

Université de Montréal

**Découverte et optimisation d'inhibiteurs pour
des enzymes DfrBs impliquées dans la résistance
bactérienne**

par Jacynthe Linda Toulouse

Département de biochimie et médecine moléculaire
Faculté de médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de *Philosophiae Doctor*
en biochimie

Décembre, 2018

© Jacynthe Linda Toulouse, 2018

Résumé

Le triméthoprim (TMP) est un antibiotique utilisé en clinique ainsi que de façon préventive dans l'élevage de bétail et l'aquaculture, causant une pression sélective auprès des bactéries. La cible de cet antibiotique est l'enzyme dihydrofolate réductase (Dfr) bactérienne chromosomique qui est essentielle dans la voie de biosynthèse du folate et donc, à la prolifération cellulaire. Plusieurs types de résistances au TMP furent répertoriés, dont le transfert d'un plasmide de résistance entre bactéries pathogènes. En fait, dans le milieu hospitalier, ce plasmide est le plus dangereux, puisque ce milieu est propice à une plus forte sélection provoquant la prolifération des bactéries résistantes au TMP.

Des bactéries pathogènes résistantes au TMP possédant les gènes de Dfr de type B (*dfrB*) ont été identifiées dans l'aquaculture des poissons, l'élevage de bétail ainsi que dans les eaux usées. Les gènes des *dfrBs* sont normalement répertoriés indirectement par l'identification des éléments des intégrons. Ceci fait en sorte que la prévalence des gènes des *dfrBs* est sous-estimée. Il est impératif de déterminer la prévalence de ces gènes. Nous avons démontré le potentiel d'accomplir ceci par une recherche directe de ces gènes dans des échantillons cliniques Nord-Américains de *Escherichia coli* (*E. coli*) résistants au TMP.

À ce jour, aucun antibiotique n'est répertorié pour vaincre la résistance causée par les DfrBs. En fait, peu de travaux sont effectués sur cette famille émergente. Nous savons que les DfrBs ne possèdent aucune similarité de séquence ni de structure avec les Dfr chromosomiques procaryotes et eucaryotes. Il est d'une grande importance de développer des inhibiteurs sélectifs envers ces enzymes pouvant potentiellement servir d'antibiotiques. Éventuellement, une bithérapie avec le TMP et ce nouvel antibiotique augmentera l'efficacité antibactérienne du TMP. Pour y arriver, nous avons déployé plusieurs moyens dont la continuité de la conception d'inhibiteurs basée sur les fragments (FBLD) qui fût rapportée précédemment, la conception d'inhibiteurs visant deux cibles thérapeutiques basée sur la structure des substrats ainsi que le criblage à haut débit d'une librairie d'extraits d'invertébrés marins. Puisque la DfrB1 est la plus étudiée de la famille des DfrBs, nous l'utilisons comme modèle pour ces approches dans la découverte et l'optimisation des inhibiteurs. Lors de l'optimisation des inhibiteurs issus du FBLD, nous avons obtenu des indices provenant du mode de liaison de ces inhibiteurs à la cible

enzymatique et de la caractérisation biophysique de cette liaison afin de diriger les efforts de synthèse de ceux-ci. Ainsi, les connaissances des domaines de l'enzymologie et de la pharmacologie nous permettent de comprendre le mode de liaison de ces inhibiteurs.

Les DfrBs possèdent une similitude de séquence protéique de 77% - 100%. À notre connaissance, aucune caractérisation cinétique des autres membres des DfrBs n'a été effectuée. En fait, pour la plupart des membres, aucune confirmation n'a été obtenue concernant la corrélation de la résistance au TMP. Nous confirmons leur activité dihydrofolate réductase ainsi que leur faible sensibilité au TMP. Avec ces résultats encourageants, nous proposons un mode d'inhibition similaire entre elles. Ainsi, les inhibiteurs de la DfrB1 testés démontrent aussi une inhibition pour les DfrBs et ce, avec la même amplitude. Nos avancées indiquent qu'il serait possible de détenir une tête de série inhibant *tous* les membres de cette famille.

Mots-clés : Dihydrofolate réductases B, résistance aux antibiotiques, triméthoprim, inhibition, cristallographie par rayons X, cinétique enzymatique, courte simulation de dynamique moléculaire *in silico*, prévalence des gènes.

Abstract

Trimethoprim (TMP) is a widely used antibiotic, both clinically as well as preventively in livestock farming and aquaculture. This broad usage causes selective pressure on bacteria. The target of this antibiotic is the chromosomal bacterial dihydrofolate reductase enzyme (Dfr), which is essential in the folate biosynthesis pathway and, therefore, in cell proliferation. Several types of resistance to TMP have been identified, including the transfer of an antibiotic-resistant plasmid between pathogenic bacteria. In clinical environments, this plasmid is of great concern, since this environment is conducive to a greater selective pressure causing the proliferation of TMP-resistant bacteria.

TMP-resistant pathogenic bacteria with the plasmids carrying Dfr type B genes (*dfrB*) have been identified in aquaculture, livestock and wastewater. The *dfrBs* genes are normally tracked indirectly by the identification of integron elements underestimating the prevalence of *dfrBs* genes. Because it is imperative to determine the prevalence of these genes, we have demonstrated the possibility to determine their prevalence by the direct search of these genes in a North American library of TMP-resistant clinical samples.

To date, no antibiotics are known to overcome the resistance caused by DfrBs. In fact, there is limited data on this emerging family. We know that DfrBs are evolutionarily and structurally unrelated to prokaryotic and eukaryotic chromosomal Dfrs. It is of great importance to develop selective inhibitors for these enzymes that could subsequently serve as potential novel antibiotics. Eventually, dual therapy with TMP and this new antibiotic could increase the antibacterial effectiveness of TMP. In order to develop such molecules, we have deployed several approaches including a follow-up of previously reported fragment-based lead design (FBLD), the design of substrate analogs as inhibitors targeting two therapeutic enzymes, as well as high-throughput screening of a library of marine invertebrate extracts. Since DfrB1 is the best studied among the DfrB family, it was used as a model for these approaches. In the optimization of inhibitors from FBLD, we studied the inhibitor's mode of binding to direct synthetic efforts for the next generation of potential inhibitors. Thus, knowledge in the fields of enzymology and pharmacology enhances our understanding of the mode of binding of these inhibitors.

Members of the DfrB family share 77% to 100% protein sequence similarity. To our knowledge, no kinetic studies of DfrB members other than DfrB1 have been performed. For

most members, the correlation of TMP resistance and their source have not been confirmed. We confirm their dihydrofolate reductase activity as well as their resistance to TMP. With these encouraging results, we propose a similar mode of inhibition for all DfrBs. We demonstrate that inhibitors of DfrB1 also inhibit other DfrBs. Our advances indicate that it would be possible to have a lead compound that inhibits *all* members of this family.

Keywords : Dihydrofolate reductases B, antibiotic resistance, trimethoprim, inhibition, X-ray crystallography, enzymatic kinetics, *in silico* low molecular dynamics, prevalence of genes.

Table des matières

RÉSUMÉ	III
ABSTRACT	V
TABLE DES MATIÈRES	VII
LISTE DES TABLEAUX	XI
LISTE DES FIGURES	XIII
LISTE DES SCHÉMAS	XV
LISTE DES ABRÉVIATIONS	XVI
REMERCIEMENTS	XIX
CHAPITRE 1 INTRODUCTION	1
1.1 . Résistance aux antibiotiques.....	1
1.1.1 . L'importance des antibiotiques	1
1.1.2 . La cible du triméthoprimé	2
1.1.3 . Le triméthoprimé et son utilisation considérable	4
1.1.4 . La résistance au triméthoprimé.....	5
1.1.5 . Les intégrons; une machinerie génétique complexe.....	6
1.2 . La famille protéique des dihydrofolates réductases de type B (DfrB)	7
1.2.1 . Les dihydrofolates réductases de type B : les moutons noirs des dihydrofolates réductases.....	7
1.2.2 . La structure secondaire, tertiaire et quaternaire de la DfrB1	9
1.2.3 . Le site actif de la DfrB1	11
1.3 . Les prérequis dans le développement d'une tête de série.....	12
1.4 . Les approches utilisées pour la découverte d'inhibiteurs de la DfrB1	13
1.4.1 . La conception d'inhibiteurs basée sur les fragments (FBLD).....	13
1.4.2 . La conception d'inhibiteurs basée sur les fragments (FBLD) : le cas de la DfrB1	16
1.4.3 . La conception d'inhibiteurs visant deux cibles thérapeutiques, basée sur la structure des substrats.....	19
1.4.4 . Les groupements chimiques importants dans la liaison des substrats	21
1.4.5 . Le criblage à haut débit d'une librairie de sources naturelles : les extraits d'invertébrés marins	22

1.4.6 . Détermination de la qualité d'un criblage à haut débit.....	24
1.5 . Conclusion	25
PRÉFACE AU CHAPITRE 2	26
CHAPITRE 2 - « INTEGRON-ASSOCIATED DFRB4, A PREVIOUSLY UNCHARACTERIZED MEMBER OF THE TRIMETHOPRIM-RESISTANT DIHYDROFOLATE REDUCTASE B FAMILY, IS A CLINICALLY IDENTIFIED EMERGENT SOURCE OF ANTIBIOTIC RESISTANCE. »	27
2.1 . Article 1	27
2.2 . ABSTRACT	28
2.3 . INTRODUCTION	28
2.4 . CLINICAL SAMPLE LIBRARY	29
2.5 . <i>IN SILICO</i> SCREENING	30
2.6 . RECONSTRUCTION OF THE <i>DFRB4</i> MOBILE GENETIC ELEMENT AND CONTIGUOUS DNA SEGMENTS	30
2.7 . THE <i>DFRB4</i> PROTEIN CONFERS TMP RESISTANCE.....	32
2.8 . CONCLUSION	32
2.9 . ACKNOWLEDGEMENTS.....	33
2.10 . ABBREVIATIONS	33
2.11 . REFERENCES	33
PRÉFACE AU CHAPITRE 3	38
CHAPITRE 3 – « DIMERIC BISBENZIMIDAZOLE INHIBITORS OF AN EMERGENT TRIMETHOPRIM-RESISTANT TYPE II DIHYDROFOLATE REDUCTASE GUIDE THE DESIGN OF MONOMERIC ANALOGS. »	40
3.1 . Article 2	40
3.2 . ABSTRACT	41
3.3 . INTRODUCTION	42
3.4 . RESULTS AND DISCUSSION.....	45
3.5 . CONCLUSION	67
3.6 . ACKNOWLEDGMENTS	67
3.7 . EXPERIMENTAL PROCEDURES.....	68
3.8 . SUPPORTING INFORMATION.....	90
3.9 . ABBREVIATION	90

3.10 . REFERENCES	91
PRÉFACE AU CHAPITRE 4	96
CHAPITRE 4 – « INHIBITION OF AN EMERGENT FAMILY OF INTRINSICALLY TRIMETHOPRIM-RESISTANT TYPE II DIHYDROFOLATE REDUCTASES WITH DUAL TARGET INHIBITORS »	98
4.1 . Article 3	98
4.2 . ABSTRACT.....	99
4.3 . INTRODUCTION	99
4.4 . RESULTS AND DISCUSSION.....	102
4.5 . CONCLUSION	110
4.6 . SUPPORTING INFORMATION.....	111
4.7 . ABBREVIATIONS	111
4.8 . REFERENCES	111
CHAPITRE 5 - DÉVELOPPEMENT D'UN CRIBLAGE À HAUT DÉBIT POUR UNE LIBRAIRIE D'EXTRAITS D'INVERTÉBRÉS MARINS.	116
5.1 . Introduction	116
5.2 . Méthodes	116
5.3 . Résultats.....	118
5.3.1 . Détermination du facteur Z'	118
5.3.2 . Criblage des extraits d'invertébrés marins	119
5.4 . Conclusion	121
CHAPITRE 6 - DISCUSSION ET CONCLUSION	122
6.1 . La prévalence et l'émergence des gènes <i>dfrBs</i>	122
6.2 . La résistance au TMP transposée aux DfrBs.....	123
6.3 . Les inconvénients présagés des inhibiteurs	124
6.4 . L'optimisation des inhibiteurs de type bisubstrat.....	126
6.5 . L'optimisation des inhibiteurs de type bisbenzimidazole	128
6.6 . La recherche d'inhibiteurs allostériques ciblant l'interface protéine-protéine	130
6.7 . Conclusion	131
BIBLIOGRAPHIE	I
ANNEXE 1 - MATÉRIEL SUPPLÉMENTAIRE DU CHAPITRE 2.....	IX

ANNEXE 2 - MATÉRIEL SUPPLÉMENTAIRE DU CHAPITRE 3.....	X
ANNEXE 3 - MATÉRIEL SUPPLÉMENTAIRE DU CHAPITRE 4.....	XVI
ANNEXE 4 - PROCÉDURE EXPÉRIMENTALE POUR LA SOLUBILISATION ET LE CRIBLAGE DE LA LIBRAIRIE DES EXTRAITS D'INVERTÉBRÉS.....	XXVI
Solubilisation des extraits d'invertébrés marins.....	xxvi
Criblage des extraits d'invertébrés marins	xxvi

Liste des tableaux

Tableau 1-I. Nomenclature des gènes encodant les Dfrs.....	8
Table 2-I. DNA sequence query coverage and identity (in brackets) of the main type II DHFRs with the expected value (in bold) (coding sequence only).....	29
Table 3-I. Inhibition upon central core substitution and alteration of the central core length of 3.2	47
Table 3-II. Inhibition with analogs lacking benzimidazole moieties.	53
Table 3-III. Inhibition upon substitution of the phenoxy moieties.....	55
Table 3-IV. IC ₅₀ upon substitution of the terminal carboxylates.....	59
Table 3-V. Inhibition with 1:1 target:ligand stoichiometry.....	60
Table 3-VI. Inhibition with non-symmetrical benzimidazole-based compounds.....	63
Table 3-VII. Kinetic parameters for DfrB1 and DfrB4.....	66
Table 3-VIII. Inhibition of DfrB1 and DfrB4 with selected inhibitors.....	66
Table 4-I. Inhibition of His ₆ -DfrB1 and HPPK with bisubstrate inhibitors 4.1 and 4.2	103
Table 4-II. Inhibition of His ₆ -DfrB1 with bisubstrate inhibitors 4.3 to 4.5	105
Table 4-III. Inhibition of DfrB1 (without His ₆ -tag) with bisubstrate and bisbenzimidazole inhibitors 4.1 , 4.3 , 4.6 and 4.7	107
Table 4-IV. Kinetic constants for the dihydrofolate reductase activity of DfrBs.....	108
Table 4-V. Inhibition of DfrBs with TMP and inhibitors 4.1 , 4.3 , 4.6 and 4.7	109
Table A1-I. DfrB Gene cassettes obtained from Genbank as templates for in silico screening.....	viii
Table A2-I. Crystallographic data for DfrB1- 3.1 and DfrB1- 3.3 (PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’).....	ix
Table A2-II. Relative binding frequency per residue of DfrB1 with all inhibitors, resulting from the compilation of simulated poses. Inhibitors are listed in the order of their appearance in the Results section.....	x
Table A2-III. Type of interactions between inhibitor segments and amino acids of DfrB1 determined using the Protein Ligand Interaction Fingerprints (PLIF) module in MOE (CCG).....	xi

Table A2-IV. The methylester of 3.38 does not inhibit DfrB1.....	xii
Table A2-V. IC ₅₀ (μM) of different inhibitors for DfrB1 and DfrB4.....	xii
Table A3-I. Protein sequence identity and query coverage (in brackets) of DfrBs with the Expected value (in bold).....	xv
Table A3-II. IC ₅₀ of His ₆ -DfrB1 with TMP and inhibitors 4.1 to 4.5	xvi
Table A3-III. IC ₅₀ of DfrB1 (untagged) with TMP and inhibitors 4.1 , 4.3 , 4.6 and 4.7	xvi
Table A3-IV. Total activity, yield and purity of DfrB purifications.....	xvi
Table A3-V. Theoretical and measured masses of purified DfrBs.....	xvii
Table A3-VI. k_{cat}^{NADPH} for the dihydrofolate reductase activity of DfrBs.....	xvii
Table A3-VII. IC ₅₀ of DfrBs with TMP and inhibitors 4.1 , 4.3 , 4.6 and 4.7	xiii

Liste des figures

Figure 1-1. Voie de biosynthèse du folate.....	3
Figure 1-2. Structure du DHF, du NADPH et du TMP.	4
Figure 1-3. Représentation de la machinerie génétique complexe que sont les intégrons.....	7
Figure 1-4. Représentation de la structure secondaire, tertiaire et quaternaire de la DfrB1.....	11
Figure 1-5. Stratégies utilisées dans l'optimisation des fragments.....	15
Figure 1-6. Structure des inhibiteurs connus de la DfrB1.	17
Figure 1-7. Structure du rouge de Congo, un inhibiteur non sélectif de la DfrB1.	18
Figure 1-8. Représentations de la conformation en U du composé inhibiteur 1.8 lors de la liaison avec la DfrB1.....	19
Figure 1-9. Structures d'un bisubstrat, le composé 1.10	21
Figure 2-1. Scheme of the 9029 bp DNA segment derived from SRX560289.....	31
Table of content graphic (Chapitre 3).....	42
Figure 3-1. Surface representation of the tetrameric DfrB1 with its substrates NADPH and DHF, in sticks representation, bound inside the active-site tunnel (PDB code 2RK1).....	44
Figure 3-2. <i>In silico</i> reconstruction of 3.3 in the active site of DfrB1 based on the crystal structure PDB ID 'Authors will release the atomic coordinates and experimental data upon article publication.'.....	48
Figure 3-3. DfrB1 with bound NADPH and DHF substrates (PDB ID 2RK1; left panel) or with the central core of a pair of molecules 3.3 (PDB ID 'Authors will release the atomic coordinates and experimental data upon article publication.'; right panel) provide evidence that binding of DHF is incompatible with binding of one molecule of 3.3 (overlay, center panel).....	49
Figure 3-4. Inhibition of DfrB1 with monomeric inhibitor 3.38	61
Figure 3-5. Panel A: Inhibition of DfrB1 with 3.12 . Inset: Hill plot of panel A. The Hill coefficient is 1.1 ± 0.1 . Panel B: Dixon plot of 3.12 with DfrB1.....	64
Figure 3-6. A) Structure of 3.12 . B) <i>In silico</i> reconstruction of 3.12 in the active site of DfrB1 (left panel) and from a top view (right panel).....	65
Figure 4-1. The homotetrameric DfrB1 structure (PDB: 2RK1) and the DfrB family sequence alignment.....	101

Figure 5-1. Détermination du facteur Z' pour l'essai d'activité dihydrofolate réductase avec la DfrB1.....	119
Figure 5-2. Distribution normale de l'activité relative (%) de la DfrB1 mesurée avec les extraits d'invertébrés solubilisés en DMSO pour les plaques A et B.....	120
Figure 6-1. Structure du FAD; les parties importantes sont identifiées.....	126
Figure 6-2. Structures de l'inhibiteur de type bisubstrat 1.10/4.1 et de l'inhibiteur non-symétrique de type bisbenzimidazole 3.3	126
Figure 6-3. Structures des groupements fonctionnels alternatifs aux carboxylates (6.1).....	129
Figure 6-4. Structures de l'inhibiteur 3.12 et de la molécule 5.10	130
Figure A2-1. Crystal structures of 3.1 PDB ID 'Authors will release the atomic coordinates and experimental data upon article publication.' (left) and of 3.3 PDB ID 'Authors will release the atomic coordinates and experimental data upon article publication.' (right).....	xiii
Figure A2-2. Matrix of the binding residues of DfrB1 with 3.3	xiii
Figure A2-3. In silico reconstruction of 3.38 in the active site of DfrB1 based on the crystal structure PDB ID 'Authors will release the atomic coordinates and experimental data upon article publication.'.....	xiv
Figure A2-4. Dixon plot of 3.12 with DfrB1.....	xiv
Figure A3-1. 12% SDS-PAGE gel of purified DfrBs.....	xiii

Liste des schémas

Scheme 3-1. Trimethoprim (TMP) and known selective inhibitors of DfrB1	43
Scheme 3-2. Synthesis of 3.3 to 3.12	46
Scheme 3-3. Synthesis of 3.17 to 3.19	53
Scheme 3-4. Synthesis of 3.20	54
Scheme 3-5. Synthesis of brominated 3.21b and its derivatives.....	56
Scheme 3-6. Synthesis of methyl 2-(3-bromophenyl)benzimidazole-6-carboxylate 3.22 and its derivatives.....	57
Scheme 3-7. Synthesis of 3.38 to 3.39c	60
Scheme 3-8. Synthesis of non-symmetrical phenylbenzimidazole-6-carboxylate compounds.....	62
Scheme 3-9. Synthesis of non-symmetrical 2-phenylbenzimidazole compounds.....	62
Scheme 4-1. Depiction of the relative binding orientation of the substrate dihydrofolate (DHF) and the reducing cofactor NADPH within the DfrB1 active-site tunnel.....	100
Scheme 4-2. Synthesis of bisubstrate inhibitor 4.3	104
Scheme 4-3. Synthesis of bisubstrate inhibitor 4.4	104
Scheme 4-4. Synthesis of bisubstrate inhibitor 4.5	105
Scheme 4-5. Structures of 4.6 and 4.7	106

Liste des abréviations

7-be	Séquence canonique de sept paires de bases
r7-be	Complément inverse de 7-be
<i>attC</i>	Site de recombinaison des gènes
<i>attI</i>	Site de recombinaison d'intégrons
BLAST	Alignement basique local ou « <i>Basic Local Alignment Search Tool</i> »
CTL+	Contrôle positif
CTL-	Contrôle négatif
Dfr	Dihydrofolate reductase
DfrA	Dfr plasmidique résistante au TMP
DfrB	Dfr de type II
DHF	7,8-dihydrofolate
DHPS	Dihydroptéroate synthétase
<i>E. coli</i>	<i>Escherichia coli</i>
FAD	Flavine adénine dinucléotide
FDA	Agence américaine des produits alimentaires et médicamenteux ou « <i>Food and Drug Administration</i> »
FBLD	Conception d'inhibiteurs basée sur les fragments
GTP	Guanosine triphosphate
hDfr	Dfr humaine
HP	6-hydroxyméthyle-7,8-dihydroptérine
HPA	α -amylase pancréatique
HPPK	6-hydroxyméthyle-7,8-dihydroptérine pyrophosphokinase
<i>intI</i>	Gène encodant l'intégrase
IntI	Intégrase
k_{cat}	Constante de vitesse catalytique
K_d	Constante de dissociation
K_i	Constante d'inhibition
K_M	Constante de Michaelis-Menten

LowModeMD Courtes simulations de dynamique moléculaire

MIC	Concentration minimale inhibitrice ou « <i>Minimal Inhibitory Concentration</i> »
μM	Micromolaire
MRSA	<i>Staphylococcus aureus</i> résistant à la <i>Méthicilline</i> ou « <i>Methicillin-resistant Staphylococcus aureus</i> »
MTX	Méthotrexate
nM	Nanomolaire
NAPDH	Nicotinamide adénine dinucléotide phosphate
OMS	Organisation mondiale de la santé
<i>Pc</i>	Promoteur des intégrons
PDB	Banque de données sur les protéines ou « <i>Protein data bank</i> »
RMN	Résonance magnétique nucléaire
RMSD	Distance moyenne entre les atomes ou « <i>root-mean-square deviation</i> »
$\sigma_{\text{CTL}+}$	Écart type contrôle positif
SAR	Relation structure-activité
THF	Tétrahydrofolate
TMP	Triméthoprim
VS	Criblage <i>in silico</i> ou « <i>virtual screening</i> »
X_s	Moyenne

À mon frère, Simon Toulouse

Remerciements

Cette thèse reflète toutes mes années au doctorat qui m'ont permis de me développer autant sur le plan personnel que scientifique. Je n'aurais jamais pu accomplir autant de choses sans toutes les personnes qui m'ont aidé directement ou indirectement à mon doctorat.

Premièrement, je voudrais remercier la Professeure Joelle Pelletier de m'avoir acceptée dans son laboratoire pour la maîtrise, puis pour le doctorat. J'ai commencé en fille inexpérimentée qui doutait de sa valeur scientifique, qui ne pouvait pas faire de présentation devant plus de deux personnes et qui venait de remporter le prix de la plus gaffeuse du baccalauréat de biochimie et médecine moléculaire. Grâce à elle, j'ai pu participer à plusieurs conférences et gagner plusieurs prix ainsi que plusieurs bourses prestigieuses. Aussi grâce à elle, je peux maintenant entrer dans une salle bondée d'étudiants pour donner un cours avec cinq minutes de préavis tout en découvrant le cours en même temps que les étudiants et ce, en français comme en anglais. Pour ce qui est de ma maladresse... Nous nous dirons que certaines choses ne changeront jamais et que : « Plus ça change, plus c'est pareil! » ;). Merci d'avoir cru en moi!

Il ne faut jamais oublier l'aide apporté par le personnel de l'université : les gars de l'atelier, le laboratoire de spectroscopie de masse et le secrétariat de chimie. À chaque bris d'instruments ou au moindre pépin, les gars de l'atelier venaient toujours à notre rescousse. Ils peuvent construire, à partir de rien, d'impressionnantes pièces afin d'améliorer nos instruments. Ils me font penser à des pères typiques dans un garage qui peuvent tout réparer avec une pièce trouvée dans le fond d'un tiroir. Ils font partie d'une des meilleures équipes d'employés que j'ai vues. Les filles du laboratoire de spectroscopie de masse font aussi un très bon travail. Elles sont toujours prêtes à prendre de leur temps pour répondre à nos questions après l'analyse nos échantillons. Elles possèdent une patience inouïe pour les non-chimistes comme moi. Étant en charge des achats du laboratoire du professeure Joelle Pelletier, j'ai, malheureusement, vécu l'entrée en fonction du logiciel Synchro finance à l'université. J'ai finalement réussi à apprivoiser ce logiciel grâce à l'aide du personnel du secrétariat de chimie. Parce que c'est évident pour tout le monde de comprendre qu'un numéro de projet est un numéro de compte et qu'un numéro de compte GL (grand livre) est un code à l'interne!

Au courant de mon doctorat, Professeure Joelle Pelletier a construit une équipe dans son laboratoire qui m'a permis de faire de belles rencontres. Merci à Dominic Bastien qui m'a

formée et inclus dans le laboratoire. Merci à mes stagiaires, Nadia Nouasria et William Sandé, qui m'ont laissé les encadrer; j'ai eu beaucoup de chance d'avoir de bons stagiaires comme vous! Merci à ma *sista* Sarah Mélissa Abraham, qui a accepté mes conseils et qui m'a donné autant de fous rires. Merci à Brahm Yachnin qui a pris de son temps et eu beaucoup de sa patience pour travailler avec moi. Merci aux membres *alumni* du laboratoire qui ont affecté mon parcours scientifique: Damien Colin, Natalie Rachel, David Charbonneau, Simon Durr et Daniela Quaglia. Je suis désolée si j'ai oublié quelqu'un!

Plusieurs collègues sont devenus des amis pendant mon doctorat, rendant l'ambiance du laboratoire plus agréable. Merci à Sophie Gobeil de m'avoir endurée, je ne sais pas comment elle a fait pour ne pas se fâcher. J'ai été comme un enfant, j'ai cherché la limite de sa patience et je n'ai pas encore trouvée... Elle a aussi pris beaucoup de temps à me lire et à me conseiller dans tous mes travaux scientifiques. Merci à Lorea Carlotta Alejaldre avec qui, j'ai passé de bons moments à la zumba, dans les bars et au laboratoire la fin de semaine. J'apprécie aussi toutes les fois où tu m'as consolé. Merci à Sara OuadiT qui a lu la plupart de mes chapitres de thèse dont, mes remerciements. Je n'aurais jamais cru trouver une meilleure amie à ce moment de ma vie! Elle est LA meilleure stagiaire que j'ai côtoyée et avoir su qu'elle était aussi bonne, je l'aurais prise sur mon projet! Depuis que je la connais, j'ai passé que des moments LIT! Merci de m'encourager quotidiennement avec tes mots chaleureux comme « Courage! ».

Je ne peux pas passer à côté de deux hommes formidables, Oliver Rousseau et Maximilian Ebert, avec qui, j'ai passé de belles soirées arrosées à jouer aux jeux de société/vidéos (LA BOMBA !!!) ainsi qu'un merveilleux *road trip* en Ontario. Nous saurons toujours nous encourager avec notre sublime devise : « T'es belle, t'es bonne, t'es fine et t'es capable! ». Ma relation avec Olivier Rousseau a porté plusieurs noms au courant de ces dernières années, malgré cela il a toujours su être là pour me calmer et remettre les choses en perspectives. Il me fascine avec sa mémoire de poisson rouge et sa grande écoute, c'est un bon contraste avec moi la bavarde avec une bonne mémoire. Il est une des personnes qui me fait le plus rire, par ses explosions de tubes eppendorfs, ses blagues sur les *vegans*, ses blagues sur ma prononciation des mots tel que « vortexter » et plusieurs autres choses qui malheureusement, ne s'écrivent pas ici... Merci d'avoir pris beaucoup de temps et eu beaucoup de patience envers moi, la tannante! Un grand merci à Maximilian Ebert, mon meilleur ami, à qui je pourrais écrire 100 pages de remerciements. De par ses idées, il a contribué directement à mon projet, ce projet qui était si

différent des autres du laboratoire. Ensemble, nous avons testé la patience de plusieurs personnes à l'université ainsi que dans la rue. Lorsque nous nous retrouvons ensemble, notre énergie devient exponentielle en devenant de vrais catalyseurs (LOL!) l'un pour l'autre! Je vais toujours me souvenir de nos jeudis soir à regarder des films, de nos brunchs et des fins de semaine passés ensemble au chalet. J'ai aimé le corriger sur son français québécois et lui faire perdre ses termes qu'il trainait de la France. Il est aussi un contributeur direct à l'amélioration de mon estime personnelle par ses phrases aussi diplomates que : « Si tu ouvrais enfin les yeux et que tu te rendais compte comment les gens te perçoivent, tu n'aurais pas une aussi mauvaise estime de toi! ». Merci d'avoir fait rentrer Inga et Julie dans ma vie! Cette merveilleuse Inga qui a toujours laissé sa porte ouverte pour moi malgré, toutes les niaiseries et les soirées bruyantes auxquelles j'ai participées. Elles méritent une médaille! Merci à Julie d'être l'enfant le plus *cool!*

Il existe aussi plusieurs personnes qui ont participé indirectement à mon doctorat. Des personnes qui ont su me faire décrocher du laboratoire. Des fois, il faut s'éloigner pour mieux revenir! Il y a ma famille : mon père (Bertrand), ma mère (Lyne) et mes deux frères (Benoit et Simon). Nos interactions ont fait de moi la personne que je suis aujourd'hui, une femme déterminée et résiliente. Mes parents m'ont appris à être indépendante et à travailler fort pour avoir ce que l'on veut dans la vie. J'espère qu'un jour ils pourront comprendre l'importance et l'utilité de l'université. Un merci spécial à mon frère, mon meilleur ami, Simon qui a toujours été là pour moi. Ce n'est pas tout le monde qui est capable de vivre et de partager une voiture avec sa sœur, *tsé* c'est long longtemps quelques décennies. D'avoir inversé les rôles et de m'avoir aidée quand j'en avais besoin lorsque normalement c'est le contraire, merci. À toutes les images qu'il a faites pour mes articles, merci. Malheureusement, je ne peux pas tout énumérer.

