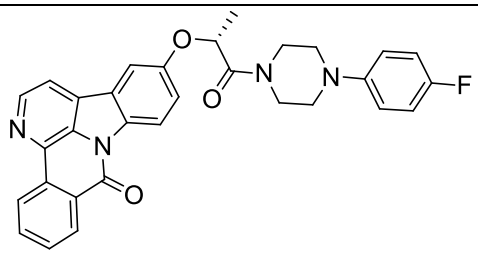
P131 (kcal/mol)	**ZINC46542062 (kcal/mol)**	**ZINC58646829 (kcal/mol)**	**ZINC89780094 (kcal/mol)**
ASP252 (-9.82)	ASP163 (-8.61)	PRO26D (-2.09)	LEU25D (-2.33)
SER164 (-2.62)	ALA165 (-2.63)	TYR358D (-1.80)	ASN171 (-1.87)
MET302 (-1.73)	MET302 (-2.38)	MET326 (-1.47)	VAL143 (-1.83)
TYR358D (-1.70)	GLY303 (-1.67)	LEU25D (-1.33)	SER164 (-1.66)
ALA165 (-1.65)	SER164 (-1.54)	ALA165 (-1.24)	SER169 (-1.66)
PRO26D (-1.28)	ASP252 (-0.9)	MET308 (-1.15)	HIS166 (-1.62)
GLY303 (-0.98)	GLU329 (-0.71)	HIS166 (-1.0)	TYR358D (-1.46)
HIS166 (-0.9)	GLY212 (-0.57)	SER354D (0.74)	PRO26D (-1.42)
SER354D (-0.9)		GLY303 (-0.5)	GLY357D (-1.14)
LEU25D (-0.7)			ILE172 (-0.9)
MET308 (-0.6)			ALA165 (-0.5)

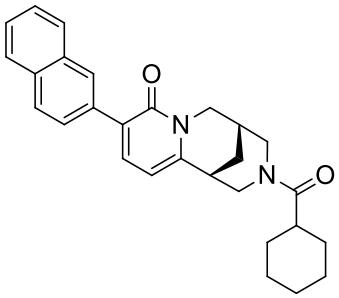
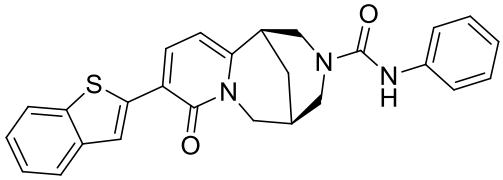
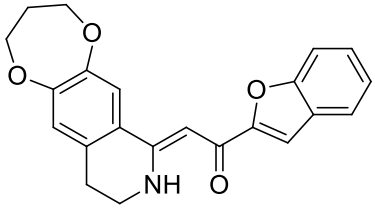
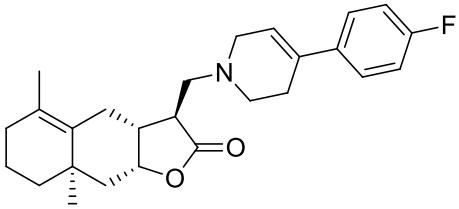
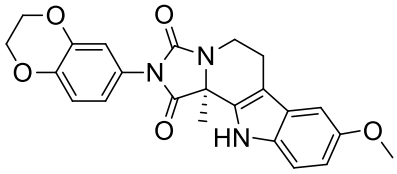
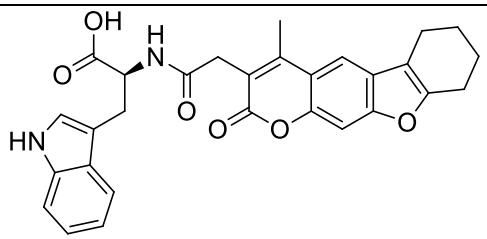
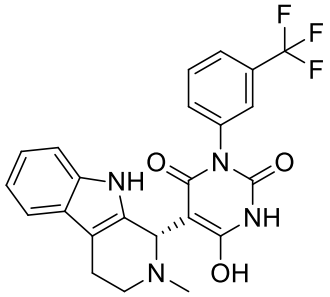
CHAPTER 7