Il y a aussi mon cercle d'amis proches: Oli (Olivier Léger-Leduc), Stéph (Stéphanie Tremblay), Fofoy (Geoffroy Desmarais), Caro (Caroline Malo-Pion), Vickychou (Victor Lambin-Iezzi), Jude (Judith Roberge), Simbine (Simon Destrembles) et MJ (Marie-Jeanne Paré). On en a fait des choses ensemble : des soirées arrosées, des fins de semaine au chalet, un voyage à Cuba, des sorties à la plage... Merci à Oli pour toutes les conversations que nous avons eues sur l'équilibre nos vies, aux petites heures du matin. Merci à Stéph pour m'avoir engraisé avec ses mets plus que bons! Merci à Fofoy de m'avoir fait découvrir la technologie comme la

réalité virtuelle, j'ai pu voir MJ battre des zombies! Merci à Dre. Caro qui m'a montré qu'on pouvait aller loin dans la vie et ce, malgré le milieu d'où on vient! Merci à Vickyhou de me motiver dans la fin de mon doctorat, j'ai bien aimé chialer sur les réparations de la rampe avec lui! Merci à Jude pour beaucoup de choses depuis les 20 dernières années, dont son support et son côté tannante, elle rend tout moins plate! Merci à Simbine de prendre soin de moi ainsi que de mon confort, et de partager MJ avec moi! Et la dernière et non la moindre, ma sœur MJ! Merci de m'avoir accueillie dans ta maison lorsque j'avais nulle part où aller, d'avoir été présente et de m'avoir supporté dans les moments difficiles de ma vie. Les dernières années ont été difficiles pour toi et moi, mais je prévois un avenir qui nous réserve de belles choses! Deux autistes avec un syndrome d'Asperger ensemble vont peut-être réussir à comprendre les gens et leur subtilité. Merci à Michael Gagnon d'être dans ma vie, je ne peux pas imaginer ma vie sans lui. Tous les moments passés avec lui sont plus qu'agréables. Malgré son horaire chargé de Monsieur le docteur résident sénior en médecine interne à McGill, il prend toujours le peu de temps qu'il a pour lire mes textes en anglais! Merci à Laura Mamani qui a été une super coloc et qui a souvent pris le thé avec moi pour me changer les idées!

Chapitre 1 Introduction

1.1 . Résistance aux antibiotiques

1.1.1 . L'importance des antibiotiques

Les antibiotiques sont considérés comme une des plus grandes avancées médicales modernes(1-3). Cette révolution médicale a marqué le 20^{ième} siècle de par la guérison d'infections bactériennes qui auparavant étaient mortelles(2). Avant l'usage de la célèbre pénicilline en 1941, le taux de mortalité par une infection sanguine au *Staphylococcus aureus* était estimé à 80% et la cause première de mortalité aux États-Unis était la pneumonie(2). Les agents antibiotiques sulfamides ont été les acteurs principaux pour prévenir les infections pendant la Deuxième Guerre mondiale(4). Ces antibiotiques ont traité les infections aux *Neisseria meningitidis* et *Pneumocystis carinii* ainsi que la lèpre causée par *Mycobacterium leprae*(5).

À la fin des années 30, soit seulement quelques années après leur introduction, la résistance aux sulfamides a été reportée(1). Dans le cas de la *Neisseria meningitidis*, une souche résistante aux sulfamides a été découverte après seulement deux années d'utilisation(5). Afin de palier à la résistance aux sulfamides, ceux-ci étaient prescrits en combinaison avec un autre antibiotique : le triméthoprim (5-(3,4,5-triméthoxybenzyl)pyrimidine-2,4-diamine; TMP) (Figure 1-2)(6, 7). Le TMP est un agent bactériostatique(8, 9). Ainsi, c'est à partir des années 60, que le TMP et le sulfaméthoxazole, un sulfamide, furent utilisés conjointement en clinique afin de combattre les infections bactériennes(7).

En milieu clinique, la combinaison de ces deux antibiotiques était utilisée pour traiter la pneumonie, la gonorrhée ainsi que les infections urinaires, respiratoires et gastro-intestinales(10). Malgré plusieurs décennies d'utilisation de cette combinaison, en plus de la résistance bactérienne, des effets secondaires indésirables importants du sulfaméthoxazole ont été découverts tels que la porphyrie, les troubles de l'appareil urinaire, les troubles hématopoïétiques et les réactions d'hypersensibilité(4, 7). Par ailleurs, des études concernant les infections urinaires et respiratoires suggèrent que le TMP est plus efficace en utilisation seule qu'en combinaison avec le sulfaméthoxazole(7). L'Agence de la santé publique du Canada

classe les antibiotiques en catégories de I à IV dont I est de très haute importance(11). Le TMP se situe dans la catégorie II avec une haute importance(11). Cette catégorie II comprend les antibiotiques à large spectre pour lesquelles des alternatives sont encore disponibles(11). L'organisation mondiale de la santé (OMS) classe le TMP parmi les antibiotiques importants dont sa résistance est considérée d'importance critique élevée(12).

1.1.2 . La cible du triméthoprime

Malgré leur introduction comme premiers antibiotiques en 1932, ce n'est qu'en 1940 que certaines observations ont pu démontrer que les sulfamides ciblent la dihydroptéroate synthétase (DHPS)(4, 13). *Hitchings et al.* ont démontré que la plupart des 2,4-diaminopyrimidines, tel que le composé synthétique TMP, inhibent compétitivement la voie de biosynthèse du folate(10, 14). La dihydrofolate réductase (Dfr) a été identifiée comme la cible de ces inhibiteurs, une cible différente des sulfamides. Plus précisément, le TMP inhibe la Dfr chromosomique bactérienne(7).

La voie de biosynthèse du folate, constituée de plusieurs enzymes, est retrouvée chez tous les organismes vivants (Figure 1-1). Plus précisément, cette voie est essentielle puisqu'elle fournit une unité de carbone à l'aide d'un cofacteur, le tétrahydrofolate (THF), permettant la biosynthèse de certains acides aminés (glycine, méthionine et histidine) et des purines (5, 13). L'inhibition de cette voie métabolique épuise le bassin de purines, arrêtant ainsi le processus de synthèse de l'ADN. La croissance et la prolifération cellulaire s'arrêtent, causant la mort de la cellule(5, 13).

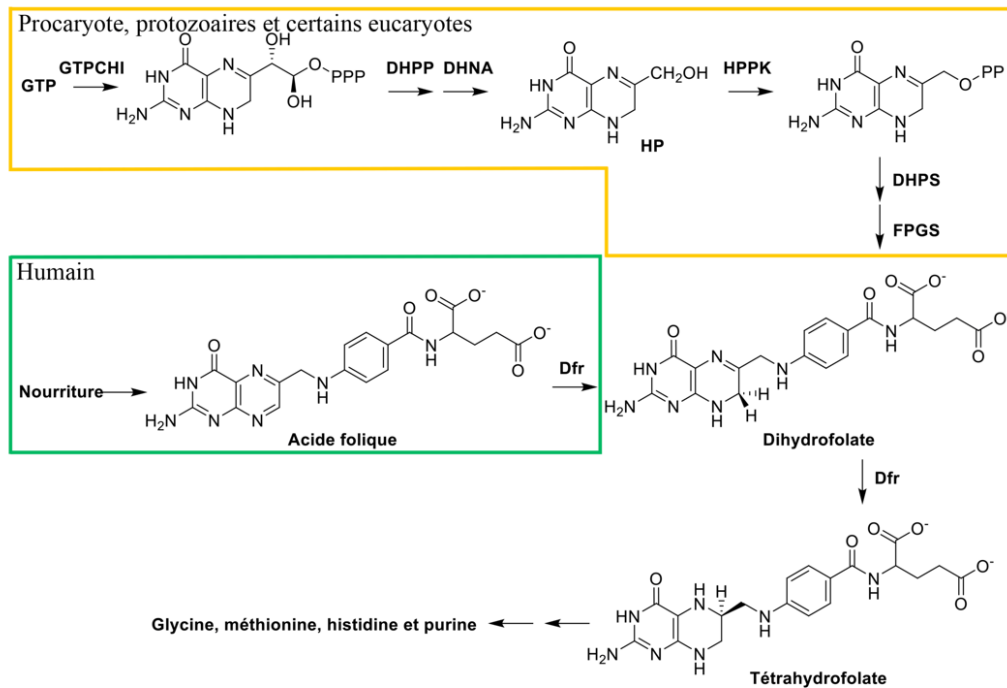


Figure 1-1. Voie de biosynthèse du folate.

L'encadré de couleur jaune représente celle retrouvée chez les procaryotes, les protozoaires et certains eucaryotes alors que l'encadré de couleur verte est pour les humains. Les étapes sans encadré sont utilisées par tous les organismes. GTP : guanosine triphosphate; GTPCHI : GTP cyclohydrolase I; DHPP : 6-hydroxyméthyle-7,8-dihydroptérine pyrophosphate; DHNA : 7,8-dihydroneoptérine; HP : 6-hydroxyméthyle-7,8-dihydroptérine; HPPK : 6-hydroxyméthyle-7,8-dihydroptérine pyrophosphokinase; DHPS : dihydroptéroate synthétase; FPGS : Folylpoly- γ -glutamate; DHF : dihydrofolate; Dfr : dihydrofolate réductase; THF : tétrahydrofolate (Figure inspirée de Wright et al. (2006-2010) et Bermingham et al. (2002))(5, 13)

La voie de biosynthèse du folate chez les procaryotes, les protozoaires et certains eucaryotes débute par la synthèse de l'acide folique(13). En effet, la plupart des micro-organismes sont autotrophes pour l'acide folique : sa synthèse est possible à partir de la guanosine triphosphate (GTP) par plusieurs réactions enzymatiques (Figure 1-1; encadré jaune)(5, 13). Cependant, ces étapes enzymatiques, dont la DHPS, sont absentes chez l'humain puisque la voie de biosynthèse du folate commence par la réduction de l'acide folique par la Dfr (hDfr) (Figure 1-1; encadré vert). Ainsi, l'alimentation permet de pallier cette lacune en fournissant l'acide folique(5, 13).

La Dfr est impliquée dans une des étapes subséquentes à celle de la DHPS (Figure 1-1)(5, 14). En fait, l'enzyme Dfr catalyse la réduction du substrat 7,8-dihydrofolate (DHF) en THF à l'aide du cofacteur nicotinamide adénine dinucléotide phosphate (NAPDH) (Figure 1-2). C'est à cause de son exclusivité et de son importance que cette voie fut une cible de choix pour le développement d'antibiotiques.

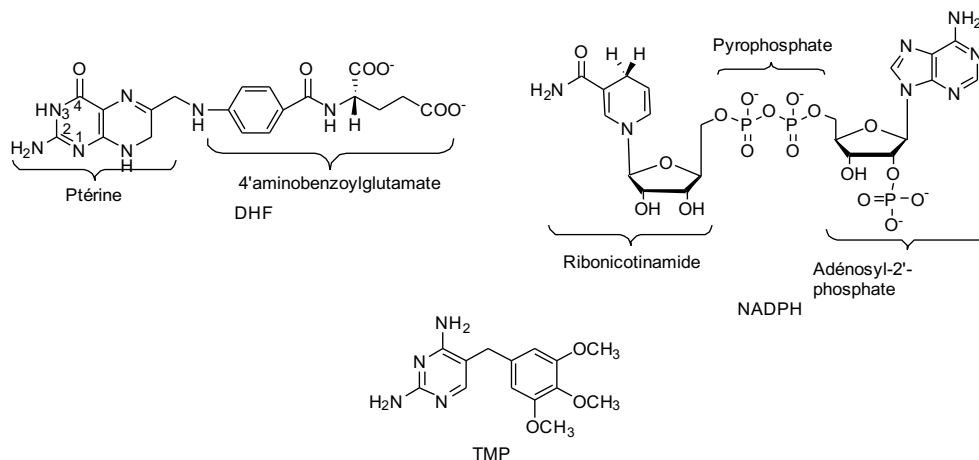


Figure 1-2. Structure du DHF, du NADPH et du TMP.

En haut, le DHF numéroté est le substrat et le NADPH est le cofacteur des Dfr. En bas, le TMP est un inhibiteur de la Dfr chromosomique bactérienne.

1.1.3 . Le triméthoprime et son utilisation considérable

Les bactéries sont impliquées dans les maladies à prolifération cellulaire rapide tout comme les cellules cancéreuses. Dans ces maladies à prolifération cellulaire rapide, les cellules nécessitent une quantité élevée de purines pour la synthèse de l'ADN et par conséquent, d'un des précurseurs des purines, le THF. En inhibant la Dfr, la quantité de THF diminue ainsi que celle des purines et donc, la synthèse d'ADN diminue. Ceci empêche la propagation de l'infection bactérienne. Comme mentionné précédemment, le TMP inhibe fortement la Dfr chromosomique des bactéries à Gram négatifs avec une constante d'inhibition (K_i) de l'ordre de 20 pM et aussi, celle de quelques bactéries à Grams positifs comme *Staphylococcus aureus*, *Streptococcus pyogenes* et *Streptococcus pneumoniae* (15, 16). De plus, le TMP affecte très peu les cellules animales. Pour la hDfr, le TMP possède une constante d'inhibition de 260 000 fois plus élevée que celle des bactéries à Gram négatifs(17).

De par son activité sur un large spectre de bactéries et de par sa faible toxicité chez l'humain, le TMP est un antibiotique de choix pour vaincre les infections bactériennes(15). De plus, son faible coût de production fait en sorte qu'il est couramment utilisé en prévention en milieu d'élevage et dans l'aquaculture des crevettes et des saumons(18, 19). Cet antibiotique est introduit entre autres dans leur nourriture afin d'éviter la diminution du rendement de l'élevage. Cette mesure de prévention résulte en une pression sélective auprès des bactéries résistantes. D'ailleurs, cette résistance peut être transférable d'un organisme à l'autre comme observé fréquemment, d'*Aeromonas salmonicida* à *Escherichia coli* (*E. coli*) et à *Salmonella*, organismes pathogènes pour l'humain(19). Malgré la réglementation dans la plupart des pays industrialisés, certains pays comme la Chine et le Chili continuent à utiliser de façon massive les antibiotiques en prévention dans l'aquaculture(19). Il est estimé qu'aux États Unis, 80% des antibiotiques utilisés en prévention en milieu d'élevage alors que 20% chez les humains(20).

Une forte corrélation a été démontrée pour les souches d'*E. coli* résistantes à plusieurs antibiotiques provenant d'échantillons sanguins humains et celles provenant des animaux d'élevage (21). Cependant, l'utilisation massive du TMP mondialement crée une résistance bactérienne à celui-ci, produisant un problème clinique important(10). En 1966, soit quelques années après le début de son utilisation médicale, une résistance au TMP est retrouvée dans 1% des infections urinaires à *E. coli* alors que cette résistance s'élève à 33% pour les infections à *Klebsiella aerogenes*(22). Selon un rapport de l'OMS effectuée sur 22 pays, les infections urinaires par *E. coli* résistante se situent entre 8 à 65%(23).

1.1.4 . La résistance au triméthoprime

Plusieurs mécanismes de résistance sont en cause tels qu'une imperméabilité accrue de la membrane envers l'antibiotique, une voie métabolique alternative, une surproduction de la cible ou encore un plasmide codant pour une enzyme résistante ou désactivant l'antibiotique(4). Cette dernière résistance au TMP peut provenir d'une variante de la Dfr chromosomique nommée DfrA, ayant acquis des mutations sans affecter son activité enzymatique, mais tout en diminuant son affinité pour le TMP(6).

La résistance au TMP peut découler d'une enzyme Dfr alternative encodée par un plasmide. À la fin des années 80, un peu moins d'une dizaine de plasmides codant pour les enzymes Dfrs de type B (DfrB) qui sont résistantes au TMP étaient répertoriés(24, 25). Ces

plasmides, possédant souvent plusieurs résistances aux antibiotiques, sont transférables par conjugaison, augmentant leur transmission(10). Jusqu'à ce jour aucun antibiotique visant les DfrBs n'existe.

1.1.5 . Les intégrons; une machinerie génétique complexe

Les plasmides multi-résistants contiennent plusieurs résistances aux antibiotiques. Ceux-ci possèdent une machinerie génétique complexe nommée intégrons, permettant la capture, l'incorporation et la dissémination de gènes de résistance, soit un clonage de gène naturel(26, 27). Les intégrons sont retrouvés en abondance en clinique dans les pathogènes(28). Les intégrons se composent principalement de trois éléments : le promoteur (*Pc*), le gène encodant l'intégrase (*intI*) et le site de recombinaison d'intégrons (*attI*)(Figure 1-3)(26, 29). *Pc* permet la transcription des cassettes de gènes mobiles(29). Les cassettes de gènes mobiles comprennent le site de recombinaison (*attC*) ainsi que la séquence codante de la protéine de résistance(26, 29) (Figure 1-3). L'intégrase (IntI) catalyse l'insertion et l'excision de cassettes de gènes mobiles(29). Cette action se produit soit par recombinaison du site d'insertion *attI* avec la séquence d'*attC*, soit par recombinaison des séquences *attC* × *attC*(26, 28, 29). La séquence d'*attC* est retrouvée en aval de la séquence codante et elle varie de 57 à 141 paires de bases(27). La séquence d'*attC* inclut, à son extrémité 3', une séquence canonique de sept paires de bases 5'-GTTRRRY (7-be) ainsi que le complément inverse de ce 7-be en amont de la séquence d'*attC* (r7-be). Ce court 7-be est connu pour être fortement conservé, contrairement au reste de l'*attC* qui est très variable. Cette machinerie génétique complexe que sont les intégrons présente deux avantages pour les bactéries, les rendant dangereuses(29). Tout d'abord, l'intégration de la nouvelle cassette de gènes au site spécifique *attI* dans le génome bactérien se produit sans perturbation des gènes déjà existants(29). De plus, l'intégration se fait près du *Pc* et elle permet la transcription rapide du nouveau gène conférant des phénotypes avantageux, suite à une pression sélective(29).

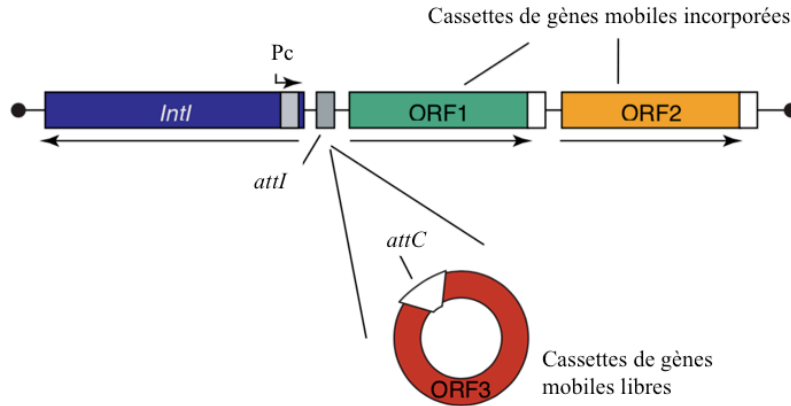


Figure 1-3. Représentation de la machinerie génétique complexe que sont les intégrons.

Les acteurs principaux sont évoqués dans la capture et l'incorporation suivie de la dissémination de gènes de résistance. (Figure adapté de Alonso et al. (2006))(26)

Dans le cas des *dfrBs*, ceux-ci sont uniquement retrouvés dans les cassettes de gènes mobiles(10, 26). La différence majeure entre les cassettes de gènes mobiles des *dfrBs* réside dans les séquences entre 7-be et la séquence codante. Bien que les régions en aval diffèrent, elles sont plus conservées que les régions en amont.

1.2 . La famille protéique des dihydrofolates réductases de type B (DfrB)

1.2.1 . Les dihydrofolates réductases de type B : les moutons noirs des dihydrofolates réductases

Peu de choses sont connues sur la famille des DfrBs outre le fait que les gènes encodant ces enzymes sont retrouvés dans des organismes résistants au TMP. Pour seulement un membre de la famille, la DfrB2, la K_i a été déterminée avec une valeur de 0,15 mM indiquant une faible sensibilité au TMP(16). Les DfrBs n'ont rien en commun avec les enzymes Dfr animale et chromosomique bactérienne outre leur capacité à catalyser la même réaction(26). Ces enzymes Dfr animale et chromosomique bactérienne monomériques se retrouvent chez les eucaryotes ainsi que chez les procaryotes avec une homologie de structure élevée (RMSD = 1,41Å) malgré leur faible identité de séquence protéique de $12,6 \pm 1,8\%$ (16, 17, 30).

Plusieurs types de Dfrs seront présentés au courant de cet ouvrage. La nomenclature de Dfrs peut porter à confusion; ainsi, une précision s'impose (Tableau 1-I). Les Dfrs incluent la hDfr, les Dfrs chromosomiques bactériennes, les DfrAs ainsi que les DfrBs, et elles sont indispensables à la cellule. Les DfrAs, qui ont évolué à partir des Dfrs chromosomiques bactériennes et leur sont fortement homologues, sont plasmidiques et confèrent une résistance au TMP(26). Malgré leur dissemblance avec les DfrAs, les DfrBs sont également plasmidiques et peu sensibles au TMP(26).

Tableau 1-I. Nomenclature des gènes encodant les Dfrs.

Gènes	Protéines	Caractéristiques
<i>dfr</i>	Dfr	<ul style="list-style-type: none"> • Essentielles dans toutes les cellules • Comprennent la Dfr humaine (hDfr), les Dfrs chromosomiques bactériennes (DfrAs) et les DfrBs
<i>dfrA</i>	DfrA	<ul style="list-style-type: none"> • Évolutivement similaires aux Dfrs chromosomiques bactériennes • Résistantes au TMP • Retrouvées sur des plasmides
<i>dfrB</i>	DfrB	<ul style="list-style-type: none"> • Évolutivement distinctes des Dfrs chromosomiques • Résistantes au TMP • Retrouvées sur des plasmides

Les Dfr chromosomiques de *Staphylococcus aureus* et d'*E. coli* détiennent respectivement, une sensibilité au TMP de l'ordre de 10^6 et 10^7 fois supérieure à la DfrB1, qui est la plus étudiée des DfrBs. Cette dernière est donc caractérisée d'enzyme résistante au TMP(10, 18, 31-35). Les neuf enzymes DfrBs (DfrB1 à DfrB9) connues sont hautement similaires (identité de séquence protéique entre 77% et 100%). Leur différence réside principalement aux 21 premiers acides aminés de l'extrémité N-terminale où il a été démontré, dans le cas de la DfrB1, que ces acides aminés ne participent pas à la structure secondaire, tertiaire ou quaternaire de l'enzyme et ce, même en présence d'un ligand(25, 36). Ces N-termini diffèrent dans la composition d'acides aminés et non dans le nombre de ceux-ci, contrairement aux C-termini qui diffèrent en longueur et en composition(36). Ceci permet de spéculer que la longueur serait un facteur plus important que la composition de ces N-termini. Les N-termini de

la DfrB1 qui est un homotétramère, interagissent ensemble afin d'augmenter la stabilité de celle-ci(36). En résumé, ceux-ci seraient important dans l'expression, la stabilité, la solubilité, l'adaptabilité évolutive ainsi que l'accessibilité des substrats(25, 36, 37).

La similitude élevée entre les séquences d'ADN des *dfrBs* suggère fortement qu'elles possèdent la même structure, soit le motif SH3(26). À ce jour, seulement la structure de la DfrB1 (R67 DHFR) a été résolue par cristallographie aux rayons X (Figure 1-4)(34, 35). Le motif SH3 est principalement retrouvé dans les protéines eucaryotes comme un domaine non catalytique permettant les interactions protéine-protéine(26, 38). Le motif SH3 est peu retrouvé chez les procaryotes et les DfrBs sont les premières protéines à utiliser ce domaine comme unité catalytique(26). Ces observations permettent d'émettre l'hypothèse que les DfrBs seraient de nouvelles enzymes émergentes(26).

La DfrB1 est peu efficace comparée aux autres Dfrs(18, 32-35). Les Dfrs chromosomiques d'*E. coli* relativement à la DfrB1, ne possèdent pas la même étape limitante dans la catalyse de la réaction. La faible efficacité catalytique s'explique par le fait que la réaction chez la DfrB1 a lieu lorsque la forme protonée du DHF se lie au site actif(39). Cette protonation est possible grâce au milieu réactionnel et dépend donc du pH(39). Elle est nécessaire puisque tout semble indiquer l'absence d'un donneur de proton dans le site actif de la DfrB1, contrairement aux Dfr chromosomiques(39). De plus, la DfrB1 possède une constante de Michaelis-Menten (K_M) pour le DHF de 8,2 μM et pour le NADPH de 1,6 μM alors que celles de la Dfr chromosomique d'*E. coli* sont de 1,2 μM et de 0,94 μM respectivement(17, 40). Donc, la DfrB1 possède une affinité environ six fois plus faible pour le DHF (K_M), une étape limitante différente ainsi qu'une vitesse catalytique (k_{cat}) de cette étape limitante plus faible que la Dfr chromosomique d'*E. coli*. Malgré ceci, la DfrB1 permet la survie des bactéries et ainsi, crée une résistance contre le TMP. Il est envisageable que l'enzyme DfrB1 agisse comme une enzyme primitive ne faisant que maintenir les substrats proches sans participer directement à la réaction(41).

1.2.2 . La structure secondaire, tertiaire et quaternaire de la DfrB1

La DfrB1 est une enzyme thermostable homotétramérique, ce qui la différencie des Dfr chromosomique d'*E. coli* et hDfr qui sont des monomères(Figure 1-4)(42). La DfrB1 possède une symétrie 222 (Figure 1-4)(43). Chaque monomère de 78 acides aminés se compose d'un

motif de type SH3 qui est constitué d'un baril à 5 brins β antiparallèles (Figure 1-4A)(16, 43). L'association entre deux monomères se produit par la contribution de trois brins β d'un monomère avec deux brins β de l'autre monomère ce qui forme un dimère(16). Ensuite, deux dimères s'unissent par le contact des histidines 62 (H62) situées sur les boucles non-structurées afin de créer un tétramère (Figure 1-4A)(16, 40). Le site actif, juste assez grand pour accueillir les substrats, se compose d'un orifice hydrophobe en forme de sablier traversant la protéine au centre(44). Chaque monomère contribue de manière égale à la formation du site actif ce qui résulte en un seul site actif symétrique central par homotétramère, ce qui distingue cette enzyme des autres enzymes multimériques qui comptent généralement le même nombre de sites actifs que de monomères (Figure 1-4A)(40, 41). Par conséquent, plusieurs complexes peuvent être formés soit $E \cdot (NADPH)_2$, $E \cdot (DHF)_2$ ou $E \cdot NADPH \cdot DHF$, ce qui indique une reconnaissance peu spécifique des substrats et qui sera expliquée ultérieurement(40, 41). La valeur de la constante de dissociation (K_d) du NADPH est 50 fois plus faible que celle du DHF envers la DfrB1 ce qui favorise la formation du complexe $E \cdot NADPH$ (40, 45). Par la suite, une coopérativité positive est observée, favorisant la formation du complexe productif $E \cdot NADPH \cdot DHF$, au détriment du complexe non-productif $E \cdot (NADPH)_2$ (45). Ce concept sera important lors de la conception d'inhibiteurs afin d'augmenter leur affinité.

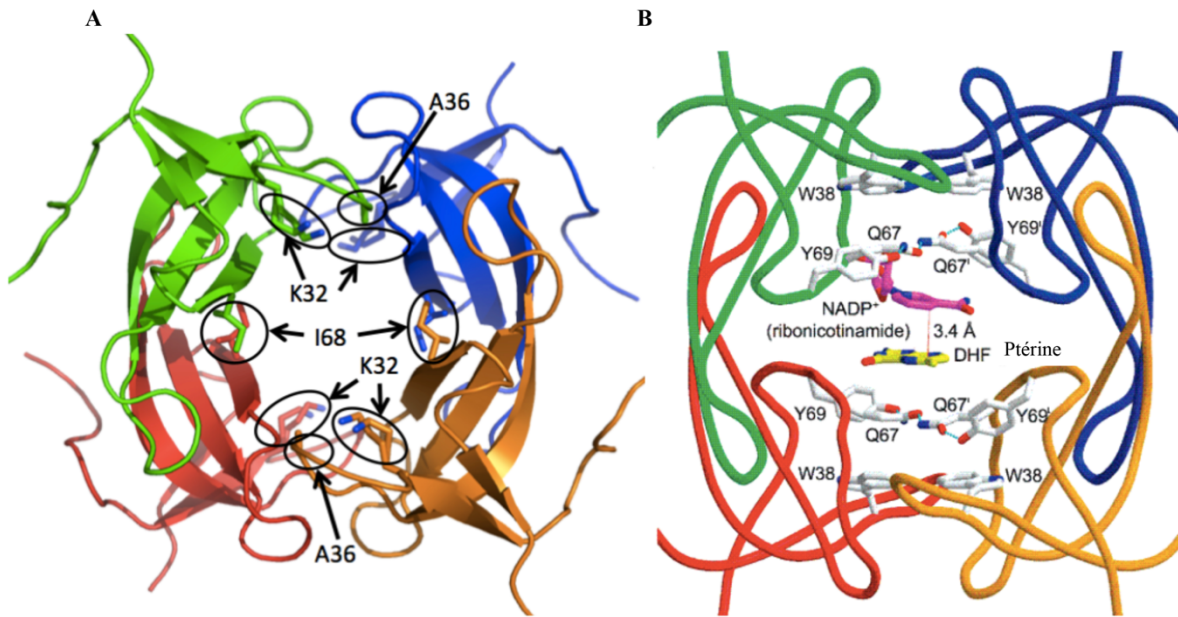


Figure 1-4. Représentation de la structure secondaire, tertiaire et quaternaire de la DfrB1.

Chaque sous-unité est représentée par une couleur différente. A) La structure secondaire et tertiaire de la protéine se compose de 5 brins β antiparallèles formant un baril, caractéristique du motif SH3. Les acides aminés décrits dans le texte sont indiqués : A36, I68 et K32. La protéine possède quatre A36, mais seulement deux sont visibles puisque les deux autres se retrouvent sur l'autre interface de la protéine. Les K32 sont positionnées à l'extérieur du site actif alors que les I68 sont à l'intérieur du site actif. B) Deux paires de résidus Y69-Q67-Q67'-Y69' forment une pince autour du substrat et du cofacteur. L'emplacement du substrat, DHF en jaune, et du cofacteur NADP+ en rose, est indiqué. (Figure adaptée de PDB 2RK1 et de Park et al. (1997))(39)

1.2.3 . Le site actif de la DfrB1

Pour la conception d'inhibiteurs compétitifs et sélectifs, la caractérisation du site actif de la cible enzymatique est importante. Le site actif de la DfrB1 est dominé par deux paires de résidus Y69-Q67-Q67'-Y69' provenant de chaque monomère (Figure 1-4B)(39). Afin de promouvoir la réaction, les deux paires de résidus Y69-Q67-Q67'-Y69' étreignent le substrat avec le cofacteur formant un empilement de plusieurs cycles (Figure 1-4B)(43, 45). Cet empilement inclut l'indole de deux tryptophanes (W38) et le phénol de deux tyrosines (Y69) d'une part et d'autre des ligands (Figure 1-4B)(43, 45). Ces Y69 sont associées par des liens hydrogènes à une paire de glutamines (Q67) centrées ces dernières(43). Ainsi, dans ce réseau

d'interactions, une tyrosine joue le rôle de donneur de pont hydrogène alors que l'autre tyrosine agit comme accepteur de pont hydrogène. Cependant, en absence de substrats, la Q67 se déplace formant un lien hydrogène avec une isoleucine (I68) (Figure 1-4A)(43). Ce changement de conformation induit une liaison entre la Y69 et une molécule d'eau(43). Aussi, le déplacement de la chaîne latérale de Q67 décondense l'empilement autour du W38 induisant une inclinaison vers le centre du site actif(43). La structure cristalline du NADP⁺ et du DHF avec la DfrB1 (PDB 2RK1) indique que la partie nicotinamide du NADP⁺ se trouve dans le cœur du site actif alors que son extrémité adénosyl-2'phosphate se trouve à l'extérieur du site actif(43, 45). Dans cette même structure cristalline, la partie ptérine du DHF se lie au cœur du site actif(45). Le transfert d'hydrure est alors catalysé par la superposition du nicotinamide avec la ptérine préprotonnée par le milieu. La reconnaissance du substrat et du cofacteur se réalise grâce à une subtile analogie structurale. En effet, deux lysines (K32) de deux sous-unités adjacentes s'approchent de l'extrémité glutamate du DHF alors que les deux autres K32 interagissent avec le pyrophosphate du NADP⁺ (Figure 1-4A)(43). Pour le DHF, la chaîne principale d'une I68 de l'enzyme reconnaît les atomes N3 et O4 de la ptérine (Figure 1-2) de la même façon qu'avec le groupement amide du NADPH en formant des liaisons hydrogènes (Figure 1-4A)(43). Pour le NADPH, l'oxygène lié à un des groupements phosphate établit une interaction électrostatique avec l'azote du squelette de l'alanine 36 (A36) (Figure 1-4A)(43). Ceci fixe solidement ces molécules pour la réaction de transfert stéréospécifique de l'hydrure(43). Cette reconnaissance du substrat DHF par l'enzyme est importante pour le mécanisme catalytique et la sélectivité(43).

1.3 . Les prérequis dans le développement d'une tête de série

Le développement d'un nouveau médicament est un processus complexe, coûteux et qui prend beaucoup de temps(20, 46, 47). À elles seules, ces étapes peuvent prendre de 12 à 15 ans et elles peuvent coûter jusqu'à 2,5 milliards de dollars américains(20, 46). Le développement d'un médicament pouvant être mis sur le marché comprend quatre grandes étapes : la découverte d'une tête de série, le développement pré-clinique, le développement clinique et l'approbation de Santé Canada ou de la FDA (*Food and Drug Administration*)(46).

Parmi les quatre étapes du développement d'un médicament, cette thèse se focalise principalement sur la première étape qui est la découverte d'une tête de série. La découverte

d'une tête de série consiste à identifier un inhibiteur ayant l'effet désiré avec une affinité suffisante pour poursuivre à la prochaine étape(46).

Plusieurs approches peuvent être utilisées afin d'identifier une tête de série inhibant une cible thérapeutique. Ces approches peuvent être *in silico*, *in vitro* ou encore *in vivo*. En chimie médicinale, ces approches sont utilisées à des étapes précises. Ainsi, les approches *in silico* sont souvent utilisées afin d'identifier les composés ayant potentiellement un effet désirable *in vitro* ou *in vivo* à partir d'une large librairie(48, 49).

Par exemple, le criblage *in silico* ou « *virtual screening* » (VS) est une approche largement utilisée dans le domaine de la chimie médicinale afin d'identifier les molécules effectrices à partir d'une large librairie(48, 49). L'approche du VS possède plusieurs avantages dont la rapidité et le faible coût pour identifier des têtes de séries(49). Cependant, il est peu probable d'identifier de puissantes têtes de séries avec une affinité dans le nM, contrairement au criblage à haut débit(49). Un VS a été effectué pour la DfrB1 résultant en la découverte de quelques inhibiteurs hydrophobes. Le site actif de la DfrB1 étant un grand tunnel hydrophobe, ceci crée un artefact *in silico* forçant le criblage à identifier les composés hydrophobes de tailles élevées. La solubilité de ceux-ci devient alors un problème lors de la validation *in vitro* ou *in vivo*.

Les prochaines sections présentent les approches utilisées au courant de mes études doctorales pour découvrir une tête de série ciblant la DfrB1. Toutes ces approches visent à inhiber l'activité de la DfrB1 en ciblant son site actif. Ainsi, des inhibiteurs sélectifs et compétitifs sont recherchés. Les approches sont la conception d'inhibiteurs basée sur les fragments (FBLD), la conception d'inhibiteurs basée sur la structure des substrats et le criblage à haut débit.

1.4 . Les approches utilisées pour la découverte d'inhibiteurs de la DfrB1

1.4.1 . La conception d'inhibiteurs basée sur les fragments (FBLD)

La FBLD est définie dans la langue anglaise par « fragment-based lead design ». Le terme « fragment-based drug design » est aussi employé, mais le terme « fragment-based lead

design » est de plus en plus utilisé puisque l'approche n'a pas toujours pour but de trouver un médicament, mais bien une tête de série pouvant éventuellement être un médicament(50). La FBLD est une approche efficace ne nécessitant pas une grande librairie de composés(51). Cette approche consiste à identifier des inhibiteurs de faible poids moléculaire (fragments) se liant de manière faible mais spécifique à l'enzyme cible. Ensuite, un ou plusieurs de ces fragments sont combinés et/ou complexifiés afin de construire des molécules ayant une affinité plus forte que celle des fragments individuels(51, 52). Les fragments possèdent une masse moléculaire normalement inférieure à 250 g/mol, mais pouvant parfois aller jusqu'à 350 g/mol(51). La petite taille des fragments fournit plus de liberté pour leur optimisation en une tête de série tout en simplifiant l'interprétation de la relation structure-activité (SAR) et ainsi, augmente les chances qu'une molécule se rende aux essais cliniques(51, 53). Cette approche génère un haut taux de succès de tête de série et elle connaît beaucoup de succès en recherche académique ainsi que dans l'industrie pharmaceutique(51).

Afin de diriger la construction d'une molécule plus complexe, plusieurs méthodes structurales sont utilisées telles que la cristallographie à rayons X et la résonance magnétique nucléaire (RMN). Une fois l'emplacement du ou des fragments déterminé, la molécule plus complexe peut être construite et optimiser par des approches différentes (Figure 1-5) : la liaison de fragments, la croissance de fragment ou la fusion de fragments(50).

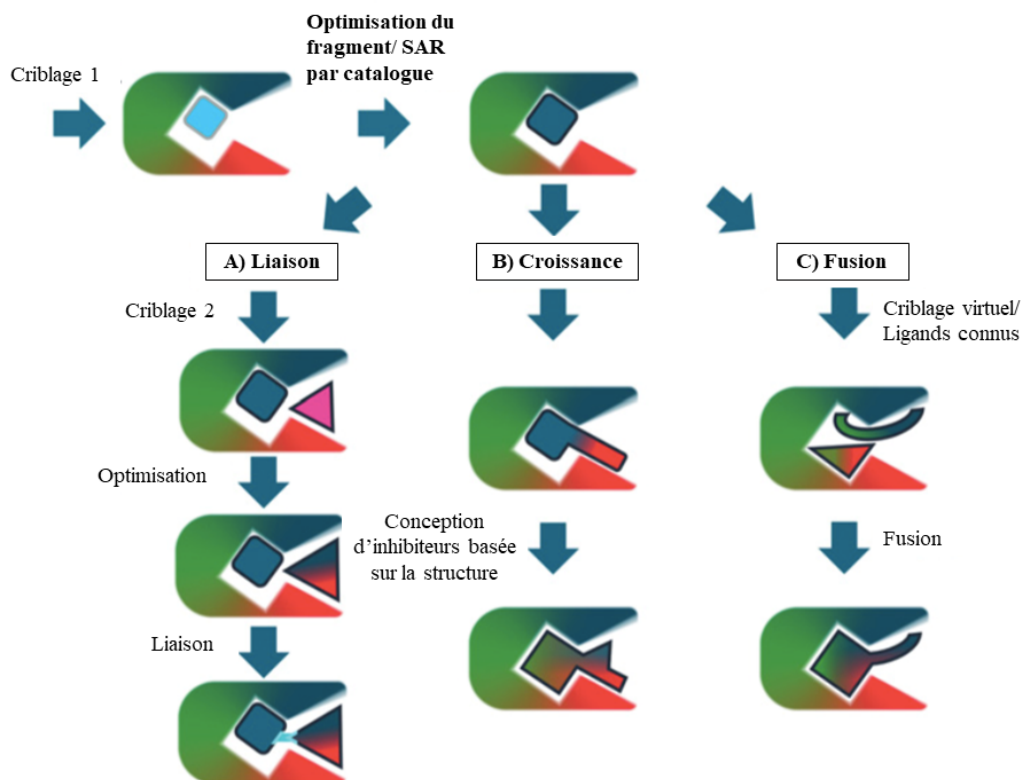


Figure 1-5. Stratégies utilisées dans l'optimisation des fragments.

A) la liaison de fragments; B) la croissance de fragment; C) la fusion de fragments. (Figure adaptée de Lamoree et al. (2017)) (50)

La liaison de fragments, comme le terme l'indique, consiste à relier les fragments ensemble (Figure 1-5)(50). Cette approche est peu utilisée puisqu'il est difficile de trouver un site de liaison distinct pour les fragments et une fois reliés, de maintenir leur position dans les sites de liaison respectifs(50). L'optimisation devient alors une étape laborieuse. *Hajduk et al.* a répertorié une tête de série inhibant une endopeptidase impliquée dans plusieurs maladies telle que l'arthrite(54, 55). Les auteurs recherchaient des inhibiteurs non peptidiques, ce qui différait de ceux retrouvés dans la littérature(54, 55). Les résultats du premier criblage rapportent deux fragments inhibiteurs possédant de K_d dans l'ordre du mM et du μ M(54, 55). La liaison des fragments a permis d'obtenir une tête de série non peptidiques avec un K_d dans l'ordre du nM(54, 55).

La croissance de fragment est l'approche la plus utilisée puisque celle-ci débute par seulement un fragment lié à la cible puis celui-ci se complexifie et est optimisé afin d'augmenter l'affinité (Figure 1-5)(50). Cette approche a été utilisé pour la découverte de Zelboraf®, un

médicament qui inhibe une kinase impliquée dans le mélanome de stade avancé(51). Cette découverte a commencé par le criblage de 20 000 fragments ayant des masses moléculaires de 150 à 350 g/mol(51). Ensuite, un fragment a été optimisé par complexification menant au médicament Zelboraf®. Une variante de cette approche consiste à chercher dans un catalogue de composés complexes ceux ayant le fragment déjà incorporé dans leur structure (*SAR par catalogue*). Ceci permet alors de rétrécir la taille de la librairie à cribler en faisant un criblage dirigé(50).

La fusion de fragments, quant à elle, consiste à construire une molécule complexe avec une efficacité de liaison élevée, à partir de plusieurs groupements fonctionnels de différents fragments(50). Le tout est dirigé par des structures résolues afin de déterminer la liaison des inhibiteurs trouvés(50). L'acquisition des structures résolues de la cible avec le fragment inhibiteur peut être une tâche laborieuse. Ceci peut aussi être jumelé avec des ligands connus dans la littérature(50). Cette approche a permis la découverte d'un inhibiteur, AUY922, présentement en essais cliniques pour le mélanome et plusieurs cancers dont celui du sein et du poumon(56). Le fragment de départ inhibait dans le mM alors qu'une fois complexifié, AUY922i inhibe la croissance des cellules cancéreuses du colon (*HCT116 cells*) avec une concentration de l'ordre du nM(56).

Les approches rationnelles du FBLD ont permis la découverte de plusieurs molécules inhibitrices ayant le potentiel de devenir un médicament. Du fragment à la tête de série, le tout est effectué de manière efficace, rationnellement.

1.4.2 . La conception d'inhibiteurs basée sur les fragments (FBLD) : le cas de la DfrB1

La FBLD a permis d'identifier deux inhibiteurs complexes sélectifs envers la DfrB1 (composés **1.8** et **1.9**; Figure 1-6) (18). D'abord, un criblage semi-automatisé fut réalisé sur plus de 100 fragments de faible masse moléculaire. Les fragments criblés étaient des petites molécules hydrophobes, planaires et aromatiques (Figure 1-6) possédant une similarité générale avec le groupement ptérine du DHF (Figure 1-2). L'hydrophobicité des fragments complémente le site actif hydrophobe de la DfrB1 ce qui augmente leur liaison. L'aromaticité permet d'augmenter la liaison avec les résidus Y69-Y69' du site actif afin d'imiter l'empilement observé

pour les substrats (Figure 1-4B). Sept inhibiteurs faibles furent identifiés (Figure 1-6A). Ceux-ci furent testés en parallèle avec la hDfr afin d'identifier ceux étant sélectifs envers la DfrB1. Ces fragments inhibiteurs furent caractérisés selon deux paramètres: la concentration inhibitrice à 50% (CI₅₀) et K_i. Ces fragments inhibiteurs possèdent une valeur de K_i qui se situe entre 51 - 530 μM, mais seulement le composé inhibiteur **1.7** parmi ceux-ci est sélectif à la DfrB1(18). Un groupement fonctionnel se retrouve dans la structure de six fragments sur sept : l'acide carboxylique, suggérant son importance dans la liaison à sa cible (Figure 1-6A).

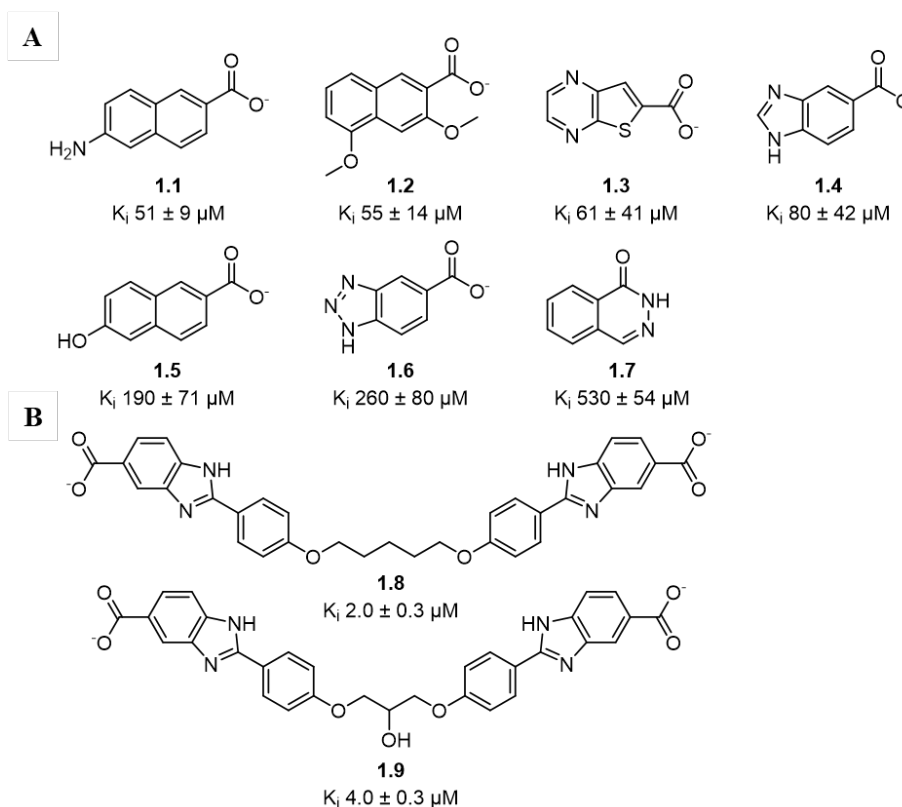


Figure 1-6. Structure des inhibiteurs connus de la DfrB1.

A) Les structures représentent les fragments inhibiteurs de la DfrB1 avec une affinité dans l'ordre du mM. B) Les composés inhibiteurs 1.8 et 1.9 découverts au laboratoire Pelletier. (Figure adaptée de Bastien et al. (2012)) (18)

Pour construire un inhibiteur plus complexe et plus puissant, l'approche de croissance des fragments fut utilisée en reliant deux fois le même fragment ensemble(18). Ces inhibiteurs symétriques sont inspirés d'un inhibiteur connu de la DfrB1, le rouge de Congo, et de la symétrie retrouvée dans le site actif (Figure 1-4). Le rouge de Congo possède une valeur de K_i de 2μM (Figure 1-7)(18). Il est fréquemment utilisé comme teinture de coupe pour les microscopes(57).

Malgré sa constante d'inhibition intéressante, ce composé n'est pas sélectif à la DfrB1 puisqu'il lie plusieurs cibles, expliquant son utilisation dans les coupes de microscopes ainsi que sa toxicité(57).

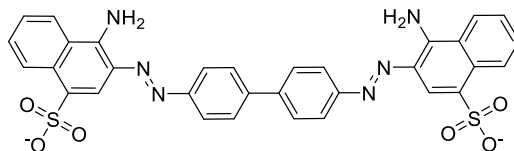


Figure 1-7. Structure du rouge de Congo, un inhibiteur non sélectif de la DfrB1.

Le rouge de Congo inhibe la DfrB1 avec une valeur de K_i de $2\mu\text{M}$ (18). (Figure adaptée de Bastien et al. (2012))(18)

Pour développer des molécules symétriques et complexes avec une masse moléculaire plus élevée, le fragment inhibiteur **1.4** fut utilisé (Figure 1-6A)(18). Malgré sa sélectivité, le fragment inhibiteur **1.7** n'est pas utilisé puisqu'il détient la valeur de K_i la plus élevée. Par l'approche de croissance des fragments, deux molécules **1.4** furent liées afin de créer les composés inhibiteurs **1.8** et **1.9**. Ceux-ci possèdent une valeur de K_i de 2 - 4 μM , ce qui est trop faible pour être considéré dans les phases cliniques sachant que la valeur de K_i du TMP pour la Dfr chromosomique d'*E. coli* est dans l'ordre du picomolaire(16, 18, 58).

Des études stœchiométriques du composé inhibiteur **1.9** indiquent que le site actif serait occupé par plus d'une molécule(18). La détermination de cette stœchiométrie a été exécutée par le graphique de Hill qui met en relation le logarithme de l'activité enzymatique en fonction de la concentration d'inhibiteur(18). Ainsi, le composé inhibiteur **1.9** avec la DfrB1 possède un coefficient de Hill de 2,2 ce qui indique la présence de plus qu'une molécule dans le site actif(18). Toutefois, des expériences *in silico* avec le logiciel Molegro Virtual Docker suggèrent la présence d'une seule molécule inhibitrice de conformation en U dans le site actif de la DfrB1 (Figure 1-8)(18). Aussi, une tentative a été effectuée afin de visualiser *in vitro* la position de l'inhibiteur dans le site actif. Pour ce faire, le patron de diffraction aux rayons X du cristal de la DfrB1 avec le composé inhibiteur **1.9** a été obtenu(18). La résolution de la structure cristalline est trop faible pour permettre la visualisation des interactions entre la DfrB1 et le composé inhibiteur **1.9**(18). Cependant, certaines informations recueillies permettent d'avancer que l'inhibiteur se retrouve en deux copies sous forme allongée, le long du site actif(18).

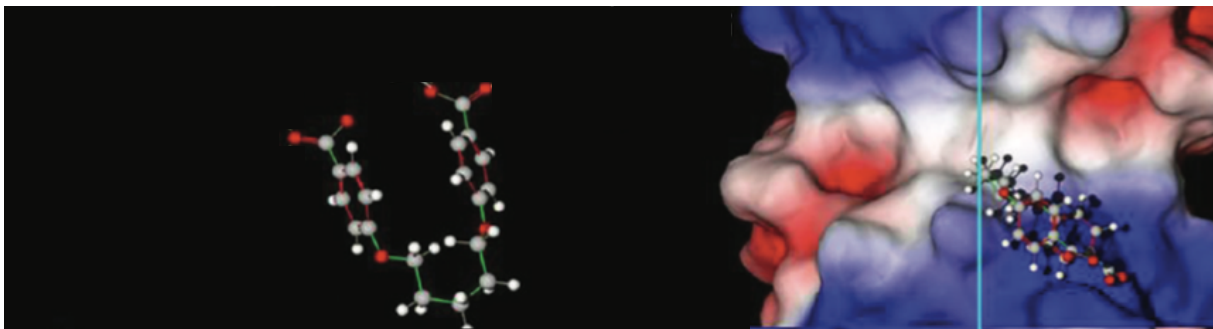


Figure 1-8. Représentations *in silico* de la conformation en U du composé inhibiteur 1.8 lors de la liaison avec la DfrB1.

À gauche, la conformation en U du composé inhibiteur 1.8 telle que retrouvée dans le complexe simulé avec la DfrB1, à droite. La ligne bleue pâle représente l'axe de symétrie de la DfrB1 (Figure adaptée de Bastien et al. (2012)) (18)

L'approche de croissance des fragments a permis d'obtenir deux composés inhibiteurs (**1.8** et **1.9**) plus puissants que les fragments individuels. Cependant, l'optimisation n'étant pas idéale, les composés inhibiteurs **1.8** et **1.9** peuvent encore être améliorés. D'avantages études sur le mode de liaison de ces inhibiteurs pourraient diriger la conception d'analogues afin d'obtenir un inhibiteur plus puissant.

1.4.3 . La conception d'inhibiteurs visant deux cibles thérapeutiques, basée sur la structure des substrats

Dans le développement d'un antibiotique, il faut toujours considérer la résistance bactérienne qui peut survenir. La thérapie combinatoire de plusieurs médicaments possède des avantages intéressants. Par exemple, la bactérie à Gram négatif peut être plus susceptible à un des antibiotiques administrés(59). Parfois, deux antibiotiques administrés possèdent une action synergique; autrement dit, la combinaison des antibiotiques est supérieure à la somme de leur action individuelle(59). Plus particulièrement, cette combinaison ralentit l'émergence de bactéries résistantes(59). Cependant, une combinaison de deux ou plusieurs médicaments peut affecter les profils pharmacocinétique et pharmacodynamique (toxicité) de ceux-ci ce qui est un inconvénient majeur(59). Dans le cas de la tête de série de la DfrB1, si celle-ci n'inhibe pas la Dfr chromosomique bactérienne, elle devra être administrée en bithérapie avec le TMP, rendant difficile le maintien de la combinaison TMP-sulfamide afin d'éviter l'émergence d'une

résistance. Une autre approche utilisée pour contrer la prise de plusieurs médicaments est le développement d'une seule tête de série qui atteint deux cibles thérapeutiques.

Dans la découverte d'inhibiteurs, la première étape consiste à identifier un ligand et ce, peu importe sa puissance inhibitrice. Cette puissance pourra être augmentée par la suite par des modifications structurelles du ligand. Les substrats étant eux-mêmes des ligands à une enzyme, il est ainsi possible de prendre les substrats comme point de départ. Le TMP est un exemple d'inhibiteur basé sur un des substrats de la Dfr(60). Afin de trouver des inhibiteurs pour celle-ci, les interactions observées entre le DHF et la Dfr ont été imitées en utilisant la structure du DHF comme point de départ. D'autres analogues du DHF ont été conçus tels que le méthotrexate (MTX)(60). Le MTX lie similairement à la hDfr, comme le DHF, le caractérisant d'inhibiteurs compétitif avec un $K_i < 1\text{nM}$ (60). La structure cristalline et la déconstruction du MTX démontrent que le 2,4-diaminopyrimidine est un pharmacophore(60). L'optimisation de ce pharmacophore par l'ajout de substituant a mené à la conception de l'inhibiteur sélectif qu'est le TMP(60).

L'enzyme 6-hydroxyméthyle-7,8-dihydroptérine pyrophosphokinase (HPPK) transfère un pyrophosphate de l'ATP à la 6-hydroxyméthyle-7,8-dihydroptérine (HP) et elle est aussi importante dans la voie des folates que les Dfrs (Figure 1-1)(61). Contrairement aux Dfrs, HPPK est présente seulement chez les microorganismes. Ceci fait de la HPPK une cible intéressante pour le développement d'un antibiotique. Les bisubstrats sont des composés constitués des groupements chimiques de deux substrats pour une enzyme ayant deux substrats. Ainsi, en utilisant la structure des substrats de la HPPK, des bisubstrats furent conçus afin d'inhiber la HPPK. Plusieurs inhibiteurs furent rapportés dont le composé **1.10** qui possède un K_d de 2,55 μM et un IC_{50} de 3,16 μM pour la HPPK (Figure 1-9A)(61-63). Par ailleurs, ces inhibiteurs (composé **1.11**; Figure 1-9B) sont constitués d'une partie ptérine et d'une partie adénosyle analogue à l'ATP, structures semblables aux substrats de la DfrB1 (Figure 1-2). Il est envisageable de concevoir des inhibiteurs de type bisubstrat qui agirait à la fois sur HPPK et DfrB1 créant des inhibiteurs doubles.

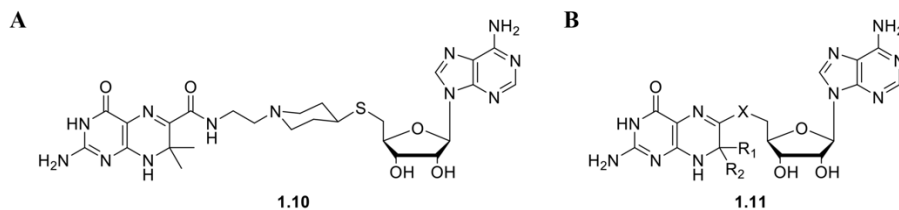


Figure 1-9. Structures d'un bisubstrat, le composé 1.10.

A) Exemple de composé synthétisé au laboratoire du Dr Xinhua Ji(61, 62, 64). $K_d = 2,55 \mu M$ et $IC_{50} = 3,16 \mu M$. B) Structure générale d'un bisubstrat. Les notations R_1 , R_2 et X représentent des groupements variables. Le noyau ptérine est à la gauche du groupement X tandis que l'extrémité adénosyle est située à la droite du groupement X . (Figure inspirée de Shi et al. (2012))

1.4.4 . Les groupements chimiques importants dans la liaison des substrats

Il existe quelques ligands de la DfrB1 rapportés dans la littérature, analogues aux NADPH et DHF(44). Les essais cinétiques effectués en présence de fragments de ces analogues a permis de déterminer les groupements chimiques importants dans la liaison des substrats. La liaison du NADPH à la DfrB1 dépend principalement de l'amide du nicotinamide, du pyrophosphate ainsi que de l'adénosyl-2'-phosphate (Figure 1-2)(44). Le DHF nécessite principalement l'amide de la ptérine et les acides carboxyliques de l'extrémité 4'-aminobenzoylglutamate (Figure 1-2)(43, 44). La ptérine tolère mieux les modifications sans affecter sa liaison que la partie 4'-aminobenzoylglutamate(44).