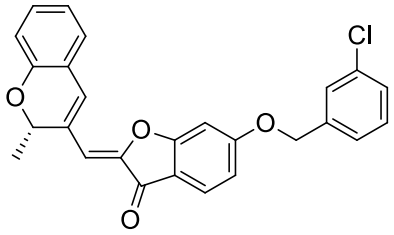
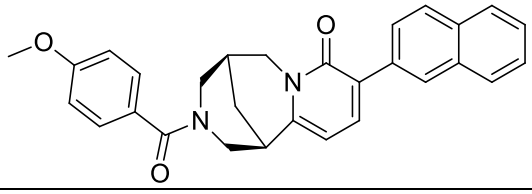
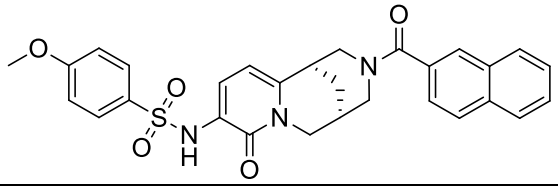
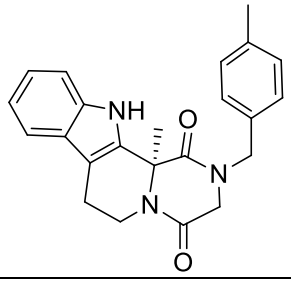
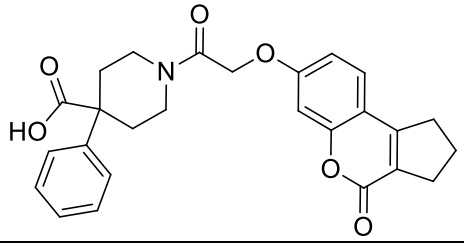
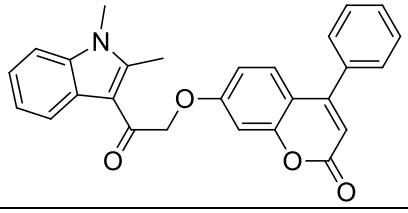
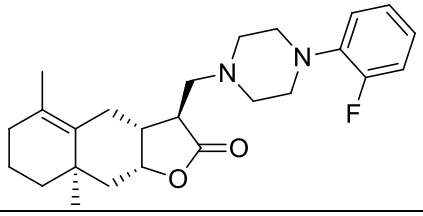
Supplementary Table 7.1: 2D structure of top 50 natural compounds with the highest binding affinity scores.

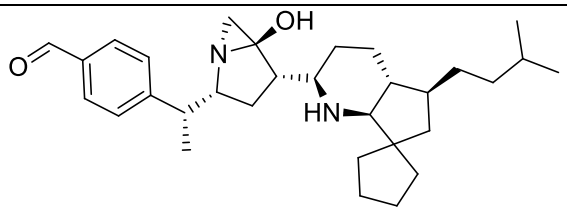
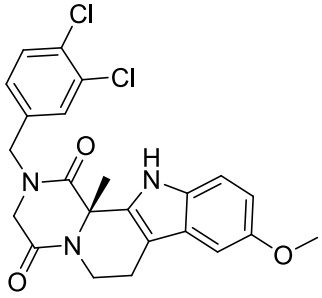
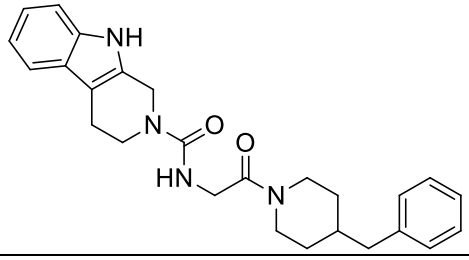
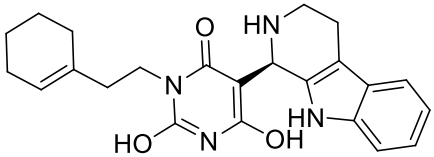
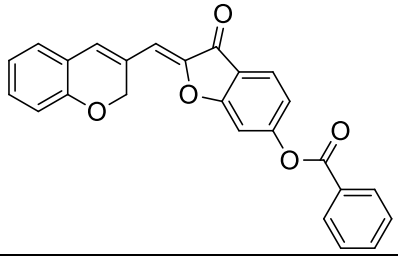
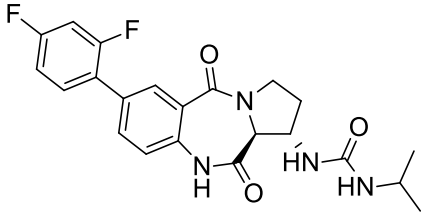
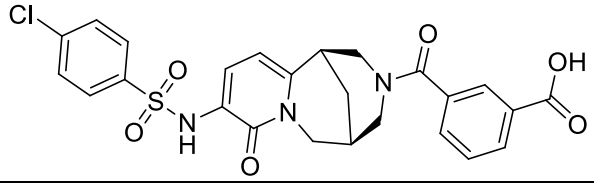
Natural Compound	2D Structure	Docking score (kcal/mol)
ZINC000005225833		-9.8
ZINC000004258873		-9.1
ZINC000003841381		-9.0
ZINC000004258877		-9.0