Ainsi, la ptérine du DHF est un des points d'ancrage au centre du site actif(44, 45). La partie adénosine du NADPH contribue faiblement à la liaison du cofacteur à l'enzyme car elle se retrouve à l'extérieur du site actif, à un site différent de la ptérine. Ainsi, en liant la ptérine et l'adénosine, il est possible de créer un composé de type bisubstrat, qui peut de surcroît être diversifié (Figure 1-9). En plus d'inhiber la HPPK, cette librairie de bisubstrats possède un potentiel pour aussi inhiber la DfrB1.

1.4.5 . Le criblage à haut débit d'une librairie de sources naturelles : les extraits d'invertébrés marins

La découverte de nouvelles entités chimiques provient de plusieurs sources dont les plantes, les mycètes, les bactéries ou encore, les espèces marines(47). Les produits naturels ont été une source pour la plupart des médicaments soit en étant utilisés tels quels, soit en constituant la charpente du médicament(65). Ces médicaments basés sur les produits naturels couvrent un large éventail de maladies telles que les infections, le cancer et le diabète(65). De plus, les produits naturels possèdent une variété de structures chimiques comprenant, entre autres, plusieurs centres chiraux, un nombre élevé d'atomes d'oxygène ainsi que de liaisons d'hydrogène, une large complexité d'encombrement stérique et une valeur d'hydrophobicité (logP) faible(66). Ces caractéristiques uniques et avantageuses peuvent devenir des désavantages pour les chimistes médicaux; leur synthèse devient alors difficile(66). Puisque pour la plupart des sources naturelles de produits d'intérêt ne sont pas indéfiniment renouvelables et qu'elles soient contrôlées par des lois très strictes, la découverte de nouvelles molécules est alors ralentie(65, 66). Ainsi, il devient impératif de développer de nouvelles méthodes de synthèse(66).

Cependant, malgré qu'ils ne respectent pas complètement les règles de Lipinski, la plupart des produits naturels possèdent une activité biologique et un profil ADMET favorable(66). Le Paclitaxel, un agent chimiothérapeutique découvert dans l'écorce d'arbre *Taxus brevifolia*, est un exemple de médicament qui n'aurait jamais été pris en considération puisqu'il ne respecte pas la règle de Lipinski(66). Malheureusement, sa source naturelle, *Taxus brevifolia*, est devenue une espèce en danger, stimulant le développement de voies de synthèse de ce composé complexe(66).

Les extraits biologiques ont la capacité de contenir un nombre incalculable de composés divers qui sont des métabolites primaires ou secondaires et qui n'ont jamais précédemment été caractérisés(67). Les extraits biologiques sont en quelque sorte des librairies de composés naturels. Plusieurs exemples de composés naturels ont prouvé leur potentiel pour l'inhibition de cibles. L'acide acétylsalicylique, commercialisé sous le nom d'Aspirine, est dérivé de l'acide salicylique. Ce dernier, anciennement dérivé des feuilles de saule qui étaient broyées, était utilisé en médecine traditionnelle afin de diminuer les maux de tête(68).

Le criblage d'extraits d'invertébrés a permis de trouver une tête de série comme alternative à des médicaments utilisés par les diabétiques(67). Ces médicaments inhibent une enzyme humaine : l' α -amylase pancréatique (HPA), qui est importante pour la régulation glycémique (67). À partir d'un criblage de 30 000 extraits biologiques, *Tarling et al.* ont découvert 25 extraits inhibant l'activité de la HPA(67). Par la suite, ils ont affiné jusqu'à un seul extrait biologique(67). Cet extrait renfermait une molécule d'une masse de 802.6 g/mol qui inhibait la HPA dans l'ordre du bas μ M(67). De plus, cet inhibiteur a permis de découvrir le mode d'inhibition de deux inhibiteurs précédemment découverts(67). Avec ces résultats encourageants, une autre librairie de 10 000 extraits biologiques a été criblée envers la HPA(69). *Tysoe et al.* ont alors découvert un inhibiteur peptidique de 44 acides aminés qui semble inhiber la HPA dans le nM, de manière sélective(69). Les peptides sont rarement utilisés comme médicaments puisqu'ils possèdent généralement une faible biodisponibilité causée par leur faible stabilité et leur faible perméabilité de la membrane intestinale(70). Sachant que la HPA se retrouve dans le système gastro-intestinal, leur imperméabilité de la membrane intestinale est un atout, nécessitant moins d'agent médicamenteux, mais leur faible stabilité reste problématique(69, 70). Cet exemple, comme celui du Paclitaxel, démontre que la règle de Lipinski n'est pas absolue dans le développement d'un médicament et qu'il existe quelques exceptions.

Un autre criblage d'extraits d'invertébrés marins tels que des éponges et des ectoproctes broyés a été testé contre une cible bactérienne, la pyruvate kinase de *Methicillin-resistant Staphylococcus aureus* (MRSA)(71). Plus d'une dizaine d'inhibiteurs spécifiques contre cette cible ont été découverts dans l'ordre du nM(71). Ces inhibiteurs sont des *bis*-alcaloïdes indoliques(71).

Ces trois criblages ont en commun que les librairies d'extraits biologiques proviennent du même laboratoire, soit celui du Prof. Raymond Andersen de l'Université de la Colombie-Britannique(67, 69, 71). Sachant que ces librairies renferment une panoplie de composés à potentiel thérapeutique, il est intéressant de tester une de ces librairies pour découvrir des inhibiteurs de la DfrB1. La méthode pour le criblage d'une librairie du Prof. Raymond Andersen sera développée grâce à la station automatisée du laboratoire.

1.4.6 . Détermination de la qualité d'un criblage à haut débit

Le criblage à haut débit est une approche très utilisée dans les compagnies pharmaceutiques puisqu'il permet de cribler facilement une grande librairie de composés en peu de temps en utilisant de faibles volumes de réactifs (72, 73). En général, l'effet d'une molécule est testé en duplicata (ou seulement une fois) à une seule concentration contre une cible biologique(74). Il est important d'obtenir une sensibilité, une reproductibilité et une précision adéquates afin d'identifier correctement les molécules effectrices parmi un large éventail de composés(72, 73). L'essai doit permettre d'identifier les molécules à faible effet tout en évitant les faux positifs et les faux négatifs(73, 74). La variabilité peut être causée par des erreurs instrumentales ou humaines(73).

Le facteur Z' ou « Z' -score » est une mesure statistique quantitative reflétant la qualité du criblage(72, 73, 75). Le facteur Z' se base sur les contrôles positifs and négatifs(73). Aussi nommé le coefficient de la fenêtre de criblage, ce facteur est souvent utilisé dans le développement et l'optimisation de méthode de criblage(73). Il permet de chiffrer un seuil de sélection adéquat afin d'identifier les données qui en diffèrent significativement. La détermination de ce seuil définit non seulement le nombre de molécules effectrices identifiées, mais aussi le nombre de faux positifs et de faux négatifs(73). Souvent, les faux négatifs sont priorisés au détriment des faux positifs afin d'augmenter les probabilités de trouver un puissant inhibiteur(74). Plus le seuil de sélection est proche de la moyenne des échantillons, plus de faux positifs seront considérés alors que le nombre de faux négatifs diminue(73).

L'automatisation du criblage ainsi que l'utilisation de plaques à 96, 384 et 1536 puits, augmentent la rapidité du criblage(76). Néanmoins, il est important de vérifier la stabilité de la cible, des effecteurs et des substrats, le cas échéant, sous les conditions et pour la durée du criblage. Le profil des molécules effectrices dépend la librairie de composés ainsi que la concentration testée de ceux-ci lors du criblage(73). Globalement, la qualité du criblage à haut débit, la détermination du seuil ainsi que le profil des molécules effectrices affectent statistiquement la nature des premières molécules effectrices identifiées(73).

1.5 . Conclusion

Il existe très peu de connaissances sur la prévalence des gènes *DfrBs*. Nous savons cependant que la découverte des gènes *dfrBs* a été possible par le suivi des intégrons et que ces gènes sont retrouvés chez les bactéries pathogènes résistantes au TMP retrouvées dans le contexte de l'aquaculture des poissons, l'élevage de bétail ainsi que dans les eaux usées(45, 77-80). La prévalence des gènes *dfrBs* dans les échantillons cliniques n'est donc pas connue. Ainsi, le chapitre 2 présente la recherche directe des gènes *dfrBs* dans des échantillons cliniques Nord-Américains de *E. coli* résistants au TMP. Nous y identifions le gène *dfrB4*. Nous démontrons la dangerosité de la DfrB4 en quantifiant son insensibilité au TMP et en démontrant son association à plusieurs autres gènes de résistance.

Le chapitre 3 exhibe l'optimisation d'inhibiteurs pour la DfrB1 par la détermination de SAR d'analogues des composés **1.8** et **1.9** inhibiteurs. Nous avons élucidé le mode d'inhibition de cette classe d'inhibiteurs par la résolution de la structure cristalline de la DfrB1 avec un des inhibiteurs, appuyée par de courtes simulations de dynamiques moléculaires.

Au courant des dernières décennies, plusieurs connaissances ont été recueillies concernant la DfrB1. Par contre, les connaissances sur les autres membres de cette famille sont faibles. Dans le chapitre 4, nous confirmons l'activité dihydrofolate réductase ainsi que la faible sensibilité au TMP de plusieurs membres de la famille des DfrBs. Nous démontrons l'inhibition de ces DfrBs par les inhibiteurs de la DfrB1, démontrant que notre découverte d'inhibiteurs de la DfrB1 s'étend à la famille des DfrBs. De plus, avec la conception d'inhibiteurs basée sur la structure des substrats, nous découvrons de nouveaux inhibiteurs de type bisubstrat pour cette famille ciblant aussi une autre enzyme de la voie des folates. Ceci illustre la possibilité de concevoir une tête de série ayant deux cibles thérapeutiques.

Enfin, le chapitre 5 présente des études préliminaires visant à cribler le potentiel inhibiteur d'extraits d'invertébrés marins. Les extraits biologiques renferment des composés intéressants ayant des structures chimiques diversifiées. Ces extraits ont démontré à maintes reprises un potentiel énorme dans la découverte de tête de série. Nous présentons au chapitre 5 le développement d'une méthode de criblage à haut débit d'extraits d'invertébrés marins contre la DfrB1.

Préface au chapitre 2

Le séquençage du génome entier d'isolats cliniques d'*E. coli* résistants au TMP a permis d'identifier un membre de la famille des gènes de la dihydrofolate réductase de type II résistante au triméthoprimine (*dfrB*). Le gène *dfrB4* était situé à l'intérieur d'un intégron de classe I flanqué de multiples gènes de résistance. Cette disposition a été précédemment rapportée dans un plasmide de 130,6 kb contenant de multiples résistances, mais excluant le gène *dfrB4*. La protéine DfrB4 a conféré une résistance accrue au triméthoprimine de 2 000 fois dans un système de surexpression chez *E. coli*. Nos résultats sont en accord avec la découverte que *dfrB4* contribue à la résistance clinique au triméthoprimine.

Ces résultats ont été publiés dans la revue scientifique « *Antimicrobial Agents and Chemotherapy* » en mai 2017 (en ligne depuis le 27 février 2017). Ma contribution, sous la supervision du Dr. Joelle Pelletier, dans cet article est l'idée générale de l'article, l'écriture de la première ébauche de l'article, la création des figures et des tableaux, l'analyse de l'identité de séquence entre les membres *dfrB*, l'hypothèse et l'analyse de la reconstruction de la séquence flanquant le gène *dfrB4*, l'alignement BLAST du segment découvert ainsi que les expériences sur la résistance du TMP chez *E. coli* (détermination de la concentration minimale inhibitrice (MIC)). Lorea Alejandre, aussi sous la supervision du Dr. Joelle Pelletier, a contribué à l'écriture de l'article ainsi qu'à la confirmation de la détermination de la MIC. Dr. Joelle Pelletier a participé à la rédaction de l'article. Thaddeus J. Edens, sous la supervision du Dr. Ameer Manges, a fait le criblage *in silico* des gènes des membres *dfrB* alors que le Dr. Ameer Manges, possédant la librairie d'échantillons cliniques d'*E. coli* résistants au TMP, a écrit les parties qui y sont associées.

Chapitre 2 - « *Integron-associated dfrB4, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance.* »

2.1 . Article 1

Jacynthe L. Toulouse^{1,2,3}, Thaddeus J. Edens⁴, Lorea Alejaldre^{1,2,3}, Ameer R. Manges⁵ and Joelle N. Pelletier^{*,1,2,3,6}

¹Département de biochimie, Université de Montréal, C.P. 6128, Succursale Centre-ville Montréal, QC H3C 3J7, Canada

²CGCC, the Center in Green Chemistry and Catalysis, Montréal, QC Canada

³PROTEO, the Québec Network for Research on Protein, Function, Engineering and Applications, Québec, QC, Canada

⁴Devil's Staircase Consulting, North Vancouver, British Columbia, Canada.

⁵School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada

⁶Département de chimie, Université de Montréal, C.P. 6128, Succursale Centre-ville Montréal, QC H3C 3J7, Canada

*Corresponding Author:

Tel: 514-343-2124. Fax: 514-343-7586. E-mail: joelle.pelletier@umontreal.ca.

Running title: *DfrB4* gene in trimethoprim-resistant clinical isolates

Antimicrobial Agents and Chemotherapy. 2017;61:1-5.

2.2 . ABSTRACT

Whole genome sequencing of trimethoprim-resistant *E. coli* clinical isolates has identified a member of the trimethoprim-resistant type II dihydrofolate reductase gene family (*dfrB*). The *dfrB4* gene was located within a class I integron flanked by multiple resistance genes. This arrangement was previously reported in a 130.6 kb multi-resistance plasmid. The DfrB4 protein conferred a > 2,000-fold increased trimethoprim resistance upon overexpression in *E. coli*. Our results are consistent with *dfrB4* contributing to clinical trimethoprim resistance.

KEYWORDS

Type II dihydrofolate reductase, trimethoprim resistance, *E. coli* clinical isolates, *dfrB4*, antibiotic-resistant genes, class I integron, urinary tract infection

2.3 . INTRODUCTION

Public health agencies worldwide rank trimethoprim (TMP) to be a broad-spectrum antibiotic of importance in human medicine (1). Widely used as a result of its low cost and effectiveness, TMP inhibits the activity of many microbial chromosomal dihydrofolate reductases (DHFR), and DHFRs thus have long served as prioritized targets of antiproliferative drugs (2). While the majority of living cells harbor a chromosomal member of the type I DHFR family, encoded by a *dfrA* homolog, the *dfrB* genes encode a family of plasmid-borne type II DHFRs that are evolutionarily unrelated to type I DHFRs. The *dfrB* genes have been found in pathogenic bacteria recovered from many food sources including fish (3), pigs (4, 5) and cows (6) where they confer TMP resistance. Bacteria carrying *dfrB* have also been identified in wastewater samples (7). Over the past decade, *dfrB* genes have been tracked indirectly in antibiotic resistance studies through identification of integron-related elements (8–10). Therefore, the importance of *dfrB* genes in TMP resistance in human pathogens may be underappreciated (11, 12).

To date, only seven members of the *dfrB* gene family are known, and are highly homologous (between 77-94% genetic identity; 77-99% amino acid identity) (**Table 2-I**). Among these, the *dfrB1* protein (also known as R67 DHFR) is the best studied type II DHFR (13–17). It is proposed to be recently-evolved (18) and confers a significant survival advantage

under trimethoprim exposure to microbes that harbor it. The family of *dfrB* genes have, to date, consistently been reported as being contained within the following mobile genetic elements: a highly variable, 57 to 141-bp recombination binding site (*attC*) and the 7-base canonical sequence (7-be) which is identical in all *dfrBs* (5'-GTTRRRY) except *dfrB2* (GTTAGGC) and includes a 7-be at the 3'-end, which is the reverse complement of the 7-be found upstream of the coding sequence (r7-be).

Table 2-I. DNA sequence query coverage and identity (in brackets) of the main type II DHFRs with the expected value (in bold) (coding sequence only).

The expected value (e) represents the probability of randomly matching two different sequences. The lower the e-value, the more significant the match.

	<i>dfrB1</i>	<i>dfrB2</i>	<i>dfrB3</i>	<i>dfrB4</i>	<i>dfrB5</i>	<i>dfrB6</i>	<i>dfrB7</i>
Truncated <i>dfrB1</i>	100%	95% (77%) 2e-47	90% (88%) 3e-77	89% (83%) 5e-61	100% (89%) 9e-90	100% (92%) 7e-98	100% (92%) 7e-98
<i>dfrB2</i>		100%	85% (86%) 2e-66	72% (85%) 2e-54	83% (79%) 3e-45	97% (78%) 4e-50	95% (77%) 2e-47
<i>dfrB3</i>			100%	94% (85%) 2e-73	96% (86%) 4e-75	97% (87%) 7e-79	96% (86%) 8e-78
<i>dfrB4</i>				100%	92% (81%) 8e-59	94% (80%) 1e-57	92% (80%) 4e-56
<i>dfrB5</i>					100%	100% (91%) 3e-96	100% (92%) 1e-100
<i>dfrB6</i>						100%	93% (94%) 7e-111

2.4 . CLINICAL SAMPLE LIBRARY

We examined whole genome shotgun sequencing data for 593 *E. coli* isolates including 380 *E. coli* isolates recovered from 324 individuals with hospital-associated human extra-intestinal infections (19) and 189 *E. coli* isolates recovered from 189 women diagnosed with a community-acquired urinary tract infection (UTI) [Manges unpublished]. Genomic *E. coli* DNA from women with UTI (Manges collection) was extracted using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific). Purified DNA was sheared in water using the Biorupter

Pico (Diagenode) and sequencing libraries were prepared using the TruSeq DNA PCR-Free library preparation kit (Illumina), according to manufacturer instructions. All *E. coli* isolates were sequenced on the Illumina HiSeq 2500 at the UBC Pharmaceutical Sciences Sequencing Center and British Columbia Genome Sciences Centre (Vancouver, BC).

2.5 . *IN SILICO* SCREENING

Seven members of the *dfrB* family (Table A1-I) were used as templates for *in silico* screening, including two variants of *dfrB1* that differ by the absence or presence of 19 additional N-terminal amino acids not essential for the reductase activity (20). The *in silico dfrB* screening was performed by aligning paired-end reads from 569 whole genome sequencing datasets to the set of *dfrB* genes (Table A1-I) in FASTA format with BWA, using standard alignment parameters. Sequencing reads from *E. coli* isolates recovered from blood samples of an individual taken two days apart aligned exactly over the entire *dfrB4* gene sequence (19) (<https://www.ncbi.nlm.nih.gov/sra/SRX560289> and <https://www.ncbi.nlm.nih.gov/sra/SRX560290>). The DNA consensus sequences were identical to the previously deposited 237 bp *dfrB4* sequence (Genbank AY968808 and KP314737.1). Three further samples held one read that overlapped with the *dfrB4* integron with 0-2 mismatches. As no contig within assemblies of these 3 datasets aligned with *dfrB4*, they were not further considered.

2.6 . RECONSTRUCTION OF THE *DFRB4* MOBILE GENETIC ELEMENT AND CONTIGUOUS DNA SEGMENTS

Read sets with significant alignment to any reference *dfrB* gene were assembled into contigs and the *dfrB* genes were aligned with BWA to these assembled contigs. The contigs from both samples have a r7-be (GTTGGGC) 61-bp upstream of the *dfrB4* coding sequence and an *attC* sequence downstream (Figure 2-1). The 7-be found in the *attC* (GCCCAAC) is 53-bp downstream from the coding sequence. These flanking sequences are identical to those of a previously reported *dfrB4* mobile genetic element (Genbank accession #AY970968.1) from *Klebsiella pneumoniae* (21). The *attC* sequence differs by one nucleotide from another *dfrB4* mobile genetic element from *E. coli* (Genbank accession #KP314737.1) (22) (Table A1-I).

Contiguous to the mobile genetic element, we identified a 8963 bp segment derived from SRX560290 and a 9029 bp segment derived from SRX560289 (Figure 2-1). The 66-bp difference did not belong to an ORF and lay outside of the *dfrB4* mobile genetic element. The *dfrB4* gene was carried with *IntI 1* (class 1 integron integrase/recombinase) and *TnpA* (transposase), which are indispensable for integration of mobile genetic elements (23). Moreover, multiple antibiotic, biocide and metal resistance genes were identified: *qacEΔ1* (biocide resistance (24, 25)), *sull* (sulfonamide resistance (26)), the *mphR(A)/mphA/mrx* gene cluster (macrolide/erythromycin resistance (27)) and *chrA* (chromate resistance (28, 29)). The *qacEΔ1* gene encodes for a truncated and less efficient cation efflux pump version of *qacE* (24).

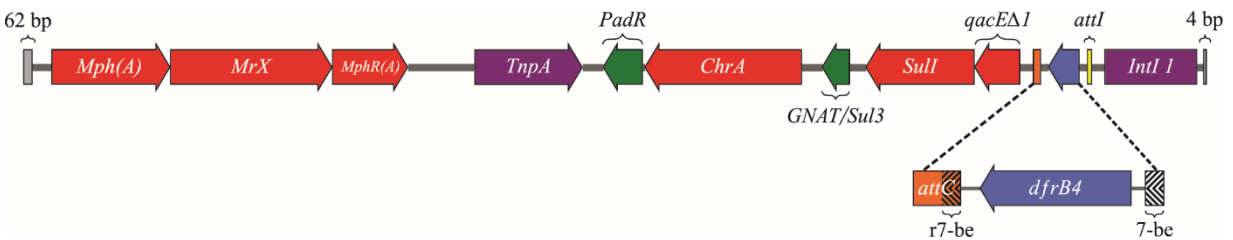


Figure 2-1. Scheme of the 9029 bp DNA segment derived from SRX560289.

The indispensable genes for integration of mobile genetic elements are coloured purple: *IntI* (integrase/recombinase) and *TnpA* (transposase), and the integron-associated recombination site (*attI*) is yellow. Resistance genes (*qacEΔ1*: biocide resistance; *sull*: sulfonamide resistance; the *mphR(A)/mphA/mrx* gene cluster: erythromycin resistance; and *chrA*: chromate resistance) are coloured red except *dfrB4* (dihydrofolate reductase: TMP resistance), which is blue. The recombination binding site (*attC*) is orange while the 7-be sequences are dashed. DNA sequences sharing similarity with known genes are in green: *GNAT/sulI3* (acetyltransferase/sulfonamide resistance) and *padR* (transcriptional regulator).

The 62 bp and 4 bp sequence differences between DNA contigs are in grey.

Additionally, both segments included similarity to the *GNAT* (acetyltransferase) and *sull3* genes (30). The DNA sequence was 96% identical to *GNAT* from *Aeromonas hydrophila* with 72% query coverage (Genbank accession # JX141473.1). The *sull3* gene was 100% identical with 55% query coverage (Genbank accession #EF382672.1), yet includes a 140-nucleotide 3'-deletion. The segments include the little-known *padR* transcriptional regulator. Nine additional ORFs were also identified as hypothetical proteins.

A BLAST alignment of the contig derived from SRX560289 against all organisms resulted in 957 hits with e value $\leq 1 \times e^{-24}$ (31, 32). Among these, a single hit (Genebank accession #CP014320.1) included the *dfrB4* mobile genetic element. It was a 130.6-kb plasmid from a clinical isolate of a patient with recurrent *E. coli* sequence type 131 (ST131) cystitis, the most prevalent *E. coli* lineage in recurrent UTI (33). Other hits include mismatches (≤ 15) and gaps (≤ 38) and exclude the *dfrB4* sequence. This suggests that the *dfrB4* mobile genetic element has recently been incorporated into the CP014320.1 multi-resistance plasmid.

2.7 . THE DFRB4 PROTEIN CONFERS TMP RESISTANCE

The high sequence identity of the DfrB4 protein to the well-characterized TMP-resistant DfrB1 (**Table 2-I**) and the 6 mm TMP zone of inhibition for these clinical samples (S. Salipante, personal communication) corresponding to a MIC $> 16 \mu\text{g/mL}$ (34) strongly argue that DfrB4 should confer TMP resistance to *E. coli*. To corroborate this hypothesis, the consensus *dfrB4* coding sequence was obtained (GeneArt Gene synthesis, Thermo Fisher) with an *N*-terminal hexahistidine affinity tag for eventual purification. It was cloned into pET24-cTEM19m between *Nde*I and *Hind*III, under IPTG-inducible overexpression in *E. coli* BL21 (DE3). The MIC value was determined at least in duplicate through broth microdilution after incubation at 37°C for 14-16 h. We observed high resistance to TMP when expressing DfrB4 ($> 600 \mu\text{g/ml}$, which is the maximal concentration of TMP that is soluble in 5% methanol) (35). The MICs for all negative controls were at least 2,000-fold lower, specifically *E. coli* BL21(DE3) (0.30 $\mu\text{g/ml}$), and *E. coli* BL21(DE3)/pET24-cTEM19m expressing an IPTG-inducible β -lactamase (36) (0.075 $\mu\text{g/ml}$). Thus, the MIC value of DfrB4 greatly surpasses the *E. coli* TMP resistance threshold (2 $\mu\text{g/ml}$), and is associated with 2.5% of all *E. coli* considered to exhibit dangerously high TMP resistance ($\geq 512 \mu\text{g/ml}$) (35, 37). In these experiments, DfrB4 was overexpressed; the expression level of DfrB4 in its native form is unknown.

2.8 . CONCLUSION

To our knowledge, this work constitutes the first report of the type II TMP-resistant *dfrB4* identified in a clinical sample by whole genome sequencing. The 2,000-fold increased MIC for TMP we observed in a transformed *E. coli* points to TMP resistance in the clinical

isolate identified carrying the *dfrB4* gene. The observation of this highly TMP-resistant DfrB4 within a known class I integron indicates that it is being disseminated, although the extent and rate of its propagation are currently unknown as it has not been systematically tracked. Our results highlight the importance of tracking the presence of this gene family in TMP-resistant clinical samples.

2.9 . ACKNOWLEDGEMENTS

We thank Steve Salipante for providing the zone of inhibition data for the original clinical samples and Simon Toulouse for assistance with drafting Figure 2-1.

This work was supported by NSERC Discovery Grant 227853 (J.N.P). J.L.T. received a scholarship from Fonds Québécois pour la Recherche sur la Nature et le Technologies (FRQ-NT). This work was supported by funds from CIHR (A.R.M) (MOP 114879). L.A. received a scholarship from PROTEO, the Québec Network for Research on Protein, Function, Engineering and Applications, which is funded by FRQ-NT. J.L.T. and L.A. received scholarships from Faculté des études supérieures et post-doctorales de l'Université de Montréal (FESP).

2.10 . ABBREVIATIONS

7-be, 7-base canonical sequence; *attC*, recombination binding site; *chrA*, chromate resistance gene; *dfrB*, type II Dfr gene; DHFR, microbial chromosomal dihydrofolate reductase; e, expected value; *GNAT*, acetyltransferase; *IntI*, integrase gene; *mphR(A)/mphA/mrx*, macrolide/erythromycin resistance genes; *qacE*, cation efflux pump gene; *qacEΔ1*, biocide resistance gene; r7-be, reverse complement of the 7-be; *sull*, sulfonamide resistance gene; TMP, trimethoprim; *TnpA*, transposase gene; UTI, urinary tract infection.

2.11 . REFERENCES

1. Government of Canada. 2015. Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) 2013 Annual Report – Chapter 1. Design and Methods. Public Health Agency of Canada, Guelph, Ontario.
2. Ho JMW, Juurlink DN. 2011. Considerations when prescribing trimethoprim-

sulfamethoxazole. *Canadian Medical Association Journal* 183:1851–1858.

3. L'Abée-Lund TM, Sørum H. 2001. Class 1 integrons mediate antibiotic resistance in the fish pathogen *Aeromonas salmonicida* worldwide. *Microbial drug resistance* 7:263–272.
4. Kadlec K, Kehrenberg C, Schwarz S. 2005. Molecular basis of resistance to trimethoprim, chloramphenicol and sulphonamides in *Bordetella bronchiseptica*. *Journal of Antimicrobial Chemotherapy* 56:485–490.
5. Sunde M. 2005. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. *Journal of Antimicrobial Chemotherapy* 56:1019–1024.
6. Barlow RS, Pemberton JM, Desmarchelier PM, Gobius KS. 2004. Isolation and characterization of integron-containing bacteria without antibiotic selection. *Antimicrobial Agents and Chemotherapy* 48:838–842.
7. Tennstedt T, Szczepanowski R, Braun S, Pühler A, Schlüter A. 2003. Occurrence of integron-associated resistance gene cassettes located on antibiotic resistance plasmids isolated from a wastewater treatment plant. *FEMS Microbiology Ecology* 45:239–252.
8. Kehrenberg C, Schwarz S. 2011. Trimethoprim resistance in a porcine *Pasteurella aerogenes* isolate is based on a *dfrA1* gene cassette located in a partially truncated class 2 integron. *Journal of Antimicrobial Chemotherapy* 66:450–452.
9. Partridge SR, Tsafnat G, Coiera E, Iredell JR. 2009. Gene cassettes and cassette arrays in mobile resistance integrons: Review article. *FEMS Microbiology Reviews* 33:757–784.
10. Levings RS, Lightfoot D, Elbourne LDH, Djordjevic SP, Hall RM. 2006. New integron-associated gene cassette encoding a trimethoprim-resistant DfrB-type dihydrofolate reductase. *Antimicrobial Agents and Chemotherapy* 50:2863–2865.
11. Miller LG, Tang AW. 2004. Treatment of uncomplicated urinary tract infections in an era of increasing antimicrobial resistance. *Mayo Clinic proceedings* 79:1048–1053.
12. Sundqvist M, Geli P, Andersson DI, Sjolund-Karlsson M, Runehagen A, Cars H, Abelson-Storby K, Cars O, Kahlmeter G. 2009. Little evidence for reversibility of trimethoprim resistance after a drastic reduction in trimethoprim use. *Journal of Antimicrobial Chemotherapy* 65:350–360.
13. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, Mayence A, Huang TL, Vanden Eynde JJ, Pelletier JN. 2012. Fragment-based design of symmetrical bis-

benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of medicinal chemistry* 55:3182–92.

14. Narayana N. 2006. High-resolution structure of a plasmid-encoded dihydrofolate reductase: Pentagonal network of water molecules in the D2-symmetric active site. *Acta Crystallographica Section D: Biological Crystallography* 62:695–706.

15. Duff MR, Chopra S, Strader MB, Agarwal PK, Howell EE. 2016. Tales of Dihydrofolate Binding to R67 Dihydrofolate Reductase. *Biochemistry* 55:133–145.

16. Yachnin BJ, Colin DY, Volpato JP, Ebert M, Pelletier JN, Berghuis AM. 2011. Novel crystallization conditions for tandem variant R67 DHFR yield a wild-type crystal structure. *Acta Crystallographica Section F* 67:1316–1322.

17. Fling ME, Elwell LP. 1980. Protein expression in *Escherichia coli* minicells containing recombinant plasmids specifying trimethoprim-resistant dihydrofolate reductases. *Journal of Bacteriology* 141:779–785.

18. Alonso H, Gready JE. 2006. Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends in Microbiology* 14:236–242.

19. Salipante SJ, Roach DJ, Kitzman JO, Snyder MW, Stackhouse B, Butler-Wu SM, Lee C, Cookson BT, Shendure J. 2015. Large-scale genomic sequencing of extraintestinal pathogenic *Escherichia coli* strains. *Genome Research* 25:119–128.

20. Reece LJ, Nichols R, Ogden RC, Howell EE. 1991. Construction of a synthetic gene for an R-plasmid-encoded dihydrofolate reductase and studies on the role of the N-terminus in the protein. *Biochemistry* 30:10895–904.

21. Tortola MT, Lavilla S, Miro E, Gonzalez JJ, Larrosa N, Sabate M, Navarro F, Prats G. 2005. First detection of a carbapenem-hydrolyzing metalloenzyme in two *Enterobacteriaceae* isolates in Spain. *Antimicrobial Agents and Chemotherapy* 49:3492–3494.

22. Delpont TC, Harcourt RG, Beaumont LJ, Webster KN, Power ML. 2015. Molecular detection of antibiotic-resistance determinants in *Escherichia coli* isolated from the endangered australian sea lion (*Neophoca cinerea*). *Journal of Wildlife Diseases* 51:555–563.

23. Woodford N, Carattoli A, Karisik E, Underwood A, Ellington MJ, Livermore DM. 2009. Complete nucleotide sequences of plasmids pEK204, pEK499, and pEK516, encoding CTX-M enzymes in three major *Escherichia coli* lineages from the United Kingdom, all belonging to the international O25:H4-ST131 clone. *Antimicrobial Agents and Chemotherapy* 53:4472–4482.

24. Helal ZH, Khan MI. 2015. QacE and QacE Δ 1 Genes and Their Correlation to Antibiotics and Biocides Resistance *Pseudomonas aeruginosa*. *American Journal of Biomedical Sciences* 7:52–62.
25. Jaglic Z CD. 2012. Genetic basis of resistance to quaternary ammonium compounds-the qac genes and their role: a review. *Veterinarni Medicina* 57:275–81.
26. Gillings MR. 2014. Integrons: past, present, and future. *Microbiology and molecular biology reviews* 78:257–77.
27. Zheng J, Sagar V, Smolinsky A, Bourke C, LaRonde-LeBlanc N, Cropp TA. 2009. Structure and Function of the Macrolide Biosensor Protein, MphR(A), with and without Erythromycin. *Journal of Molecular Biology* 387:1250–1260.
28. Caballero-Flores GG, Acosta-Navarrete YM, Ramirez-Diaz MI, Silva-Sanchez J, Cervantes C. 2012. Chromate-resistance genes in plasmids from antibiotic-resistant nosocomial enterobacterial isolates. *FEMS Microbiology Letters* 327:148–154.
29. Ramirez-Diaz MI, Diaz-Perez C, Vargas E, Riveros-Rosas H, Campos-Garcia J, Cervantes C. 2008. Mechanisms of bacterial resistance to chromium compounds. *BioMetals* 21:321–332.
30. Wachino JI, Shibayama K, Kurokawa H, Kimura K, Yamane K, Suzuki S, Shibata N, Ike Y, Arakawa Y. 2007. Novel plasmid-mediated 16S rRNA m1A1408 methyltransferase, NpmA, found in a clinically isolated *Escherichia coli* strain resistant to structurally diverse aminoglycosides. *Antimicrobial Agents and Chemotherapy* 51:4401–4409.
31. Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. 2012. GenBank. *Nucleic Acids Research* 40.
32. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of molecular biology* 215:403–10.
33. Johnson TJ, Aziz M, Liu CM, Sokurenko E, Kisiela DI, Paul S, Andersen P, Johnson JR, Price B. 2016. Complete genome sequence of a CTX-M-15-producing *Escherichia coli* Strain from the H30Rx subclone of sequence type 131 from a patient with recurrent urinary tract infections, closely related to a lethal urosepsis isolate from the patient's sister. *Genome Announce* 4:15–16.
34. CLSI. 2011. Performance standards for antimicrobial susceptibility testing: twenty-first informational supplement. *Clinical and Laboratory Standards Institute* 31:1–172.

35. Wiegand I, Hilpert K, Hancock REW. 2008. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature protocols* 3:163–75.
36. Gobeil S, Gagné D, Doucet N, Pelletier J. 2016. ¹⁵N, ¹³C and ¹H backbone resonance assignments of an artificially engineered TEM-1/PSE-4 Class A β -lactamase chimera and its deconvoluted mutant. *Biomolecular NMR Assignments* 10:93–99.
37. The European Society of Clinical Microbiology and Infectious Diseases. 2016. EUCAST. www.eucast.org.

Préface au chapitre 3

L'utilisation à l'échelle mondiale du triméthoprim (TMP), un antibiotique à large spectre, a provoqué la prolifération de microorganismes résistants à celui-ci. En plus des mutations provoquant une résistance de la dihydrofolate réductase chromosomique microbienne (Dfr), des Dfrs de type II (DfrB), évolutivement et structurellement différentes des Dfrs, ont été identifiées dans des microorganismes résistants au TMP. Les DfrB sont intrinsèquement résistantes au TMP et elles sont disséminées via des plasmides multi-résistances, ce qui en fait des cibles importantes pour le développement d'inhibiteurs. Nous avons précédemment rapporté des bisbenzimidazoles symétriques qui sont les premiers inhibiteurs sélectifs de la DfrB1, la plus étudiée des DfrB1s. Dans cet article, la modification de ces inhibiteurs de la DfrB1 fournit une nouvelle série d'inhibiteurs ($K_i = 1,7 - 12,0 \mu\text{M}$). Nos résultats révèlent deux caractéristiques importantes: les carboxylates terminaux ainsi que la longueur des inhibiteurs permettent d'établir les interactions essentielles avec DfrB1. Deux structures cristallines démontrent la liaison de deux molécules d'inhibiteur dans le site actif. Ces observations ont inspiré la conception d'analogues qui se lient en un seul exemplaire tout en offrant un pouvoir d'inhibition similaire ($K_i = 1,1$ et $7,4 \mu\text{M}$). L'inhibition d'un deuxième membre de la famille des DfrBs suggère un effet large de ces inhibiteurs. Ces résultats fournissent des informations clés sur l'inhibition des DfrBs hautement résistants au TMP et ils ouvrent des pistes pour l'amélioration de ces inhibiteurs.

Les molécules **3.1** et **3.2** de ce chapitre ont déjà été présentées au chapitre 1 et ont été nommées **1.9** et **1.8** respectivement.

Pour le manuscrit soumis le 3 janvier 2019 « *Dimeric bisbenzimidazole inhibitors of an emergent trimethoprim-resistant type II dihydrofolate reductase guide the design of monomeric analogs* », l'idée générale de l'article, soit l'optimisation des inhibiteurs bisbenzimidazoles déjà connus, provient de Dominic Bastien et moi-même sous la supervision du Dr. Joelle Pelletier qui a participé activement à la rédaction de l'article. Ma contribution, dans cet article, est l'écriture de la première ébauche de l'article, la réalisation des expériences cinétiques et la création des figures et des tableaux à l'exception du tableau avec les données cristallographiques. Aussi, j'ai préparé les échantillons protéiques et réalisé la cristallisation d'un des inhibiteurs alors que Dr. Damien Y. Colin a fait la deuxième structure cristalline. Dr. Brahm J. Yachnin, sous la supervision du Dr. Albert Berghui à l'Université de McGill, a analysé les données cristallographiques afin d'obtenir

le modèle final déposé dans la base de données des structures de protéines (PDB) ainsi que la création du tableau qui y est associé. Maximilian Ebert de Chemical Computing Group, a exécuté les simulations de conformations des inhibiteurs alors que j'y ai analysé les données obtenues. Daniel Deon, Marc Gagnon, Edward Ruediger et Kévin Saint-Jacques, sous la supervision du Dr. Anne Marinier, ont synthétisé les séries de composés bromés et de composés asymétriques ainsi qu'un des composés vedettes de cet article. De plus, Marc Gagnon et Edward Ruediger ont contribué à l'écriture des sections « discussion » et « matériel et méthode » concernant ces deux séries de composés. Delphine Forge, sous la supervision de Jean Jacques Vanden Eynde, a synthétisé tous les autres composés.

Chapitre 3 – « *Dimeric bisbenzimidazole inhibitors of an emergent trimethoprim-resistant type II dihydrofolate reductase guide the design of monomeric analogs.* »

3.1 . Article 2

Jacynthe L. Toulouse^{1,2,3}, Brahm J. Yachnin^{2,4}, Edward Ruediger⁵, Daniel Deon⁵, Marc Gagnon⁵, Kévin Saint-Jacques⁶, Maximilian C. C. J. C. Ebert⁷, Delphine Forge⁸, Dominic Bastien^{1,2,3}, Damien Y. Colin^{2,3,9}, Jean Jacques Vanden Eynde⁸, Anne Marinier⁵, Albert M. Berghuis^{2,4} and Joelle N. Pelletier (joelle.pelletier@umontreal.ca)^{1,2,3,9}

¹Département de biochimie et de ⁹chimie, Université de Montréal, QC, Canada

²PROTEO, the Québec Network for Research on Protein, Function, Engineering and Applications

³CGCC, the Center in Green Chemistry and Catalysis, Montréal, QC, Canada

⁴Department of Biochemistry, McGill University, Montréal, QC, Canada

⁵Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, QC, Canada

⁶Département de chimie, Université de Sherbrooke, QC, Canada

⁷Chemical Computing Group, QC, Canada

⁸Laboratoire de chimie organique, Université de Mons, Belgium

*Corresponding Author:

Tel: 514-343-2124. Fax: 514-343-7586. E-mail: joelle.pelletier@umontreal.ca.

Article soumis le 3 janvier 2019, au *Journal of Medicinal Chemistry*.

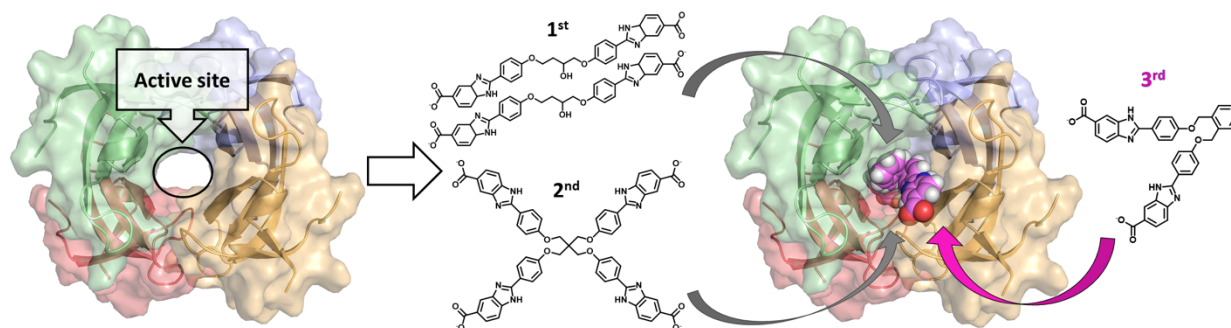
3.2 . ABSTRACT

The world-wide use of the broad-spectrum antimicrobial trimethoprim (TMP) has induced the rise of TMP-resistant microorganisms. In addition to resistance-causing mutations of the microbial chromosomal dihydrofolate reductase (Dfr), the evolutionarily and structurally unrelated type II Dfrs (DfrBs) have been identified in TMP-resistant microorganisms. DfrBs are intrinsically TMP-resistant and are disseminated *via* multi-resistance plasmids, making DfrBs important targets for inhibitor development. We previously reported symmetrical bisbenzimidazoles that are the first selective inhibitors of the best studied DfrB, DfrB1. Here, their diversification provides a new series of inhibitors ($K_i = 1.7 - 12.0 \mu\text{M}$). Our results reveal two prominent features: terminal carboxylates and inhibitor length allow the establishment of essential interactions with DfrB1. Two crystal structures demonstrate the simultaneous binding of two inhibitor molecules in the symmetrical active site. Observations of those dimeric inhibitors inspired the design of monomeric analogues that bind as a single copy yet offer similar inhibition potency ($K_i = 1.1$ and $7.4 \mu\text{M}$). Inhibition of a second member of the DfrB family suggests the generality of these inhibitors. These results provide key insights into inhibition of the highly TMP-resistant DfrBs and open avenues to further improvements.

KEYWORDS

Antibiotic resistance, R67 DHFR, DfrB, selective inhibitor, LowModeMD, inhibition optimization.

TABLE OF CONTENT GRAPHIC



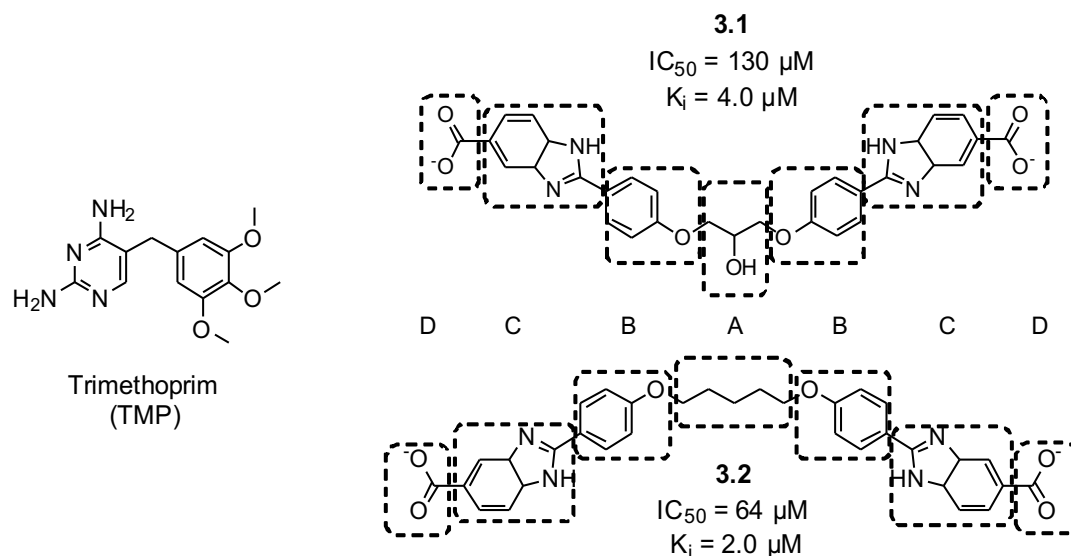
3.3 . INTRODUCTION

Trimethoprim (TMP) is a broad-spectrum diaminopyrimidine antimicrobial agent (Scheme 3-1)(1-3). As a result of its simple structure and efficacy, this low-cost antibiotic has been in constant and heavy use world-wide for decades(1, 2, 4). The combination of TMP and sulfonamide (SUL) is broadly used against aerobic bacteria and occasionally against protozoa(1). The World Health Organization has classified TMP-SUL as highly important for its effectiveness in treatment of human urinary and respiratory tract infections(1, 2, 4).

The TMP-SUL combination is also part of the primary front-line treatment for infections of known or unknown origin in veterinary applications(1) and is used in routine preventive measures, further increasing its distribution. For example, TMP is heavily used in aquaculture of shrimp(5) and fish(5, 6) where it is disseminated into surrounding waters. It is also used in livestock. This is of particular concern because TMP resistance in livestock(7), especially cattle(1) and swine(8, 9), is known to be transferable to humans.(1, 10) As a result, the global incidence of TMP resistance in humans is increasing(11).

Dihydrofolate reductases (Dfr) are ubiquitous. They have long been primary targets for the treatment of proliferative disorders such as cancer and infectious diseases because they are essential for synthesis of DNA precursors(12). Dfrs catalyze the reduction of dihydrofolate (DHF) to tetrahydrofolate with the reducing cofactor NADPH(4). TMP specifically inhibits the bacterial chromosomal Dfr; TMP resistance can result from mutations of the chromosomal Dfr. In addition, a second route to TMP resistance has evolved: the evolutionarily and structurally unrelated type II Dfr (DfrB). DfrB1, also known as R67 DHFR, is the best studied among the DfrB family. It catalyzes the same essential metabolic reaction as chromosomal Dfrs, with low but sufficient efficiency to support bacterial proliferation. However, as a result of sharing no sequence or

structural relationship with the chromosomal Dfrs, it is inert to high concentrations of TMP(13-16). As a result, Gram-negative bacteria harbouring plasmid-borne DfrB are highly resistant to TMP(4, 6, 8, 9, 15, 17, 18). It follows that co-administration of a DfrB inhibitor along with TMP would enhance treatment efficacy.



Scheme 3-1. Trimethoprim (TMP) and known selective inhibitors of DfrB1(19).

The general structure of the symmetrical bisbenzimidazole-type inhibitors 3.1 and 3.2 is defined by A: the central core; B: phenoxy moieties; C: benzimidazole moieties and D: terminal carboxylates.

DfrB1 was the first member of the DfrB family to have been studied and remains the only DfrB to have been structurally and kinetically characterized(19-23). In contrast to all chromosomal Dfrs (and their plasmid-borne, close variants), including the human Dfr (hDfr), that share a conserved monomeric fold with clearly defined substrate and cofactor binding regions, DfrB1 is a doughnut-shaped homotetrameric enzyme with a single, central active site (Figure 3-1). Each protomer consists of an SH3-like domain that contributes equally to the formation of the active site(13, 22, 24). The hourglass-shaped active site tunnel of DfrB1 has a central neck that opens into the opposing tunnel mouths (Figure 3-1). Its calculated volume is greater than twice that of the *E. coli* DfrA active site, despite catalyzing the same reaction(14). As a result of its symmetry, it can bind two DHF substrate molecules, two NADPH cofactor molecules or can bind one of each with positive cooperativity(25-27).

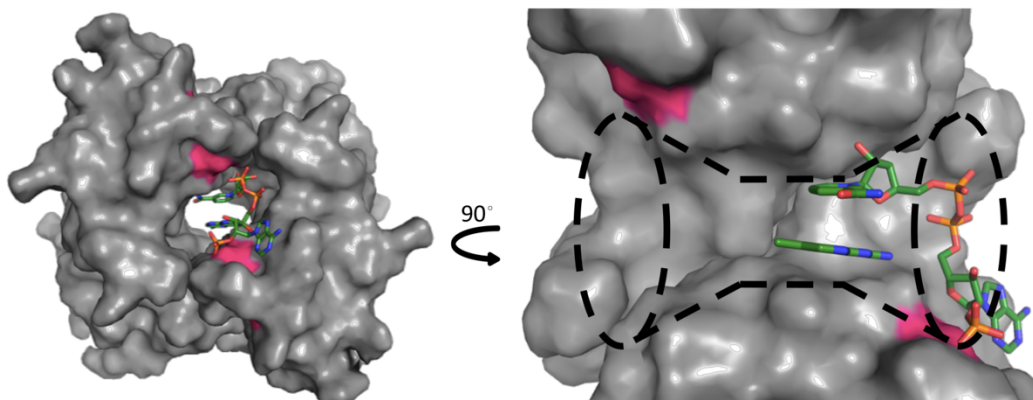


Figure 3-1. Surface representation of the tetrameric DfrB1 with its substrates NADPH and DHF, in sticks representation, bound inside the active-site tunnel (PDB code 2RK1(28)). *Left: Overview. Right: close-up of the tunnel; only two DfrB1 monomers are shown. The pABA-Glu tail of the DHF was not resolved as a result of its dynamic behaviour(29); simulations indicate that it salt-bridges with a Lys32 at the flared active-site mouth (pink surface, left opening), opposite the Lys32 where the NADPH 2'-phosphate forms a similar interaction (pink surface, right opening)(29).*

This unusual binding promiscuity has resulted in DfrBs being considered as poorly evolved enzymes and led to the identification of non-selective inhibitors of DfrB1 some of which bind in two copies(12, 14, 27). 2,4-Diaminopyrimidines, including methotrexate (MTX) and aminopterin (AMT), known for decades to be antimicrobial or antineoplastic inhibitors that target chromosomal Dfrs(3, 30, 31), bind weakly to apo-DfrB1 although no binding was observed with holo-DfrB1(3, 6, 8, 9, 15-18, 27). Pemetrexed (PMTX), 2-desamino-5,8-dideazafolic acid, 5,8-dideazapterin analogs and 5-deazafolic acid analogs non-specifically inhibit DfrB1 in the low μM range but also inhibit hDfr and other mammalian enzymes from the folate pathway with higher affinity, precluding their use as antibiotic agents(12, 27, 32, 33). In addition, the mainly hydrophobic active-site cavity of DfrB1 procures micromolar binding of distinct symmetrical or asymmetrical polycyclic molecules including novobiocin(12, 14, 34).

We previously reported the first selective inhibitors of DfrB1, obtained by fragment-based design(19). A library of > 100 low molecular weight fragments, chosen for their loose resemblance to the NADPH nicotinamide ring or the DHF pteroyl ring, was screened(19). Seven weakly-inhibiting fragments were identified, among which 1H-benzimidazole-5-carboxylic acid was selected to build symmetrical compounds that reflect the symmetry of the DfrB1 active site(19). The resulting bisbenzimidazole-based compounds displayed 100- to 1000-fold tighter binding than

the initial fragment, with K_i in the low μM range (Scheme 3-1) (19). Preliminary crystallographic results suggested that two molecules of **3.1** bind to DfrB1, consistent with kinetic analyses that also determined the competitive nature of binding(19).

Here, we investigated the contribution to binding of the substructures of **3.1** and **3.2**: the central core, the phenoxy moieties, the benzimidazole moieties and the terminal carboxylates (Scheme 3-1). In addition, the length of the molecule and geometric constraints were explored. LowModeMD simulations and steady-state kinetic analyses provide insights into key functional groups. Two crystal structures resolved with either **3.1** or one among the most effective inhibitors confirmed the previously proposed 2:1 ligand:target binding stoichiometry. Only the central cores of the dimeric inhibitors were well resolved, consistent with dynamic binding of the termini. These findings inspired the design of monomeric analogs that were confirmed to bind with 1:1 stoichiometry. Finally, we demonstrate that a second member of the DfrB family, identified in a TMP-resistant North American clinical sample(16), is also effectively inhibited by these novel bisbenzimidazole inhibitors.

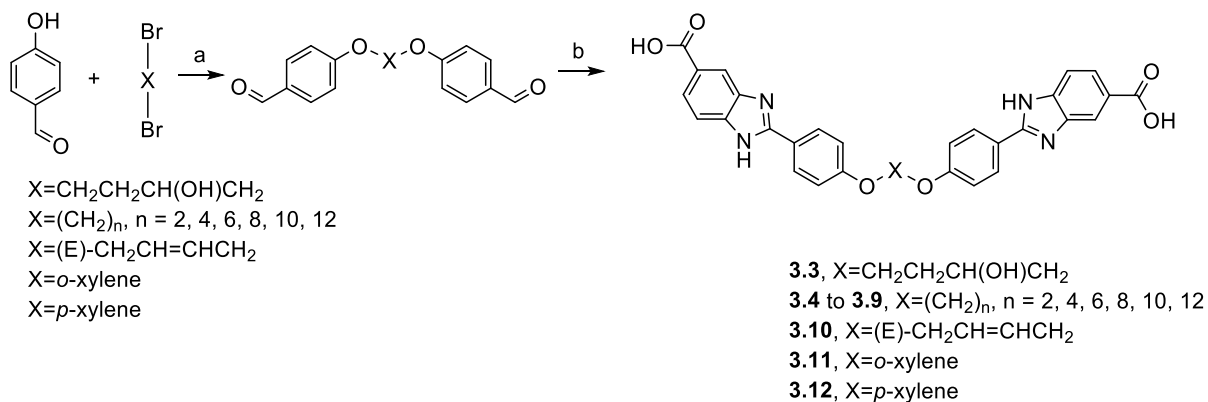
3.4 . RESULTS AND DISCUSSION

The DfrB1 homotetramer is defined by a 222 symmetry where the center of the active site is situated at the intersection of the three axes of symmetry. The asymmetric unit contains one quarter of the active site pore (PDB ID 2RK1). As a result, the active site of DfrB1 is a 24 Å-long tunnel that bisects the tetrameric structure (Figure 3-1)(35). In our prior work, we identified **3.1** (Figure 3-1) as a competitive inhibitor of DfrB1, with a $K_i = 4.0 \mu\text{M}$ (19). Its Hill coefficient of 2.1 and competitive mode of binding are consistent with the simultaneous binding of more than one molecule of **3.1** within the active site cavity.

Crystal structures reveal two bound molecules of inhibitor. The particular symmetry of DfrB1 poses a challenge in the course of analyzing electron density of ligands crystallized in the active site. It has been previously shown(28) that ligand binding typically breaks the crystallographic symmetry of the pore, either by binding two different ligands (NADPH and DHF PDB ID: 2RK1(28)) or by binding a ligand that itself is asymmetric or whose symmetry does not coincide with the crystallographic symmetry axis (PDB ID: 2RK2)(28). In such cases, the electron density observed in asymmetric unit is the average density over the four symmetric quarters of the

active site, rendering interpretation of the electron density difficult, particularly in cases with low ligand occupancy.

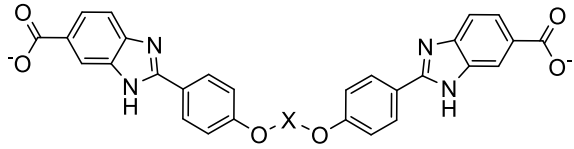
Consistent with this, we previously reported preliminary crystallographic data suggesting that two molecules of **3.1** lie lengthwise in the tunnel(19). In order to confirm these findings, we synthesized **3.3** and validated that inhibition ($K_i = 3.5 \mu\text{M}$) is essentially unchanged from **3.1** ($K_i = 4.0 \mu\text{M}$) (Scheme 3-2; Table 3-I). We obtained a crystal structure of DfrB1 in complex with **3.3**, which we report here along with the completed analysis of the crystal structure of **3.1** (Table A2-I). The inhibitors **3.1** and **3.3** have comparable K_i and are both selective, as no inhibition of hDfr was observed (Figure 3-2). In both structures, two molecules of the soaked inhibitor are juxtaposed within the tetrameric active site, in an extended conformation, providing insight into the binding mode (Figure A2-1). Only the 1,3-dioxopropan-2-ol or 1,4-dioxobutan-2-ol central cores could be resolved. The cores lie in the narrow center of the active site tunnel where they interact with the protein. The quality of the electron density decreased toward the benzimidazole extremities, in the direction of the active-site mouths. This is similar to crystal structure of DfrB1 in complex with DHF, where only the pteridin moiety was resolved in the center of the active site(28), caused by the dynamic behaviour of its *p*ABA-Glu tail(20, 29). A racemate of **3.3** was used in soaking, and due to the ambiguity of the electron density described above, we were unable to definitely differentiate between the two enantiomers in the crystal structure. As such, we chose to model the *R*-enantiomer in the active site.

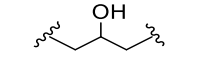
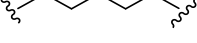
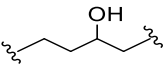
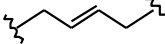
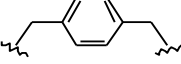


Scheme 3-2. Synthesis of **3.3** to **3.12**.

Reagents and conditions: a) EtOH/H₂O (9:1), NaOH, 120°C, 20 min. b) Na₂S₂O₅, 3,4-diaminobenzoic acid, EtOH/H₂O (3:1), 140°C, 15 min, 19-43%.

Table 3-I. Inhibition upon central core substitution and alteration of the central core length of 3.2.^a



Compound	X	DfrB1		hDHFR	
		IC ₅₀ (μM)	K _i (μM) ^b	IC ₅₀ (μM) ^c	K _i (μM) ^b
3.1		130 ^d	4.0 ^d	>1000	ND
3.2		64 ^d	2.0 ^d	>400	ND
3.3		110 ± 10	3.5 ± 1.7	>500	ND
3.4	-(CH ₂) ₂ -	160 ± 50	4.9 ± 1.6	ND	ND
3.5	-(CH ₂) ₄ -	60 ± 10	1.9 ± 0.3	ND	ND
3.6	-(CH ₂) ₆ -	180 ± 95	5.5 ± 3.0	ND	ND
3.7	-(CH ₂) ₈ -	70 ± 30	2.1 ± 0.9	ND	ND
3.8	-(CH ₂) ₁₀ -	> 100 ^c	ND	ND	ND
3.9	-(CH ₂) ₁₂ -	> 200 ^c	ND	ND	ND
3.10		90 ± 20	2.8 ± 0.8	>400	ND
3.11		[200 ± 130] ^e	[6.1 ± 3.9] ^e	>100	ND

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

^b K_i was calculated from IC₅₀ values.