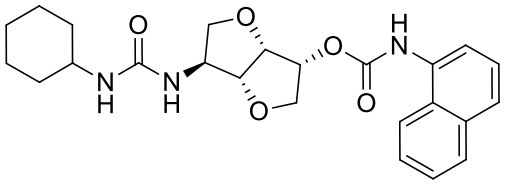
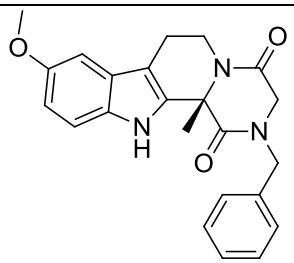
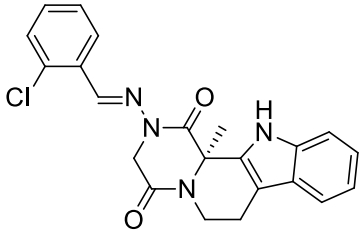
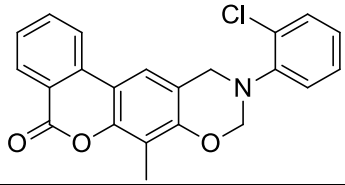
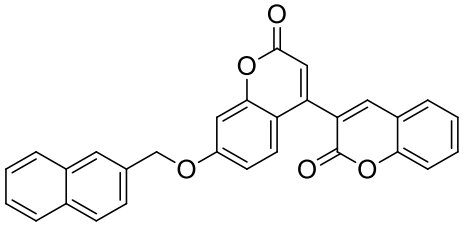
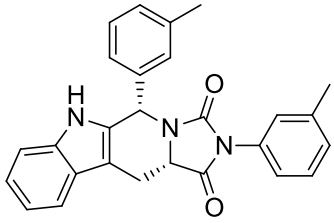
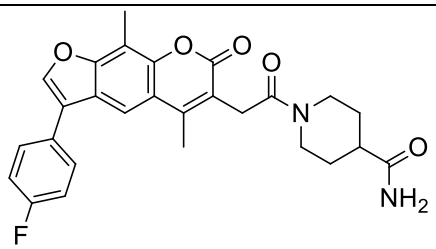
ZINC000004258884		-9.0
ZINC000004237096		-9.0
ZINC000003841380		-8.9
ZINC000035399316		-8.9
ZINC000004237082		-8.9
ZINC000004237087		-8.9
ZINC000004270981		-8.8

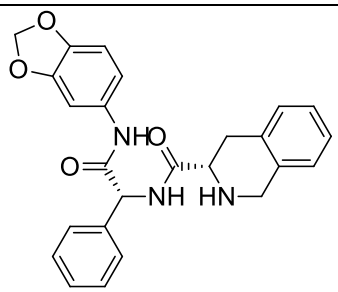
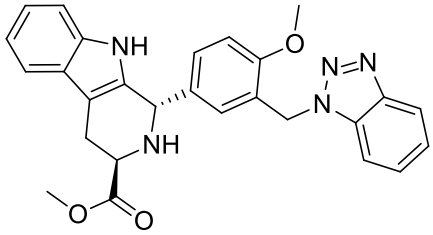
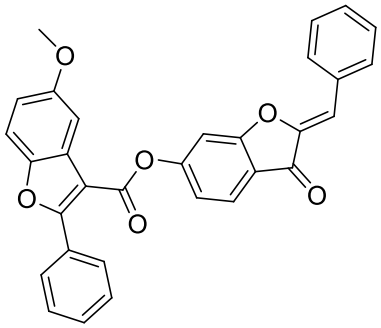
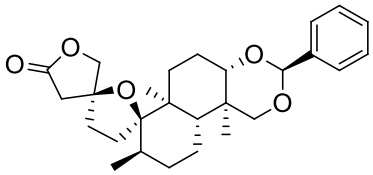
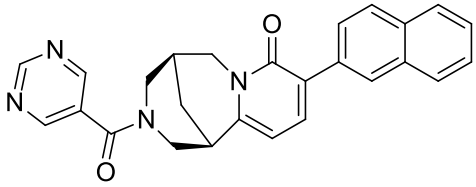
ZINC000070680551		-8.7
ZINC000070680471		-8.7
ZINC000095476565		-8.6
ZINC000009033420		-8.6
ZINC000012296682		-8.6
ZINC000008792179		-8.5