^c Inhibitors were tested at the highest possible concentration, considering constraints due to high absorbance or to precipitation.

^d Values taken from Table 2 in (19, 26).

^e Extrapolated value, as the compound was insoluble above 100 μM in 10% DMSO.

ND, not determined.

As the binding modes of both **3.1** and **3.3** are slightly offset from the center of symmetry, the inhibitors can be observed in two slightly different binding modes when the asymmetric unit is expanded to show the entire tetramer. The inhibitors were modelled at 0.5 occupancy per asymmetric unit, where four copies of the asymmetric unit are required to form the active-site tunnel. This approximated 2 fully occupied binding sites.

In our initial analysis of the crystal structure of **3.1** in complex of DfrB1(19), the data available were not sufficient for us to confidently conclude that the density could be definitively

modelled as **3.1**. We were able to observe very similar electron density in the crystal structure in complex with the chemically similar **3.3**; furthermore, this density was never visible when no soaking was performed. Given these results, combined with prior reports of ligand-bound DfrB1 structures (28), we are now confident in modelling **1** and **3**, respectively, in these crystal structures.

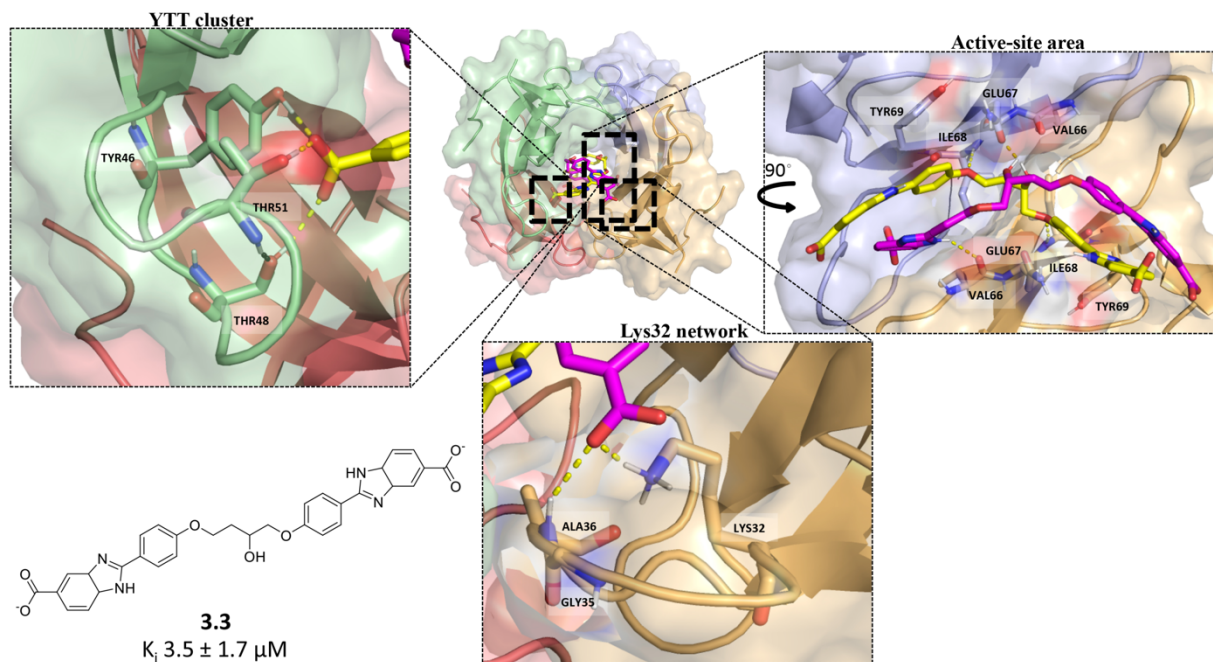


Figure 3-2. *In silico* reconstruction of **3.3** in the active site of DfrB1 based on the crystal structure PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’.

Interactions between residues of DfrB1 with 3.3 are highlighted.

***In silico* reconstruction of the network of interactions upon inhibitor binding.** To model the missing portion of the bound inhibitors, we built the two co-crystallized inhibitor molecules outward from the central core coordinates and performed LowModeMD (Molecular Operating Environment). The crystal structure with **3.3** (PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’) was selected as a template for building all other inhibitors, reserving the crystal structure with **3.1** to validate the models. Throughout the hundreds of simulated poses obtained for the full-length modelled inhibitors, the two molecules of **3.3** consistently adopted a helical conformation along the length of the active site tunnel (Figure 3-2). During the majority of the simulation, the terminal carboxylates formed electrostatic bonds with either of the two Lys32 ϵ -amino groups at each tunnel mouth and established hydrogen bonds with the backbone of the nearby Gly35 and Ala36 (Figure 3-2; Figure

A2-2; Table A2-II; Table A2-III). We will refer to this predominant bonding pattern as the ‘Lys32 network’. On some occasions, the terminal carboxylates also hydrogen bonded with the hydroxyl side chain of Ser34. Alternatively, the terminal carboxylates interacted with the hydroxyl of two out of the three following: Tyr46, Thr48 and Thr51, which we will refer to as the ‘YTT cluster’ (Figure 3-2). This confirms the dynamic binding of the inhibitor extremities consistent with the lack of clear electron density in the bound crystal structures. The benzimidazole moieties established multiple interactions: the amine hydrogen bonded with the backbone carbonyl of Val66, and the phenyl and imidazole rings of **3.3** formed intermittent π -hydrogen bonds or edge-to-face aromatic interactions with the Val66 backbone amine and Tyr69, respectively (Figure 3-2; Table A2-II; Table A2-III). Occasionally, the benzenes from the phenoxy moieties of **3.3** formed π -hydrogen bonds with the backbone amine of Ile68. Finally, the central core hydroxyl established hydrogen-bonding interactions with the side-chain amide of Gln67, as observed in the co-crystal structure.

Geometry of inhibitor binding is orthogonal to that of the bound substrates. Positive cooperativity favors binding of DHF onto the NADPH-bound DfrB1 (holo-DfrB1) to form the productive (DfrB1•NADPH•DHF) complex. NADPH binds first (quadrants I and II) followed by the binding of DHF (quadrants III and IV) according to an ordered mechanism of ternary complex formation (Figure 3-3)(36). Naturally, the symmetrical active site allows binding in the opposite quadrants (NADPH in quadrants III/IV and DHF in I/II). The active site can also accommodate the binding of two DHF or two NADPH molecules(26, 27, 29, 36), illustrating its promiscuity; binding occurs in the same mode (quadrants I/II and III/IV).

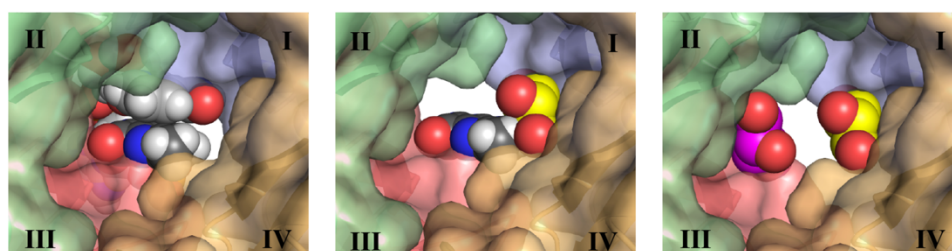


Figure 3-3. DfrB1 with bound NADPH and DHF substrates (PDB ID 2RK1(28); left panel) or with the central core of a pair of molecules **3.3 (PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’; right panel) provide evidence that binding of DHF is incompatible with binding of one molecule of **3.3** (overlay, center panel).**

The crystal structure of DfrB1 demonstrates that **3.3** binds with the orthogonal geometry: quadrants I/IV and II/III are occupied by either of the bound inhibitor molecules (Figure 3-3). As a result of this orthogonal binding mode, binding of an inhibitor molecule is incompatible with binding of NADPH or DHF (Figure 3-3).

Despite the orthogonal binding geometry of the substrates and inhibitors, and despite the differences in their respective chemical groups, it is remarkable that the substrates and **3.3** share key binding residues: Lys32, Ala36, Tyr46, Thr51, Val66, Gln67, Ile68 and Tyr69. Among these, Lys32, Gln67, Ile68 and Tyr69 are functional residues, justifying the inhibitory effect of **3.3**.

In silico reconstruction of the network of interactions with inhibitors 3.1 and 3.2. We used the model of **3.3** bound to DfrB1 as a template to predict binding interactions with **3.1** and **3.2**. One molecule of **3.3** was replaced by a first molecule of the query inhibitor and a first LowModeMD simulation was performed. Substitution of the second molecule of **3.3** by the second molecule of **3.1** or **3.2** was then followed by a second LowModeMD simulation. RMSD of a 2.4 ± 0.6 Å between the eight atoms constituting the 1,3-dioxopropan-2-ol core of **3.1** in the crystal structure (PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’) and in the simulated ensemble of nearly 3,000 conformers obtained validated the model.

The interaction pattern observed for the four segments of **3.1** and of **3.2** throughout the hundreds of poses obtained was similar to that observed for **3.3** (Table A2-II; Table A2-III). The terminal carboxylates formed electrostatic and hydrogen bonds with the Lys32 network, and alternatively hydrogen bonds with the YTT cluster. As with **3.3**, the Lys32 network was more solicited than the YTT cluster (Table A2-II; Table A2-III). The benzimidazole amines and phenol rings again established hydrogen and π -hydrogen bonds with backbones of Val66 and Ile68, and with the side-chain of Tyr69. The hydrogen bond involving the benzimidazole amine and the Gln67 side-chain was lost, yet the hydrogen and π -hydrogen bonds of the phenoxy segment were maintained. The phenoxy oxygens of **3.1** bound the Gln67 side-chain amine as well as the Ile68 backbone amine while the phenoxy oxygens of **3.2** bound instead to the Val66 backbone amine. The phenoxy benzenes established π -hydrogen bonds with the Ile68 backbone amine; the phenoxy benzenes of **3.1** formed a novel edge-to-face interaction with Tyr69. The hydrogen bonds observed with the hydroxyl in the central core of **3.3** and DfrB1 were conserved as for **3.1**. These interactions

were not maintained with the pentyl central core of **3.2** as it lacks H-bonding capacity. These models indicate that the four segments generally contribute to inhibitor binding.

Does inhibitor length matter? The bound crystal structure and simulations indicate that inhibitors **3.1**, **3.2** and **3.3** bind lengthwise within the active-site tunnel in a helical conformation and their carboxylates interact alternatively with the Lys32 network or the YTT cluster at both tunnel mouths. The location of these binding hot-spots suggests that inhibitor length will be a determinant of affinity. To verify this, central cores from two to 12 methylene units in length (**3.4** to **3.9**) (Scheme 3-2) were tested. Inhibition was observed in the low μM range for compounds with central cores of up to eight methylene units ($\leq 36.1 \text{ \AA}$, calculated with Chemdraw 3D; Table 3-I).

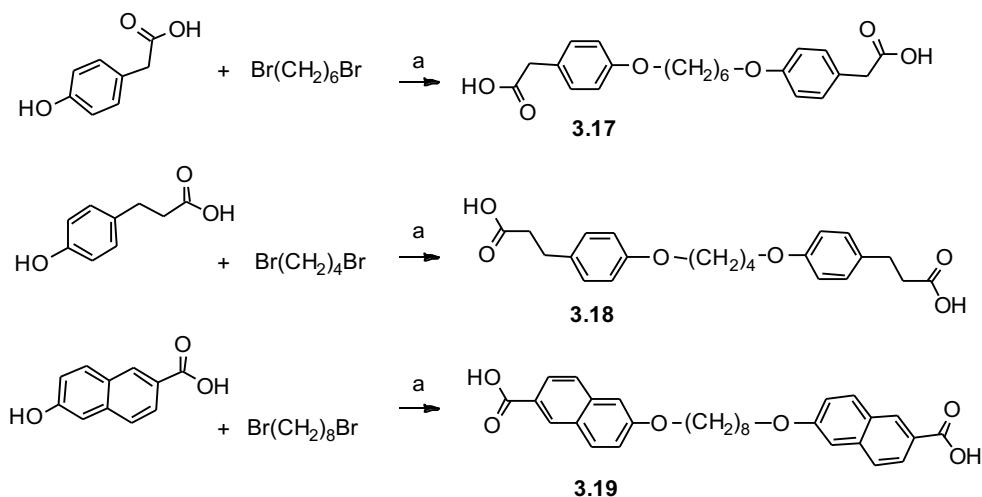
We performed LowModeMD simulations with **3.4** to **3.7** as described above. As with **3.2** the alkyl central cores of those inhibitors cannot establish electrostatic interactions. As the length of the central core increased, the terminal carboxylates bound to the Lys32 network more frequently than to the YTT cluster. The hydrogen and π -hydrogen bonds of benzimidazole moieties of **3.4** to **3.7** remained comparable to the interactions observed with **3** (Table A2-II; Table A2-III). The phenoxy oxygens and phenyls of **3.4** and **3.6** formed interactions with the same amino acids as **3**. In addition, the phenoxy oxygens and phenyls of **3.5** bound to the Val66 and Ile68 backbone amines. The eight-methylene-unit central core of **3.7** prevented interactions of the phenoxy ring with DfrB1. The sole remaining interaction was the π -hydrogen bond with the Ile68 backbone amine. Increasing hydrophobicity of the longest central cores ($> \text{eight methylene units}$, or $> 36 \text{ \AA}$) reduced solubility, limiting the concentrations that could be assayed; nonetheless, it was possible to determine that they procured no inhibition at concentrations where shorter analogs clearly did. As a result, the length of 36 \AA was considered the maximal threshold for the synthesis of further bisbenzimidazole-type compounds.

Central core substitution. The hydrogen bonding established between the hydroxylated central core of **3.1** and **3.3** and the active site suggested that the central core segment was a target for diversification. Substitution with a rigid (*trans*-2-butene; **3.10**) or a bulky (phenyl; **3.11**) moiety (Table 3-I; Scheme 3-2) procured inhibition similar to **3.1** (Table 3-I). Consistent with the good affinity of **3.2**, this demonstrates that hydrogen bonding in the central core is not a key to affinity. Selectivity was maintained, as **3.10** and **3.11** did not inhibit hDfr.

LowModeMD with **3.10** and **3.11** was performed (Table A2-II; Table A2-III). No interactions were observed with the rigid central core of **3.10**, which modified interactions of the benzimidazole moieties and the terminal carboxylates. The benzimidazoles no longer interacted with the Val66 backbone amine but established a π -hydrogen bond with the Ile68 backbone amine. The terminal carboxylates of **3.10** still formed hydrogen bonds with the YTT cluster, except for Thr48.

The phenyl central core of **3.11** established π -hydrogen bonds with the backbone amines of Val66 and Ile68; the former also hydrogen-bonded with the phenoxy oxygen, rather than the benzimidazole nitrogen seen with **3.3**, due to a slight reorganisation. Our results demonstrate that chemical diversity is tolerated in the central core of the bisbenzimidazole inhibitors and suggests that further diversification of the core is possible.

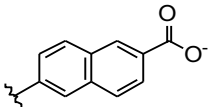
Contribution of the benzimidazole moieties and conformational freedom of the carboxylates. In our previous report(19), inhibition was lost when the benzimidazole moieties were removed from **3.2** (Scheme 3-1). Loss of inhibition may result from removal of key functional groups or from shortening the molecule. To verify this, analogs lacking benzimidazole moieties and of central core length varying from six to twelve methylene units (**3.13** to **3.16**) were assayed (Table 3-II, Scheme 3-3). Only **3.15** demonstrated detectable inhibition of DfrB1. The length of **3.15** (26.9 Å) is slightly less than that of the shortest bisbenzimidazole inhibitor of Table 3-I, **3.4** (28.9 Å). The difference of affinity between those inhibitors resulting from the removal of the benzimidazoles (**3.15**: $K_i = 8.8 \mu\text{M}$; **3.4**: $K_i = 4.9 \mu\text{M}$) is significant yet affinity is largely maintained. Longer analogs lacking benzimidazoles could not be assayed due to insufficient solubility. No inhibition was observed with further analogs lacking benzimidazoles where additional length was built in between the phenoxy groups and the terminal carboxylates (**3.17** and **3.18**) (Scheme 3-3). This is consistent with the benzimidazole moieties contributing to binding, as modeled by LowModeMD of **3.4**, yet demonstrates that they are not critical components of DfrB1 inhibitors (Table A2-II; Table A2-III).



Scheme 3-3. Synthesis of 3.17 to 3.19.

Reagents and conditions: a) EtOH/H₂O (9:1), NaOH, 120°C, 20 min, 8-10%.

Table 3-II. Inhibition with analogs lacking benzimidazole moieties.^a

Compound	X	R ¹	Length (Å)	DfrB1	
				IC ₅₀ (μM)	K _i (μM) ^b
3.13	-(CH ₂) ₆ -	-COO ⁻	21.7	> 500 ^c	ND
3.14	-(CH ₂) ₈ -	-COO ⁻	24.4	≥ 500 ^c	ND
3.15	-(CH ₂) ₁₀ -	-COO ⁻	26.9	285 ± 80	8.8 ± 2.6
3.16	-(CH ₂) ₁₂ -	-COO ⁻	29.4	> 250 ^c	ND
3.17	-(CH ₂) ₆ -	-CH ₂ COO ⁻	25.1	>500	ND
3.18	-(CH ₂) ₄ -	-(CH ₂) ₂ COO ⁻	23.9	>200	ND
3.19	-(CH ₂) ₈ -		36.6	>200	ND

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

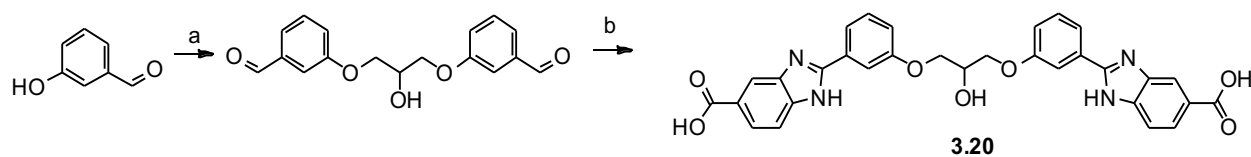
^b K_i was calculated from IC₅₀ values.

^c Inhibitors were tested at the highest possible concentration, considering constraints due to high absorbance or to precipitation.

ND, not determined.

Substitution of the benzimidazoles with 2-naphthoic acid (**3.19**) (Scheme 3-3) also abolished inhibition. Benzimidazole had been identified as a weak inhibitor of DfrB1 through fragment-based design, as had 2-naphthoic acid(19), yet in the context of these disubstituted symmetrical molecules their contribution to binding is not equivalent. The structural analogy of benzimidazole to the adenine moiety of NADPH could be the basis for its interactions with DfrB1. However, the adenine moiety in the holoenzyme (PDB ID 2RK1) is solvent exposed, establishing no interactions with DfrB1. Nonetheless, the adenine moiety is involved in binding as its removal decreases the affinity of NADPH(27). Notwithstanding the ambiguous role of the adenine moiety in NADPH binding, our observations establish that the benzimidazole moieties are significant contributors to DfrB1 inhibition. LowModeMD performed with **3.15** indicated interactions similar to those determined above although the terminal carboxylates formed hydrogen bonds more frequently with the YTT cluster than with the Lys32 network (Table A2-II; Table A2-III). In this sole case, no hydrogen bonding was observed between the terminal carboxylates and the Gly35 backbone amine. Finally, the phenoxy moieties made π -hydrogen bonds with the Ser65 β -methylene, the Val66 backbone amine and the aromatic C–H of Tyr69.

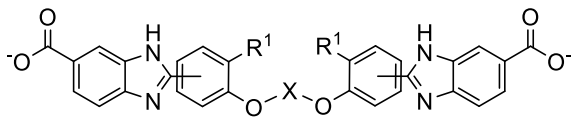
Substitution of the phenoxy moieties. We investigated the impact of the phenoxy moieties on binding by modulating their geometric constraints relative to the benzimidazole moieties. The *p*-orientation of **3.1** was modified to a *m*-orientation in **3.20** (Scheme 3-4). Despite the *m*-configuration reflecting the geometry of the flared edges of the active site tunnel, inhibition decreased (3-fold greater K_i) (Table 3-III).

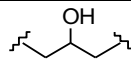
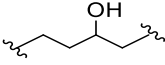
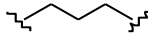


Scheme 3-4. Synthesis of 3.20.

Reagents and conditions: a) 1,3-dibromopropan-2-ol, EtOH/H₂O (9:1), NaOH, 120°C, 20 min.

b) Na₂S₂O₅, 3,4-diaminobenzoic acid, EtOH/H₂O (3:1), 140°C, 15 min.

Table 3-III. Inhibition upon substitution of the phenoxy moieties.^a

Compound	X	R ¹	Orientation of benzimidazole ^b	DfrB1		hDHFR	
				IC ₅₀ (μM)	K _i (μM) ^c	IC ₅₀ (μM) ^d	K _i (μM) ^c
3.20		-H	<i>m</i> -	400 ± 100	12 ± 3	>500	ND
3.21b		-Br	<i>p</i> -	80 ± 20	2.6 ± 0.7	>400	ND
3.22		-Br	<i>p</i> -	60 ± 20	1.7 ± 0.6	>400	ND

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

^b Orientation of benzimidazole relative to the O-linker.

^c K_i was calculated from IC₅₀ values.

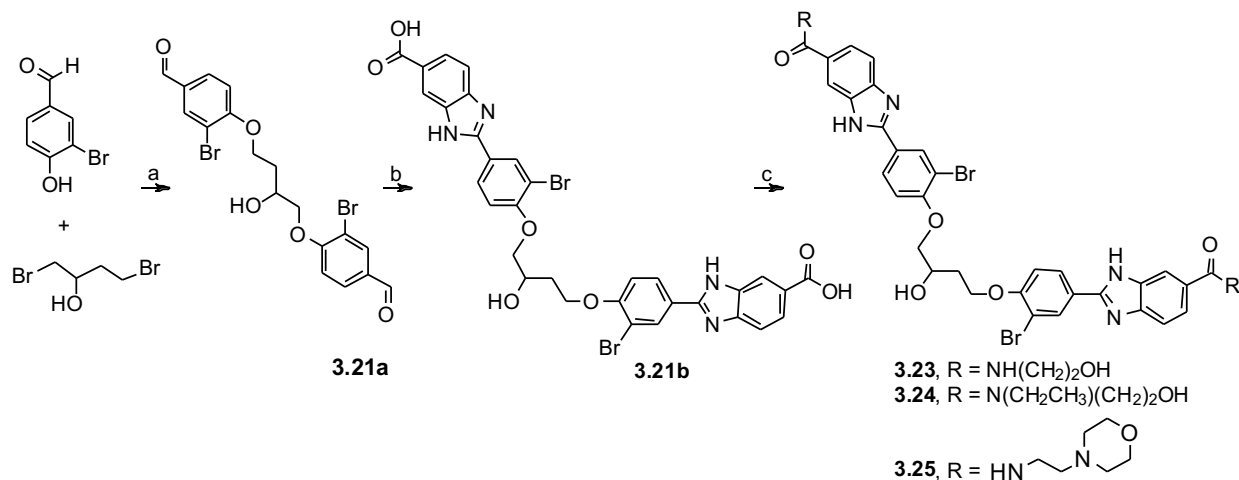
^d Inhibitors were tested at the highest possible concentration, considering constraints due to high absorbance.

ND, not determined.

LowModeMD with two molecules of **3.20** demonstrated conservation of the intertwined helical conformation. The *m*-configuration was not consistent with the hydrogen bonding observed with **3.1**, between the benzimidazole amines and the backbone amine and carbonyl of Val66 (Table A2-II; Table A2-III). Instead, the benzimidazole amines of **3.20** hydrogen bonded with the backbone carbonyl and amine of Ile68. The benzimidazole rings interacted through π -hydrogen bonds with the β -carbon hydrogen of Ser65 and the backbone amine of Val66 (Table A2-II; Table A2-III). Nearly all the binding contacts of the phenoxy moieties observed with **3.1** were abolished; only the hydrogen bond of the phenoxy oxygen with the Ile68 backbone amine was maintained. The altered hydrogen bonding may be reflected in the 3-fold greater K_i of **3.20** (12 μ M) than **3.1** (4.0 μ M).

Addition of halogens may enhance affinity with a target(37, 38). Bromination of the phenoxy ring was undertaken in an attempt to complement the large hydrophobic active site of DfrB1. A slight improvement of affinity was observed upon bromination of **3.3** (K_i = 3.5 μ M) to yield **3.21b** (K_i = 2.6 μ M) (Scheme 3-5). Furthermore, bromination of an analog of **3.2** with a

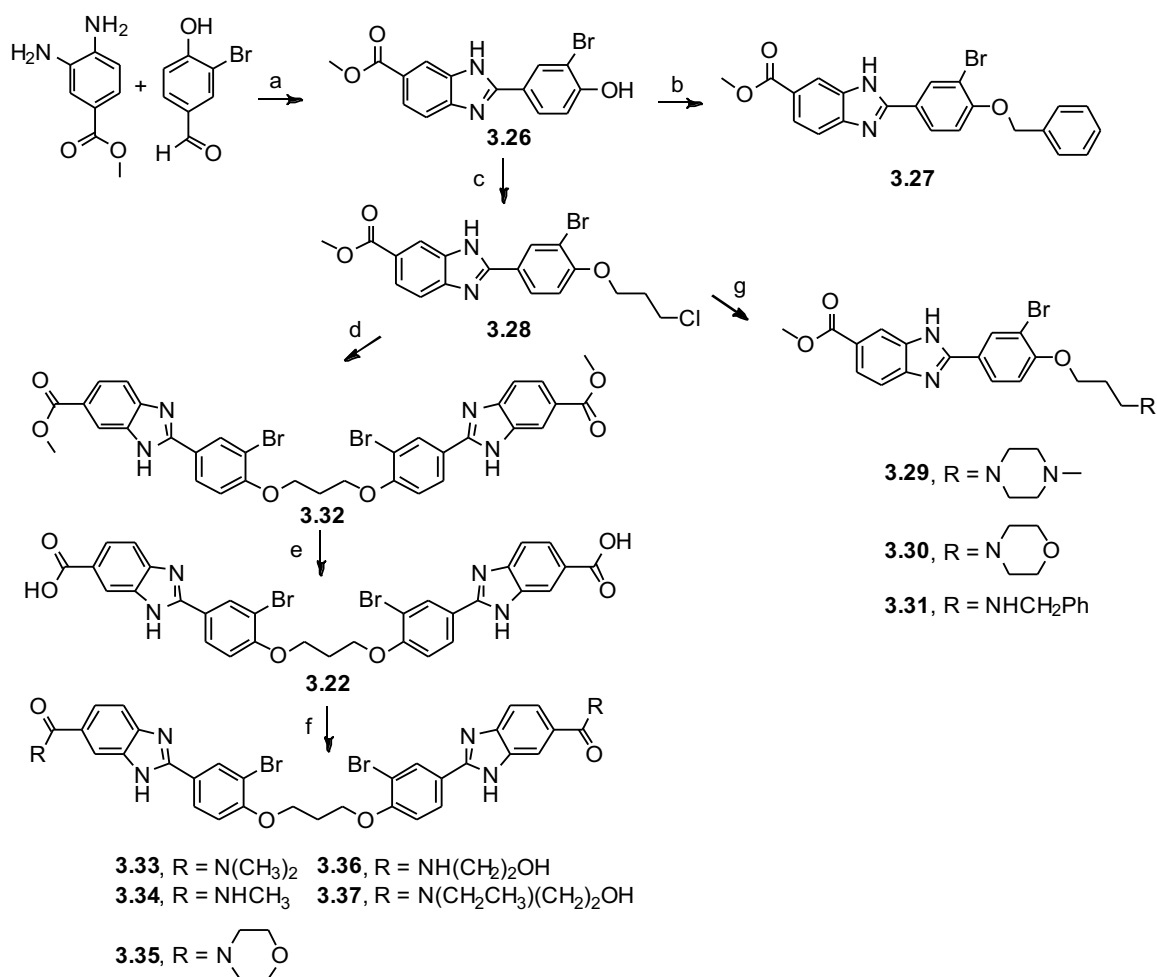
shorter ((CH₂)₃) central core to yield **3.22** (Scheme 3-6) also resulted in good affinity (K_i = 1.7 μM; Table 3-III), demonstrating the tolerance of DfrB1 to bromination.



Scheme 3-5. Synthesis of brominated 3.21b and its derivatives.

*Reagents and conditions: a) EtOH, K₂CO₃, reflux, 8h, 39% b) Na₂S₂O₅, 3,4-diaminobenzoic acid, EtOH/H₂O, 140°C, 15 min, 84% c) ethanolamine for **3.23**, 2-(ethylamino)ethanol for **3.24** or 2-morpholinoethanamine for **3.25**, HATU, DIPEA, DMF, rt, 3h, 50-92%.*

LowModeMD of **3.21b** and of **3.22** with DfrB1 suggests a hydrogen-bonding pattern similar to **3.2** and **3.3** (Table A2-II; Table A2-III), where the hydrogen bonds involving the phenoxy oxygens were substituted by halogen bonding. The phenoxy segment of **3.21b** formed electrostatic interactions between the bromine electropositive σ-hole and the oxygen lone pair electrons of the Val66 carbonyl and the Gln67 side-chain carbonyl. The bromine electropositive σ-hole in **3.22** created electrostatic interactions with the oxygen lone pair electrons of the Val66 and Ile68 carbonyls, and of the Ser65 hydroxyl. Those newly formed interactions did not significantly modify affinity.

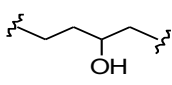
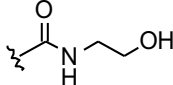
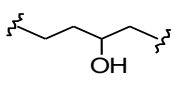
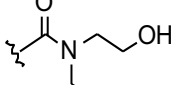
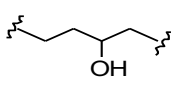
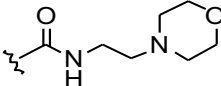



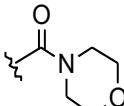
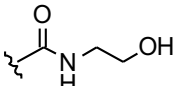
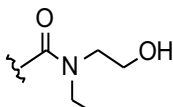


Scheme 3-6. Synthesis of methyl 2-(3-bromophenyl)benzimidazole-6-carboxylate 3.22 and its derivatives.

Reagents and conditions: a) DMF, Na₂S₂O₅, 80°C, 3h, 70%. b) Cs₂CO₃, acetonitrile, benzyl bromide, 70°C, 16h, 9%. c) Cs₂CO₃, acetone, 1-bromo-3-chloropropane, 70°C, 16h, 27%. d) 3.27, NaI, DMF, 90°C, 22h, 26-44%. e) NaOH/MeOH/THF (1:1:3), 70°C, 1h, 72%. f) methylamine for 3.33, dimethylamine for 3.34, morpholine for 3.35, ethanolamine for 3.36 or 2-(ethylamino)ethanol for 3.37, HATU, DIPEA, DMF, rt, 3h, 42-56%. g) 1-methylpiperazine for 3.29, morpholine for 3.30 or benzylamine for 3.31, Cs₂CO₃, NaI, DMF, 100°C, 20h, 17%.

Substituting the terminal carboxylates. All simulations presented above predicted that electrostatic interactions of the terminal carboxylates with the Lys32 ϵ -amino group and hydrogen bonds with backbone amines of Ser34 and of Gly35, were predominant (Table A2-II; Table A2-III). In an effort to better understand the role of the terminal carboxylates and to design compounds that would obviate potential issues associated with carboxylates within pharmacophores(39), we explored a greater diversity of terminal substituents by converting these acids into a series of amides. The introduction of neutral moieties at this position, with or without hydrogen-bonding substituents, was an effort to gain further insight into the structural requirements at the termini (Scheme 3-5 and Scheme 3-6). We verified the impact on affinity for DfrB1 of a series of bisbenzimidazolium amide derivatives (**3.23** to **3.25** and **3.32** to **3.37**; Table 3-IV). Their poor solubility precluded testing at concentrations greater than 50-100 μ M, as indicated. No inhibition was observed, supporting the hypothesis that the terminal carboxylates play an important role in binding to DfrB1. Despite their shortcomings within pharmacophores, they prove difficult to substitute in this context.

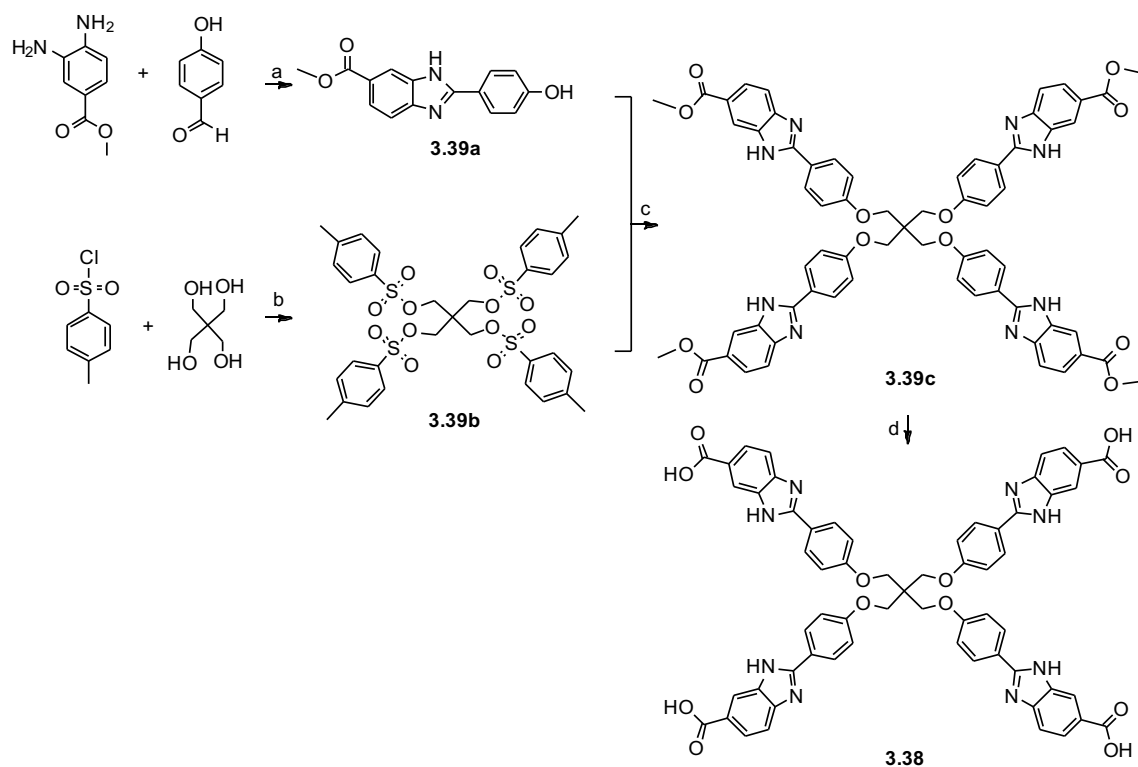
Table 3-IV. IC₅₀ upon substitution of the terminal carboxylates.

Compound	X	R ¹	IC ₅₀ (μM) ^a
3.23			>50
3.24			>50
3.25			>50
3.32	-(CH ₂) ₃ -		>30
3.33	-(CH ₂) ₃ -		>100
3.34	-(CH ₂) ₃ -		>100
3.35	-(CH ₂) ₃ -		>100
3.36	-(CH ₂) ₃ -		>50
3.37	-(CH ₂) ₃ -		>100

^a Compounds were tested at the highest possible concentration, considering constraints due to precipitation.

Reducing entropic cost by binding one molecule of inhibitor per active site. Enzyme kinetics with **3.1** and crystallography with **3.3** demonstrated that two molecules simultaneously occupy the DfrB1 active-site tunnel(19). In an attempt to reduce entropic cost of binding a dimer of bisbenzimidazole inhibitors, we designed a tetrabenzimidazole analog **3.38** (Scheme 3-7). It is

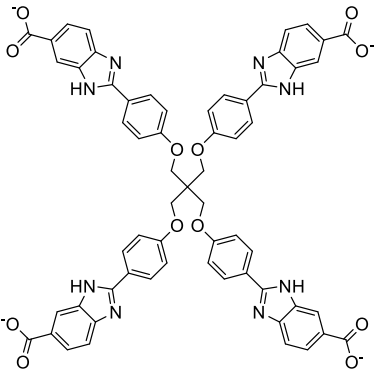
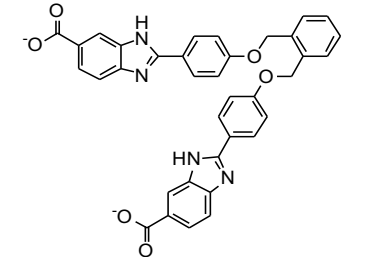
expected to display similar binding enthalpy and should reduce entropic loss by binding as a single copy. The inhibition potency of **3.38** was slightly improved (K_i 2- to 12-fold lower than the best inhibitors). A Hill coefficient of 1.1 ± 0.4 was determined, confirming the desired 1:1 binding stoichiometry (Figure 3-4). Its methyl ester analog, **3.39c**, did not inhibit, reinforcing the essential nature of the terminal carboxylates (Table A2-IV).



Scheme 3-7. Synthesis of 3.38 to 3.39c.

Reagents and conditions: a) $\text{Na}_2\text{S}_2\text{O}_5$, DMF, 80°C , 3h, 69%. b) pyridine, rt, 48h, 64%. c) DMF, K_2CO_3 , reflux, 24h, 23%. d) 1M NaOH/MeOH/THF (1:1:3), 70°C , 1h, 42%.

Table 3-V. Inhibition with 1:1 target:ligand stoichiometry.^a

Inhibitor	DfrB1		hDHFR		Hill coefficient
	IC ₅₀	K _i	IC ₅₀	K _i	
	(μ M)	(μ M) ^b	(μ M)	(μ M) ^b	
<p>3.38</p> 	35 ± 9	1.1 ± 0.3	ND	ND	1.1 ± 0.4
<p>3.12</p> 	240 ± 90	7.4 ± 2.7	>250	ND	1.1 ± 0.1

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

^b K_i was calculated from IC₅₀ values.

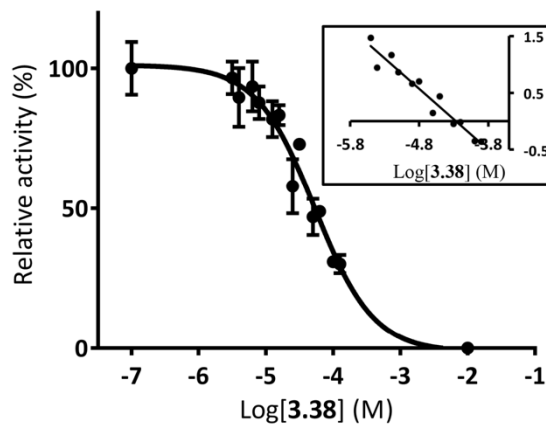
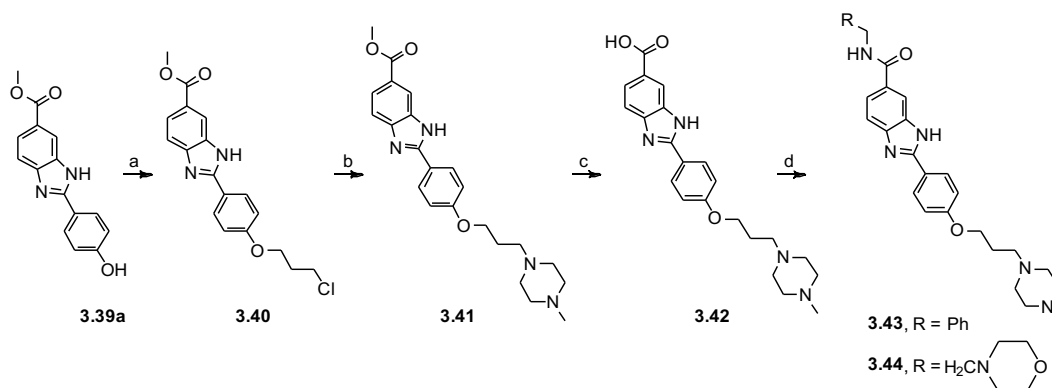


Figure 3-4. Inhibition of DfrB1 with monomeric inhibitor 3.38.

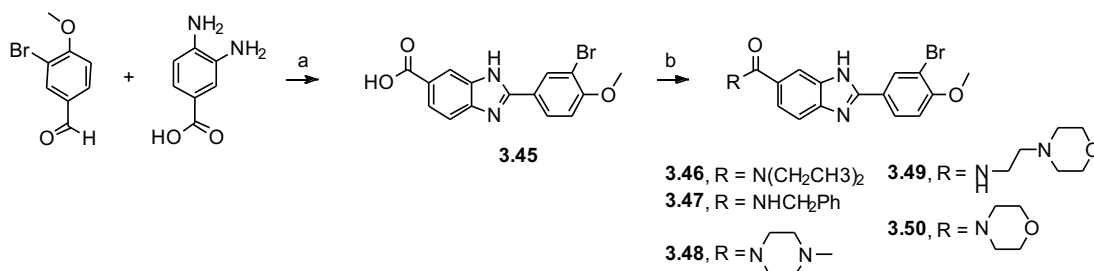
Inset: Hill plot of the inhibition curve below. The Hill coefficient is 1.1 ± 0.4.

Repeated attempts at crystallization of DfrB1 with **3.38** were unsuccessful. LowModeMD was executed with **3.38** (Figure A2-3; Table A2-II; Table A2-III). Electrostatic interactions, hydrogen and π -hydrogen bonding were all as those observed with **3.3**, apart from the congested quaternary central core. Only a weak hydrogen bond was modeled between a core methylene of **3.38** and the side-chain amide oxygen of Gln67.



Scheme 3-8. Synthesis of non-symmetrical phenylbenzimidazole-6-carboxylate compounds.

Reagents and conditions: a) Cs_2CO_3 , acetone, 1-bromo-3-chloropropane, 80°C , 16h, 33%. b) 1-methylpiperazine, Cs_2CO_3 , NaI, DMF, 100°C , 20h, 56% c) 1M NaOH/MeOH/THF (1:1:3), 70°C , 1h. d) benzylamine for **3.43** or 2-morpholinoethanamine for **3.44**, HATU, DIPEA, DMF, rt, 3h, 36%.



Scheme 3-9. Synthesis of non-symmetrical 2-phenylbenzimidazole compounds.

Reagents and conditions: a) anhydrous DMF, $\text{Na}_2\text{S}_2\text{O}_5$, 80°C , 3h. b) diethylamine hydrochloride for **3.46**, phenylmethanamine for **3.47** or 1-methylpiperazine for **3.48**, 2-morpholinoethanamine for **3.49** or morpholine for **3.50**, HATU, DIPEA, DMF, rt, 3h, 45-63%.

A second strategy to improve entropic cost with lower molecular weight than **3.38** involved non-symmetrical phenoxybenzimidazole-based compounds (Table 3-VI; Scheme 3-6, Scheme 3-8 and Scheme 3-9). They include the benzimidazole and phenoxy moieties that procure affinity (Table 3-II and Table 3-III) and take advantage of the tolerance to modification of the central core moiety (Table 3-I). However, none inhibited DfrB1 at the highest concentration where they were soluble (Table 3-VI).

Table 3-VI. Inhibition with non-symmetrical benzimidazole-based compounds.

#	R ¹	R ²	R ³	IC ₅₀ (μM) ^a
3.26	-OMe	-Br	-H	>30
3.27	-OMe	-Br		>30
3.39a	-OMe	-H	-H	>30
3.29	-OMe	-Br		>630
3.30	-OMe	-Br		>630
3.31	-OMe	-Br		>30
3.41	-OMe	-H		>500
3.43		-H		>500
3.44		-H		>500
3.45	-OH	-Br	-Me	>630
3.46		-Br	-Me	>160
3.47		-Br	-Me	>30
3.48		-Br	-Me	>2,500
3.49		-Br	-Me	>630
3.50		-Br	-Me	>630

^a Compounds were tested at the highest possible concentration, considering constraints due to precipitation.

Pursuing efforts to improve entropic cost with lower molecular weight, we designed a V-shaped analog of **3.11** based on an *o*-substituted central phenyl core (**3.12**) (Scheme 3-2). This configuration imitates half of **3.38** and was designed to fill half of the active-site tunnel and competitively prevent DfrB1 activity. It was originally inspired by results of a simulation where

an analog of **3.11** folded over itself to adopt a similar conformation(19). Binding a single molecule of **3.12** should reduce entropic cost yet would suffer from lower binding enthalpy due to loss of contacts. Upon verification, V-shaped **3.12** lost little affinity ($K_i = 7.4 \mu\text{M}$, 2-fold weaker than **3.1**) and was surprisingly similar to its linear, *p*-substituted analog, **3.11** ($K_i = 6.1 \mu\text{M}$). The Hill coefficient of 1.1 ± 0.1 confirmed the desired 1:1 binding stoichiometry of **3.12** (Figure 3-5A). Inhibitor **3.12** was selective, displaying no inhibition of hDHFR (Table 3-V).

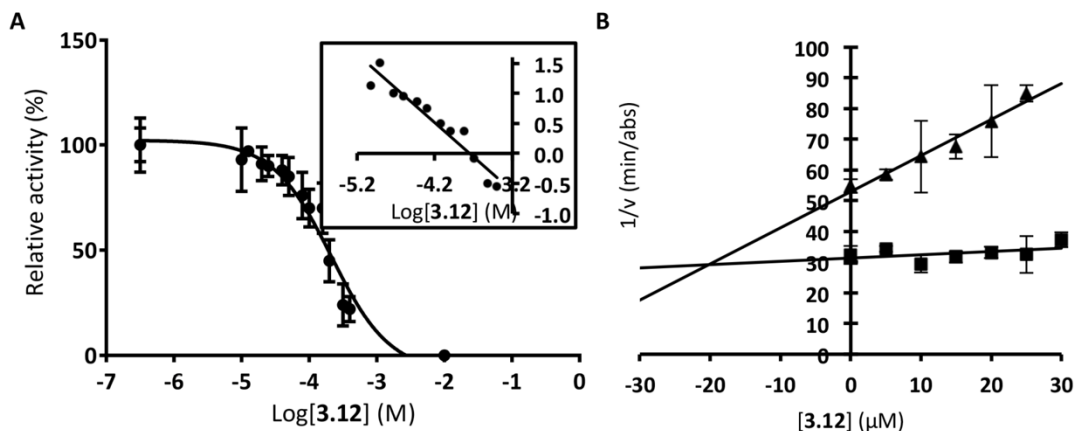


Figure 3-5. Panel A: Inhibition of DfrB1 with 3.12. Inset: Hill plot of panel A. The Hill coefficient is 1.1 ± 0.1 . Panel B: Dixon plot of 3.12 with DfrB1.

The reciprocal velocity of DfrB1 as a function of inhibitor concentration determines the type of inhibition. DHF was held constant at $50 \mu\text{M}$ and NADPH was $5 \mu\text{M}$ ($\sim 1 \times K_M^{\text{NADPH}}$; ▲) or $80 \mu\text{M}$ ($22 \times K_M^{\text{NADPH}}$; ■). The intercept of the slopes gave K_i . Values are given as the mean \pm standard deviation for triplicate results.

Dixon plot analysis of the inhibition of DfrB1 with **3.12** at different NADPH concentrations indicated a competitive inhibition mode (Figure 3-5B), with a $K_i = 20.1 \mu\text{M}$, in the same range as that calculated from IC_{50} ($7.4 \mu\text{M}$). This demonstrates that binding of **3.12** competes with NADPH binding. Dixon analysis upon varying DHF concentrations did not reflect a classical form of inhibition (Figure A2-4). Uncompetitive inhibition is expected to produce parallel lines with a non-zero slope while non-competitive inhibition gives intersecting lines with a x-intercept of $-K_i(40)$, neither of which corresponds to the observed pattern of nearly horizontal lines when varying DHF. This demonstrates that saturating NADPH precluded binding of subsaturating concentrations of **3.12**. The preferential binding of NADPH on the apo-DfrB1 ($K_{d1} = 2.5 \mu\text{M}$) relative to DHF ($K_{d1} = 120 \mu\text{M}$) justifies the greater impact of NADPH than DHF on binding of **3.12**(36).

LowModeMD simulations are consistent with one molecule of **3.12** preventing binding of NADPH or DHF to DfrB1 by occupying the active-site tunnel (Figure 3-6; Table A2-II; Table A2-III). As a result of its *o*-substituted central core, V-shaped **3.12** adopted an elongated, twisted conformation, precluding simultaneous binding of two molecules of **3.12**. Because a single molecule of **3.12** was bound, the frequency of binding contacts between **3.12** and DfrB1 was lower than for other inhibitors (Table A2-II; Table A2-III). The nature of the interactions was the same as for **3.3** except near their central cores. The phenyl core of **3.12** established π -hydrogen bonds with the alkyl side-chain of Gln67 and the Tyr69 hydroxyl group. Those weak interactions did not suffice to improve the affinity of **3.12**.

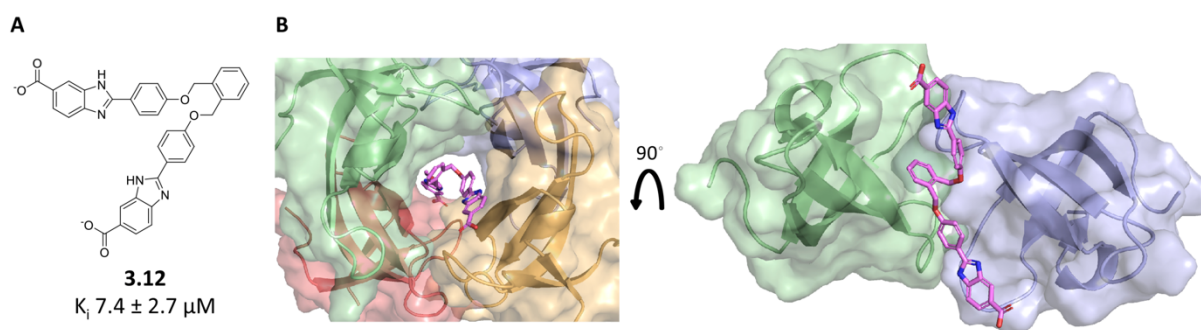


Figure 3-6. A) Structure of 3.12. B) *In silico* reconstruction of 3.12 in the active site of DfrB1 (left panel) and from a top view (right panel).

The benzimidazolium-based inhibitors are not aggregators. Various observations allow us to demonstrate that the DfrB1 inhibitors are not aggregation-based inhibitors(41). First, we have demonstrated selectivity, where DfrB1 inhibitors active between 1-10 μM do not significantly reduce hDfr activity at concentrations 10 to 100-fold greater (Tables 3-I, 3-III and 3-V). Second, inhibition curves displayed steepness consistent with soluble behaviour (Figure 3-4 and Figure 3-5), rather than the unusually steep pitch associated with aggregation-based inhibition. Third, the competitive mode of inhibition determined by observation of the bound crystal structures with **3.1**(19) or **3.3**, and kinetics with **3.1**(19) or **3.12**, is not characteristic of aggregators(41).

DfrB1 inhibitors inhibit the clinically-identified, TMP-resistant homolog DfrB4. DfrB4 shares 77% amino acid identity with DfrB1; the catalytic core is highly conserved while the loops and termini differ in 18 of their residues(20, 29, 35, 42). The key residues of DfrB1 responsible for binding of the benzimidazole-type inhibitors: the Lys32 network, the YTT cluster and the active site residues, are conserved in DfrB4, suggesting that DfrB4 may be inhibited by the same compounds. We previously identified the DfrB4 coding sequence, flanked by further

antibiotic resistance genes, in a TMP-resistant North American clinical sample(16). We further demonstrated that expression of DfrB4 in *E. coli* conferred complete resistance to the highest concentration of TMP that could be dissolved in the growth medium (600 µg/mL)(16). Here, we verified whether inhibitors of DfrB1 also inhibit DfrB4.

DfrB4 has not been biochemically characterized. We first confirmed that DfrB4 possesses dihydrofolate reductase activity. Its kinetic parameters K_m^{NADPH} and k_{cat} were similar to DfrB1, and K_m^{DHF} of DfrB4 was 3-fold greater (Table 3-VII). Both DfrBs have a catalytic efficiency of ≈ 100 -fold lower than *E. coli* chromosomal Dfr; although inefficient, this is sufficient to ensure bacterial host survival in presence of TMP(26). We determined their inhibitory constants for TMP. Both exhibit impressively low sensitivity to TMP (Table 3-VII)(27).

Table 3-VII. Kinetic parameters for DfrB1 and DfrB4.^a

	K_m^{DHF} (µM)	K_m^{NADPH} (µM)	k_{cat} (s ⁻¹)	IC ₅₀ ^{TMP} (mM)	K_i^{TMP} (mM) ^b
DfrB1	8.2 ± 0.11 ^c	1.6 ± 0.02 ^c	0.30 ± 0.04 ^c	19 ± 4	0.60 ± 0.14
DfrB4	2.6 ± 1.0	2.8 ± 0.6	0.45 ± 0.03	12 ± 2	0.55 ± 0.096

^a Values are given as the average ± standard deviation from the mean of at least duplicate of triplicate.

^b K_i was calculated from IC₅₀ values.

^c Values taken from Table 2 in (26).

Five efficient DfrB1 inhibitors (**3.1**, **3.3**, **3.12**, **3.21b** and **3.38**), representing different features, were assayed for inhibition of DfrB4 (Table 3-VIII, Table A2-V). The K_i values were similar for both DfrBs (1.1- to 1.7-fold difference). Thus, both DfrBs are similarly inhibited by the simplest inhibitors, **3.1** and **3.3**, by the V-shaped **3.12**, tolerate bromination of the phenoxy moiety (**3.21b**) and accept the bulky tetrabenzimidazole-based **3.38**.

Table 3-VIII. Inhibition of DfrB1 and DfrB4 with selected inhibitors.^a

	K_i (µM)				
	3.1	3.3	3.12	3.21b	3.38
DfrB1	4.0 ± 0.3 ^b	3.5 ± 1.7	7.4 ± 2.7	2.6 ± 0.7	1.1 ± 0.3
DfrB4	3.5 ± 1.1	3.9 ± 0.9	11 ± 1.3	3.7 ± 0.7	1.9 ± 0.5

^a K_i was calculated from IC₅₀ values. Values are given as the average ± standard deviation from the mean of at least duplicate of triplicate.

^b Values taken from Table 2 in (19).

The sequence differences between DfrB1 and DfrB4 have no significant impact on their catalytic activity, TMP resistance or capacity to be inhibited by the benzimidazole-based compounds tested here. Importantly, this immediately suggests that the inhibitors reported here also inhibit all other known members of the DfrB family, as they share 77-99% of protein sequence identity(16).

3.5 . CONCLUSION

Our report has uncovered several features that contribute significantly to the inhibition of DfrB1. The inclusion of terminal carboxylates is a key feature as their substitution abolished inhibition. They appear to establish frequent hydrogen bonds or electrostatic interactions with either the YTT cluster or the Lys32 network. The length of the inhibitor is a second key feature, to allow the above interactions to be established at both mouths of the active-site tunnel simultaneously. A third feature is the contribution of the benzimidazoles: although not crucial to binding, their impact is significant. A final feature is packing within the large, symmetrical active site: simultaneous binding of two linear bisbenzimidazole inhibitors could be substituted with one tetrabenzimidazole inhibitor (**3.38**), or, with little loss of affinity, by a single V-shaped bisbenzimidazole inhibitor (**3.12**) that was modelled in a twisted conformation. The weak contacts between the *o*-substituted central core of **3.12** and DfrB1 reveal its potential for future optimization. We have thus validated a series of features characterizing benzimidazole-based inhibitors that selectively inhibit the highly TMP-resistant DfrB1 and DfrB4 without affecting hDfr.

3.6 . ACKNOWLEDGMENTS

We thank Prof. Steven Laplante and Prof. Roberto Chica for technical support, Nadia Nouasria for technical assistance, the University of Montreal NMR laboratory and Yves Gareau for assistance in analyzing NMR spectra. We also thank Simon Toulouse for assistance with drafting Figure A2-2. We thank Pawel Grochulski at the Canadian Light Source (beamline CMCF-ID) for collecting the data for PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’.

This work was supported by Natural Sciences and Engineering Research Council of Canada grants 227853 and 2018-04686 awarded to J. N. P., Canada Foundation for Innovation

grant 11510 to J. N. P., CIHR operating grant MOP-13107 to A. M. B. and the Fondation Marcel et Rolande Gosselin to A. M. J. L. T. received scholarships from les Fonds Québécois pour la Recherche sur la Nature et les Technologies (FQRNT), from Hydro-Québec and Faculté des études supérieures et post-doctorales de l'Université de Montréal (FESP) and from PROTEO, the Québec Network for Research on Protein, Function, Engineering and Applications. B.J.Y. has held scholarships from the Natural Sciences and Engineering Research Council of Canada, McGill University, the CIHR Strategic Initiative in Chemical Biology, and the NSERC CREATE Program in Bionanomachines. D. B. received scholarship from PROTEO.

3.7 . EXPERIMENTAL PROCEDURES

Materials. Dihydrofolate was synthesized as previously reported(43). Unless otherwise mentioned, DNA-modifying enzymes were purchased from Thermo Scientific (Waltham, MA, USA). Enterokinase was from Feldan (Québec, QC) or from PROSPEC (USA). Isopropyl b-D-1-thiogalactopyranoside and other reagents for enzyme assays were from BioShop (Burlington, ON, Canada). DMSO was purchased from Fisher Scientific (Ottawa, ON, Canada), and β -NADPH from Alexis biochemicals (San Diego, CA).

General procedure A. As previously described(44), the selected hydroxybenzaldehyde (2.0 eq) and the appropriate dibrominated compound (1.0 eq) in a mixture of EtOH/H₂O (9:1) containing NaOH (2.0 eq) was heated under microwave irradiation in an Initiator® Biotage oven for 20 min at 120°C. After cooling, the precipitate was filtered, washed with EtOH/H₂O and dried in vacuo. The resulting bisbenzaldehydes were sufficiently pure to proceed to the next step.

General procedure B. A mixture of the resulting bisbenzaldehyde (1.0 eq), Na₂S₂O₅ (1.0 eq), 3,4-diaminobenzoic acid (2.0 eq), and EtOH/H₂O (3:1) was microwave irradiated (Biotage) for 15 min at 140°C. After cooling, the precipitate was filtered, thoroughly washed with EtOH/H₂O/ether and dried in vacuo.

General procedure C. As previously described(44), the selected phenol derivative (3.0 eq) and the appropriate dibromoalkane (1.0 eq) were dissolved in a mixture of EtOH/H₂O (9:1) containing NaOH (6.0 eq). The mixture was heated under microwave irradiation in an Initiator® Biotage oven for 20 min at 120°C. After cooling, the precipitate was acidified with 6 mL of 37% HCl and thoroughly washed with water and ethanol.

General procedure D. To a solution of the selected mono- or bis-acid (1.0 eq), HATU (3.0 eq) and DIPEA (10 eq) stirred in DMF (0.1 M) at room temperature, the amine (2.2 eq) was added and the resulting reaction mixture was stirred at room temperature for 3 h. The resulting mixture was then diluted with ethyl acetate, washed with water and brine, dried over MgSO₄, filtered and concentrated. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the desired product.

General procedure E. To a solution of the selected acid (1.0 eq) and amine (1.1 eq) in DMF (0.1 M) were added DIPEA (5 eq) and HATU (1.5 eq) at room temperature and the resulting mixture was stirred for 3 h. The resulting mixture was then diluted with ethyl acetate, washed (water, brine), dried (MgSO₄), filtered and concentrated. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound.