ZINC000004237086		-8.5
ZINC000005410605		-8.5
ZINC000005427070		-8.5
ZINC000004073532		-8.5
ZINC000004073831		-8.5
ZINC000002153223		-8.5
ZINC000008790102		-8.4

ZINC000008791187		-8.4
ZINC000004237083		-8.4
ZINC000004150073		-8.4
ZINC000004029681		-8.4
ZINC000002119921		-8.4
ZINC000002093846		-8.4
ZINC000061992848		-8.4

ZINC000085547785		-8.4
ZINC000008918329		-8.4
ZINC000012296859		-8.4
ZINC000008742613		-8.3
ZINC000006111385		-8.3
ZINC000008296871		-8.3
ZINC000004150056		-8.3

ZINC000004235416		-8.3
ZINC000004044286		-8.3
ZINC000004045305		-8.3
ZINC000002159716		-8.3
ZINC000002127034		-8.3
ZINC000002152734		-8.3
ZINC000002113936		-8.3

ZINC000035485265		-8.3
ZINC000036368047		-8.3
ZINC000008918002		-8.3
ZINC000008952589		-8.3
ZINC000004260408		-8.2

APPENDICES

APPENDIX A

J Parasit Dis
https://doi.org/10.1007/s12639-021-01436-4

REVIEW ARTICLE

A meta-analysis of *Cryptosporidium* species in humans from southern Africa (2000–2020)

Kehinde Foluke Omolabi[1] · Paul Olalekan Odeniran[2] · Mahmoud E. Soliman[1]

Received: 31 March 2021 / Accepted: 5 August 2021

APPENDIX B

Cell Biochemistry and Biophysics
https://doi.org/10.1007/s12013-020-00950-1

ORIGINAL PAPER

Molecular Basis of P131 Cryptosporidial-IMPDH Selectivity—A Structural, Dynamical and Mechanistic Stance

Kehinde F. Omolabi[1] · Clement Agoni[1] · Fisayo A. Olotu[1] · Mahmoud E. S. Soliman[1]

Accepted: 3 October 2020

APPENDIX C

Original Paper | Published: 09 January 2021

A probable means to an end: exploring P131 pharmacophoric scaffold to identify potential inhibitors of *Cryptosporidium parvum* inosine monophosphate dehydrogenase

Kehinde F. Omolabi, Emmanuel A. Iwuchukwu, Clement Agoni, Fisayo A. Olotu & Mahmoud E. S. Soliman

Journal of Molecular Modeling **27**, Article number: 35 (2021) | Cite this article

APPENDIX D


Molecular Simulation ›
Latest Articles
Enter keywords, authors, DOI, ORCID etc
Submit an article
Journal homepage
Articles
‘Finding the needle in the haystack’- will natural products fit for purpose in the treatment of cryptosporidiosis? – A theoretical perspective
Kehinde F. Omolabi, Clement Agoni, Fisayo A. Olotu & Mahmoud E.S. Soliman
Received 07 Dec 2020, Accepted 16 Feb 2021, Published online: 03 Mar 2021
Download citation
https://doi.org/10.1080/08927022.2021.1895435
Check for updates
Full Article
Figures & data
References
Supplemental
Citations
Metrics
Reprints & Permissions
Get access
ABSTRACT
The limitations that have inundated mainstream anticryptosporidials have necessitated the exploration of alternative approaches towards the design of novel
Related articl
Recommended articles
Investigation of ac