General procedure F. A vial was charged with methyl 2-(3-bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.28**) (1.0 eq), cesium carbonate (3.0 eq), sodium iodide (1.0 eq), the selected amine (3.0 eq) and DMF (0.5 mL). The resulting mixture was heated at 100°C for 20 h. The cooled mixture was filtered, and the filtrate was directly purified by prep HPLC (Preparative HPLC- General method A) to give the title compound.

Analytical HPLC- General method A. Analytical HPLC analyses were performed on a Agilent 1200 Series LC system with a Zorbax XDB-C18 column (4.6 x 30 mm, 3.5 μM) under the following conditions: flow rate 3 mL/min at room temperature; 10 μL of target compound solution in MeOH; solvent A 5% MeOH, 95% Water + 0.05% TFA; solvent B 95% MeOH, 5% Water + 0.05% TFA; 2 minutes gradient from 0 to 100% solvent B and 2 minutes 100% solvent B; detection $\lambda = 220/254$ nm. The purity was determined by HPLC with UV detection at 220 nm.

Analytical HPLC- General method B. Analytical HPLC analyses were also achieved on a Agilent 1200 Series LC system with a Kinetex C18 column (3.0 x 30 mm, 2.6 μM) under the following conditions: flow rate 1.5 mL/min at room temperature; 2 μL of target compound solution in MeOH; solvent A 5% MeOH, 95% water + 0.1% AcOH; solvent B 95% MeOH, 5% water + 0.1% AcOH; 0.5 minute gradient from 0 to 100% solvent B and 1.5 minute 100% solvent B; detection $\lambda = 220/254$ nm. The purity was determined by HPLC with UV detection at 220 nm.

Preparative HPLC- General method A. Preparative HPLC analyses were realised on a Agilent 1200 Series LC system with a Kinetex C18 column (21.2 x 100 mm, 5 μM) under the following conditions: 20 mL/min at room temperature; 900 μL of 25 mg/mL of compound in DMF; solvent

A 5% MeOH, 95% water + 0.05% TFA; solvent B 95% MeOH, 5% water + 0.05% TFA; 2 min isocratic 30% solvent B followed by 8 minutes gradient to 100% solvent B and 3 minutes 100% solvent B; detection $\lambda = 254$ nm.

2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.3). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,4-dibromo-2-butanol (1.73 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid (0.91 g, 2.9 mmol, 19%). ¹H-NMR (700 MHz, DMSO) δ ppm 13.08 (s, 2H), 12.49 (br, 1H), 8.20 (br, 1H), 8.15 (m, 4H), 8.10 (br, 1H), 7.84 (d, $J = 5.8$ Hz, 2H), 7.66 (br, 1H), 7.56 (br, 1H), 7.16 (m, 4H), 5.02 (s, 1H), 4.29 (t, $J = 6.6$ Hz, 2H), 4.11 (m, 3H), 2.12 (m, 1H), 1.95 (m, 1H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.05, 160.97, 154.7 (br), 153.8 (br), 147.8 (br), 144.1 (br), 138.8 (br), 135.0 (br), 128.9, 124.9, 124.0 (br), 123.5 (br), 122.8, 122.7, 120.7 (br), 118.5 (br), 115.6, 115.5, 113.2 (br), 111.2 (br), 73.0, 66.2, 65.2, 33.7. Purity > 95%.

2,2'-((Ethane-1,2-diyl)bis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.4). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,2-dibromoethane (1.40 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a white solid. ¹H-NMR (700 MHz, DMSO) δ ppm 12.92 (br), 8.17 (m, 4H), 8.14 (br, 2H), 7.83 (d, $J = 8.4$ Hz, 2H), 7.62 (br, 1H), 7.59 (br, 1H), 7.21 (m, 4H), 4.50 (s, 4H). ¹³C-NMR (176 MHz, DMSO) $\delta = 168.4, 160.6, 154.0, 128.9, 125.6$ (br), 125.4 (br), 123.8, 123.2, 115.6, 67.2. Purity 93%.

2,2'-((Butane-1,4-diyl)bis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.5). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,4-dibromobutane (1.60 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid. ¹H-NMR (700 MHz, DMSO) δ ppm 12.89 (s, 2H), 12.44 (br, 2H), 8.21 (br, 1H), 8.14 (d, $J = 8.1$ Hz, 4H), 8.10 (br, 1H), 7.82 (d, $J = 5.0$ Hz, 2H), 7.66 (br, 1H), 7.56 (br, 1H), 7.15 (d, $J = 8.1$ Hz, 4H), 4.19 (t, $J = 5.4$ Hz, 4H), 1.97 (t, $J = 5.4$ Hz, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.0, 154.7 (br), 153.8 (br), 147.8 (br), 144.1 (br), 138.9 (br), 135.2 (br), 128.9, 124.9, 124.0 (br), 123.6 (br), 122.7, 120.7 (br), 118.5 (br), 115.6, 113.2 (br), 111.2 (br), 68.1, 25.9. Purity > 95%.

2,2'-((Hexane-1,6-diyl)bis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.6). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,6-dibromohexane (1.82 g, 7.5 mmol)

were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid. ¹H-NMR (700 MHz, DMSO) δ ppm 12.89 (br, 2H), 12.38 (br, 2H), 8.35 (br, 2H), 8.13 (d, *J* = 8.3 Hz, 4H), 7.82 (d, *J* = 8.0 Hz, 2H), 7.60 (br, 2H), 7.13 (d, *J* = 8.3 Hz, 4H), 4.12 (t, *J* = 6.3 Hz, 4H), 1.82 (m, 4H), 1.56 (t, *J* = 6.9 Hz, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.0, 154.5 (br), 153.9 (br), 128.9, 125.0, 123.8 (br), 122.6, 115.5, 68.3, 29.1, 25.7. Purity > 95%.

2,2'-((Octane-1,8-diylbis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.7). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,8-dibromooctane (2.03 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a yellow solid. ¹H-NMR (700 MHz, DMSO) δ ppm 13.07 (br, 1H), 13.04 (br, 1H), 12.68 (br, 2H), 8.20 (br, 1H), 8.13 (m, 4H), 8.09 (br, 1H), 7.83 (br, 2H), 7.66 (br, 1H), 7.55 (br, 1H), 7.12 (m, 4H), 4.10 (t, *J* = 6.5 Hz, 4H), 1.78 (m, 4H), 1.48 (m, 4H), 1.41 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.0, 154.7 (br), 153.8 (br), 147.8 (br), 144.1 (br), 138.8 (br), 135.2 (br), 128.9, 124.9, 124.0 (br), 123.4 (br), 122.6, 120.7 (br), 118.5 (br), 115.5, 113.2 (br), 111.2 (br), 68.4, 29.13, 29.10, 25.9. Purity > 95%.

2,2'-((Decane-1,10-diylbis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.8). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,10-dibromodecane (2.24 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid. ¹H-NMR (700 MHz, DMSO) δ ppm 13.06 (br, 1H), 13.02 (br, 1H), 12.64 (br, 2H), 8.21 (br, 1H), 8.12 (m, 4H), 8.09 (br, 1H), 7.82 (br, 2H), 7.66 (br, 1H), 7.56 (br, 1H), 7.11 (m, 4H), 4.08 (t, *J* = 6.5 Hz, 4H), 1.77 (m, 4H), 1.45 (m, 4H), 1.38 (m, 4H), 1.34 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.0, 154.7 (br), 153.8 (br), 147.8 (br), 144.2 (br), 138.9 (br), 135.2 (br), 128.9, 124.9, 124.0 (br), 123.5 (br), 122.6, 120.7 (br), 118.5 (br), 115.5, 113.2 (br), 111.2 (br), 68.4, 29.3, 29.2, 29.1, 25.9. Purity > 95%.

2,2'-((Dodecane-1,12-diylbis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.9). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,10-dibromododecane (2.45 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid. ¹H-NMR (700 MHz, DMSO) δ ppm 13.00 (br, 1H), 12.68 (br, 2H), 8.16 (br, 2H), 8.12 (m, 4H), 7.82 (dd, *J* = 1.4, 8.3 Hz, 2H), 7.62 (br, 1H), 7.59 (br, 1H), 7.11 (m, 4H), 4.07 (t, *J* = 6.5 Hz, 4H), 1.75 (m, 4H), 1.45 (m, 4H), 1.36 (m, 4H), 1.31 (m,

8H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 161.0, 154.2, 128.9, 124.9, 123.7 (br), 122.6, 115.5, 68.3, 29.39, 29.37, 29.2, 29.1, 25.9. Purity > 95%.

(E)-2,2'-((But-2-ene-1,4-diylbis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.10). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and *trans*-1,4-dibromo-2-butene (1.60 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid (1.55 g, 5.3 mmol, 35%). ¹H-NMR (700 MHz, DMSO) δ ppm 13.17 (br, 2H), 8.19 (br, 1H), 8.14 (m, 4H), 8.09 (br, 1H), 7.83 (d, *J* = 5.9 Hz, 2H), 7.65 (br, 1H), 7.59 (br, 1H), 7.17 (m, 4H), 6.15 (t, *J* = 2.7 Hz, 2H), 4.74 (d, *J* = 2.7 Hz, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.4, 160.4, 154.7 (br), 153.9 (br), 147.8 (br), 144.2 (br), 138.6 (br), 135.2 (br), 128.9 (2C), 124.7, 124.1 (br), 123.4 (br), 122.7, 120.8 (br), 118.6 (br), 115.6, 113.4 (br), 111.3 (br), 67.9. Purity > 94%.

2,2'-(((1,4-Phenylenebis(methylene))bis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.11). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,4-bis(bromomethyl)benzene (1.97 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a light brown solid (1.92 g, 5.6 mmol, 37%). ¹H-NMR (700 MHz, DMSO) δ ppm 12.91 (br, 2H), 12.18 (br, 2H), 8.18 (br, 2H), 8.15 (m, 4H), 7.82 (d, *J* = 8.3 Hz, 2H), 7.61 (br, 2H), 7.54 (s, 4H), 7.23 (m, 4H), 5.25 (s, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 160.6, 137.0, 128.9, 128.3, 124.9, 123.9 (br), 123.0, 115.9, 69.8. Purity > 94%.

2,2'-(((1,2-phenylenebis(methylene))bis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.12). 4-Hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,3-bis(bromomethyl)benzene (1.97 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a brown solid (0.79 g, 6.5 mmol, 43%). ¹H-NMR (700 MHz, DMSO) δ ppm 13.16 (br, 2H), 8.15 (m, 4H), 8.12 (br, 2H), 7.81 (d, *J* = 8.5 Hz, 2H), 7.63 (br, 2H), 7.60 (m, 2H), 7.43 (m, 2H), 7.26 (m, 4H), 5.38 (s, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 168.3, 160.6, 154.6 (br), 153.7 (br), 135.5, 129.2, 128.9, 128.7, 125.0, 124.0 (br), 123.5 (br), 123.1, 120.8 (br), 118.5 (br), 115.9, 68.0. Purity > 95%.

4,4'-(hexane-1,6-diylbis(oxy))dibenzoic acid (3.13). Synthesis of **3.13** was done as described previously(45). ¹H-NMR (700 MHz, DMSO) δ ppm 12.59 (br, 2H), 7.87 (m, 4H), 7.01 (m, 4H), 4.05 (t, *J* = 6.5 Hz, 4H), 1.76 (m, 4H), 1.49 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 167.5, 162.8, 131.8, 123.3, 114.7, 68.2, 29.0, 25.7.

4,4'-(octane-1,8-diylbis(oxy))dibenzoic acid (3.14). Synthesis of **3.14** was done as described previously(45). ¹H-NMR (700 MHz, DMSO) δ ppm 12.59 (br, 2H), 7.88 (m, 4H), 7.00 (m, 4H), 4.04 (t, *J* = 6.5 Hz, 4H), 1.73 (m, 4H), 1.43 (m, 4H), 1.36 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 167.5, 162.8, 131.8, 123.2, 114.7, 68.2, 29.1, 29.0, 25.9.

4,4'-(Decane-1,10-diylbis(oxy))dibenzoic acid (3.15). Synthesis of **3.15** was done as described previously(45). ¹H-NMR (700 MHz, DMSO) δ ppm 12.37 (br, 2H), 7.88 (m, 4H), 7.00 (m, 4H), 4.05 (t, *J* = 6.4 Hz, 4H), 1.73 (m, 4H), 1.43 (m, 4H), 1.35 (m, 4H), 1.30 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 167.4, 162.8, 131.8, 123.5, 114.8, 68.4, 29.3, 29.1, 29.0, 25.9. Purity > 95%.

4,4'-(Dodecane-1,12-diylbis(oxy))dibenzoic acid (3.16). Synthesis of **3.16** was done as described previously(45). ¹H-NMR (700 MHz, DMSO) δ ppm 12.60 (br, 2H), 7.88 (m, 4H), 6.99 (m, 4H), 4.05 (t, *J* = 6.5 Hz, 4H), 1.73 (m, 4H), 1.42 (m, 4H), 1.34 (m, 4H), 1.29 (m, 8H). ¹³C-NMR (176 MHz, DMSO) δ ppm 167.4, 162.8, 131.8, 123.5, 114.8, 68.4, 29.35, 29.32, 29.1, 29.0, 25.9. Purity > 95%.

2,2'-((hexane-1,6-diylbis(oxy))bis(4,1-phenylene))diacetic acid (3.17). Synthesis of **3.17** was done as described previously(45). Briefly, 4-hydroxyphenylacetic acid (0.76 g, 5 mmol) and 1,6-dibromohexan-2-ol (0.65 g, 2.5 mmol) were reacted according to General procedure C. The purified product was a white solid (0.17 g, 0.45 mmol, 9%). ¹H-NMR (700 MHz, DMSO) δ ppm 12.24 (br, 2H), 7.15 (m, 4H), 6.86 (m, 4H), 3.94 (t, *J* = 6.4 Hz, 4H), 3.48 (s, 4H), 1.73 (m, 4H), 1.47 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 173.5, 157.9, 130.8, 127.3, 114.7, 67.8, 40.3, 29.1, 25.8. Purity > 95%.

3,3'-((butane-1,4-diylbis(oxy))bis(4,1-phenylene))dipropionic acid (3.18). Synthesis of **3.18** was done as described previously(45). Briefly, 3-(4-hydroxyphenyl)propanoic acid (0.83 g, 5 mmol) and 1,4-dibromobutan-2-ol (0.58 g, 2.5 mmol) were reacted according to General procedure C. The purified product was a white solid (0.15 g, 0.4 mmol, 8%). ¹H-NMR (700 MHz, DMSO) δ ppm 12.08 (br, 2H), 7.12 (m, 4H), 6.84 (m, 4H), 3.98 (m, 4H), 2.75 (t, *J* = 7.6 Hz, 4H), 2.48 (t, *J* = 7.6 Hz, 4H), 1.84 (m, 4H). ¹³C-NMR (176 MHz, DMSO) δ ppm 174.3, 157.4, 133.2, 129.7, 114.8, 67.5, 36.1, 30.0, 26.0. Purity ≈ 82%.

6,6'-((octane-1,8-diylbis(oxy))bis(4,1-phenylene))bis(2-naphthoic acid) (3.19). Synthesis of **3.19** was done as described previously(45). Briefly, 6-hydroxy-2-naphthoic acid (0.94 g, 5 mmol) and 1,8-dibromooctan-2-ol (0.72 g, 2.5 mmol) were reacted according to General procedure C. The purified product was a white solid (0.34 g, 0.5 mmol, 10%). ¹H-NMR (700 MHz, DMSO) δ

ppm 12.87 (br, 2H), 8.51 (s, 2H), 8.01 (d, $J = 9.0$ Hz, 2H), 7.92 (d, $J = 8.4$ Hz, 2H), 7.87 (d, $J = 8.4$ Hz, 2H), 7.40 (s, 2H), 7.23 (d, $J = 8.4$ Hz, 2H), 4.13 (t, $J = 6.2$ Hz, 4H), 1.81 (m, 4H), 1.49 (m, 4H), 1.41 (m, 4H). ^{13}C -NMR (176 MHz, DMSO) δ ppm 168.0, 159.0, 137.2, 131.3, 130.8, 127.9, 127.3, 126.2, 126.1, 120.1, 107.1, 68.2, 29.2, 29.0, 26.0. Purity > 95%.

2,2'-(((2-hydroxypropane-1,3-diyl)bis(oxy))bis(3,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (3.20). 3-hydroxybenzaldehyde (1.83 g, 15 mmol) and 1,3-dibromopropan-2-ol (1.62 g, 7.5 mmol) were reacted according to General procedure A followed by General procedure B. The purified product was a grey solid. ^1H -NMR (700 MHz, DMSO) δ ppm 13.01 (br, 2H), 8.20 (br, 2H), 7.85 (m, 4H), 7.80 (d, $J = 7.9$ Hz, 2H), 7.65 (br, 2H), 7.49 (dd, $J = 7.9, 7.9$ Hz, 2H), 7.17 (d, $J = 7.9$ Hz, 2H), 4.30 (m, 3H), 4.24 (m, 2H). ^{13}C -NMR (176 MHz, DMSO) δ ppm 168.5, 159.7, 131.6, 130.6, 119.8, 117.5, 113.3, 70.3, 68.2; 7. Purity = 90%.

4,4'-((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromobenzaldehyde) (3.21a). A solution of 3-bromo-4-hydroxybenzaldehyde (1.820 g, 9.06 mmol), 1,4-dibromobutan-2-ol (1.00 g, 4.31 mmol) and potassium carbonate (0.596 g, 4.31 mmol) in ethanol (11 mL) was stirred while being heated to reflux for 8 h. The cooled mixture was filtered, and the filter-cake was washed with water, ethanol and diethyl ether and then dried in vacuo to give the title compound as a white solid (0.800 g, 1.69 mmol, 39%). The product obtained was used as such in the next step without further characterization.

2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.21b). A mixture of 4,4'-((2-hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromobenzaldehyde) (**3.21a**) (0.15 g, 0.318 mmol), 3,4-diaminobenzoic acid (0.102 g, 0.667 mmol), sodium metabisulfite (0.060 g, 0.318 mmol), water (0.315 mL, 17.47 mmol) and ethanol (0.835 mL, 14.30 mmol) was heated at 140°C in a microwave reactor for 15 min. The cooled mixture was filtered, and the filter-cake was washed with water, ethanol and diethyl ether and then dried in vacuo to give the title compound as a white solid (0.196 g, 0.266 mmol, 84%). LC (Analytical HPLC- General method A): 1.890 min. HRMS (ESI): calcd for $\text{C}_{32}\text{H}_{25}\text{Br}_2\text{N}_4\text{O}_7$ $[\text{M}+\text{H}]^+$ m/z 735.0085, found 735.0077. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 13.15 (br s, 2H), 8.40 (t, $J = 1.76$ Hz, 2H), 8.02 - 8.26 (m, 4H), 7.83 (d, $J = 8.61$ Hz, 2H), 7.50 - 7.74 (m, 1H), 7.30 - 7.41 (m, 1H), 5.21 (br s, 1H), 4.28 - 4.43 (m, 2H), 4.10 - 4.25 (m, 3H), 2.13 - 2.28 (m, 1H), 1.99 (dt, $J = 4.70, 8.41$ Hz, 1H). Purity = 88%.

2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22). A mixture of dimethyl 2,2'-((propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylate) (**3.32**) (0.100 g, 0.136 mmol), 1 M aqueous NaOH (2.18 mL, 2.18 mmol), MeOH (2.18 mL) and THF (6.54 mL) was heated at 70°C for 1 h. The cooled mixture was then concentrated to dryness and the residue was treated with a 1 M HCl solution. The solid present was collected by filtration, washed with water and dried under high vacuum to give the title compound as a white solid (0.069 g, 0.098 mmol, 72% yield). LC (Analytical HPLC- General method B): 1.329 min. HRMS (ESI): calcd for C₃₁H₂₃Br₂N₄O₆ [M+H]⁺ *m/z* 706.9961, found 706.9967. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 12.92 (br s, 2h), 8.39 - 8.53 (m, 2H), 8.09 - 8.33 (m, 4H), 7.90 (dd, *J* = 1.57, 8.61 Hz, 2H), 7.70 (d, *J* = 8.61 Hz, 2H), 7.45 (d, *J* = 8.61 Hz, 2H), 4.32 - 4.53 (m, 4H), 2.34 (dd, *J* = 5.48, 6.26 Hz, 3H). Purity = 85%.

2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(N-(2-hydroxyethyl)-1H-benzo[d]imidazole-6-carboxamide) (3.23). 2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (**3.21b**) (0.025 g, 0.034 mmol) and ethanolamine (4.52 μl, 0.075 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.016 g, 0.019 mmol, 57%). LC (Analytical HPLC- General method A): 1.720 min. HRMS (ESI): calcd for C₃₆H₃₅Br₂N₆O₇ [M+H]⁺ *m/z* 821.0928, found 821.0925. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.49 (t, *J* = 5.48 Hz, 2H), 8.42 (t, *J* = 1.76 Hz, 2H), 8.19 (td, *J* = 2.74, 8.61 Hz, 2H), 8.13 (s, 2H), 7.82 (d, *J* = 8.61 Hz, 2H), 7.65 (d, *J* = 8.61 Hz, 2H), 7.40 (dd, *J* = 3.72, 8.80 Hz, 2H), 4.33 - 4.42 (m, 2H), 4.13 - 4.25 (m, 3H), 3.51 - 3.58 (m, 4H), 3.37 (q, *J* = 6.26 Hz, 4H), 2.17 - 2.26 (m, 1H), 1.95 - 2.05 (m, 1H). Purity > 99%.

2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(N-ethyl-N-(2-hydroxyethyl)-1H-benzo[d]imidazole-6-carboxamide) (3.24). 2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (**3.21b**) (0.025 g, 0.034 mmol) and 2-(ethylamino)ethanol (7.24 μl, 0.075 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.015 g, 0.017 mmol, 50%). LC (Analytical HPLC- General method A): 1.773 min. HRMS (ESI): calcd for C₄₀H₄₃Br₂N₆O₇ [M+H]⁺ *m/z* 877.1554, found 877.1562. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.41 (t, *J* = 1.96

Hz, 2H), 8.18 (td, $J = 2.54, 8.61$ Hz, 2H), 7.58 - 7.69 (m, 4H), 7.40 (dd, $J = 3.72, 8.80$ Hz, 2H), 7.29 (d, $J = 8.61$ Hz, 2H), 4.32 - 4.42 (m, 2H), 4.14 - 4.26 (m, 3H), 3.34 (br s, 12H), 2.14 - 2.29 (m, 1H), 1.90 - 2.07 (m, 1H), 1.14 (br s, 6H). Purity > 99%.

(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-

benzo[d]imidazole-2,6-diyl))bis(morpholinomethanone) (3.25). 2,2'-(((2-Hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (**3.21b**) (0.025 g, 0.034 mmol) and 2-morpholinoethanamine (9.80 μ l, 0.075 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.030 g, 0.031 mmol, 92%). LC (Analytical HPLC- General method A): 1.646 min. HRMS (ESI): calcd for $C_{44}H_{49}Br_2N_8O_7$ $[M+H]^+$ m/z 959.2085, found 959.2094. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.75 (t, $J = 5.67$ Hz, 2H), 8.41 (t, $J = 1.96$ Hz, 2H), 8.19 (td, $J = 2.74, 8.61$ Hz, 2H), 8.14 (s, 2H), 7.79 (d, $J = 9.39$ Hz, 2H), 7.66 (d, $J = 8.61$ Hz, 2H), 7.38 (dd, $J = 2.93, 8.80$ Hz, 2H), 4.36 (t, $J = 6.65$ Hz, 2H), 4.12 - 4.25 (m, 3H), 3.95 - 4.09 (m, 4H), 3.53 - 3.74 (m, 12H), 3.30 - 3.41 (m, 4H), 3.08 - 3.24 (m, 4H), 2.16 - 2.28 (m, 1H), 1.93 - 2.07 (m, 1H). Purity > 99%.

Methyl 2-(3-bromo-4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (3.26). A 100 mL sealed tube was charged with methyl 3,4-diaminobenzoate (1.901 g, 11.44 mmol), 3-bromo-4-hydroxybenzaldehyde (2.3 g, 11.44 mmol) and DMF (48.7 mL). Sodium metabisulfite (2.175 g, 11.44 mmol) was then added and the resulting mixture was heated at 80°C for 3 h. After cooling, the reaction mixture was concentrated to dryness and the solid residue obtained was suspended in water-acetonitrile (1:1) and sonicated for 30 min. The resulting suspension was filtered, and the residue was dried in vacuo to give the title compound (3.097 g, 8.03 mmol, 70% yield) as a yellow solid. LC (Analytical HPLC- General method A): 1.686 min. HRMS (ESI): calcd for $C_{15}H_{12}BrN_2O_3$ $[M+H]^+$ m/z 347.0026, found 347.0035. 1H NMR (400 MHz, DMSO- d_6) δ ppm 11.01 (br s, 1H), 8.34 (d, $J = 2.35$ Hz, 1H), 8.15 (d, $J = 1.57$ Hz, 1H), 8.04 (dd, $J = 2.15, 8.41$ Hz, 1H), 7.85 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.65 (d, $J = 8.22$ Hz, 1H), 7.13 (d, $J = 8.61$ Hz, 1H), 3.88 (s, 3H). Purity > 99%.

Methyl 2-(4-(benzyloxy)-3-bromophenyl)-1H-benzo[d]imidazole-6-carboxylate (3.27). To a mixture of methyl 2-(3-bromo-4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.26**) (0.100 g, 0.288 mmol) and cesium carbonate (0.141 g, 0.432 mmol) in acetonitrile (1.35 mL) was added benzyl bromide (0.038 mL, 0.317 mmol) and the resulting mixture was heated at 70°C for

16 h. The cooled mixture was then diluted with chloroform (15 mL) and the insolubles were removed by filtration. The filtrate was concentrated under reduced pressure and the crude residue was taken up in DMF and purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.011 g, 0.025 mmol, 9%). LC (Analytical HPLC- General method A): 2.144 min. HRMS (ESI): calcd for $C_{22}H_{18}BrN_2O_3$ $[M+H]^+$ m/z 437.0495, found 437.0522. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.44 (d, $J = 2.35$ Hz, 1H), 8.15 - 8.22 (m, 2H), 7.87 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.68 (d, $J = 8.22$ Hz, 1H), 7.48 - 7.54 (m, 2H), 7.40 - 7.47 (m, 3H), 7.34 - 7.39 (m, 1H), 5.33 (s, 2H), 3.88 (s, 3H). Purity = 92%.

Methyl 2-(3-bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (3.28). To a solution of methyl 2-(3-bromo-4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.26**) (0.300 g, 0.864 mmol) and cesium carbonate (0.422 g, 1.296 mmol) stirred in acetone (5.7 mL) was added 1-bromo-3-chloropropane (0.094 mL, 0.951 mmol). The resulting mixture was then stirred at 70°C for 16 h. After cooling, the reaction mixture was diluted with 15 mL of chloroform and the solids were removed by filtration. The filtrate was concentrated to dryness and the residue obtained was purified by flash chromatography (24 g cartridge) eluting with a gradient of ethyl acetate in dichloromethane (from 0 to 25%) to give the title compound as a white solid (0.097g, 0.229mmol, 27% yield). LC (Analytical HPLC- General method A): 2.048 min. HRMS (ESI): calcd for $C_{18}H_{17}BrClN_2O_3$ $[M+H]^+$ m/z 423.0106, 425.0084, found 423.0129, 425.0106. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.42 (d, $J = 2.35$ Hz, 1H), 8.15 - 8.22 (m, 2H), 7.87 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.68 (d, $J = 8.61$ Hz, 1H), 7.39 (d, $J = 8.61$ Hz, 1H), 4.30 (t, $J = 5.87$ Hz, 2H), 3.83 - 3.90 (m, 5H), 2.25 (quint, $J = 6.16$ Hz, 2H).

Methyl 2-(3-bromo-4-(3-(4-methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (3.29). Methyl 2-(3-Bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.28**) (0.020 g, 0.047 mmol) and 1-methylpiperazine (0.014 g, 0.142 mmol) were reacted according to General procedure F. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.010 g, 0.021 mmol, 44%). LC (Analytical HPLC- General method A): 1.547 min. HRMS (ESI): calcd for $C_{23}H_{28}BrN_4O_3$ $[M+H]^+$ m/z 487.1339, found 487.1368. 1H NMR (400 MHz, Methanol- d_4) δ ppm 8.35 (d, $J = 1.17$ Hz, 1H), 8.39 (d, $J = 2.35$ Hz, 1H), 8.06 - 8.16 (m, 3H), 7.75 (d, $J = 8.61$ Hz, 1H), 7.33 (d, $J = 8.61$ Hz, 1H), 4.30 (t, $J = 5.87$ Hz, 2H), 3.97 (s, 3H), 3.21 - 3.29 (m, 4H), 2.87 - 3.07 (m, 6H), 2.85 (s, 3H), 2.10 - 2.21 (m, 2H). Purity > 99%.

Methyl 2-(3-bromo-4-(3-morpholinopropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (3.30). Methyl 2-(3-Bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.28**) (0.030 g, 0.071 mmol) and morpholine (0.019 g, 0.212 mmol) were reacted according to General procedure F. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.009 g, 0.019 mmol, 27%). LC (Analytical HPLC- General method A): 1.495 min. HRMS (ESI): calcd for $C_{22}H_{25}BrN_3O_4$ $[M+H]^+$ m/z 476.1006, found 476.1034. 1H NMR (400 MHz, Methanol- d_4) δ ppm 8.39 (d, $J = 2.35$ Hz, 1H), 8.31 - 8.35 (m, 1H), 8.13 (dd, $J = 2.35, 8.61$ Hz, 1H), 8.08 (dd, $J = 1.57, 8.61$ Hz, 1H), 7.73 (d, $J = 8.61$ Hz, 1H), 7.33 (d, $J = 8.61$ Hz, 1H), 4.34 (t, $J = 5.67$ Hz, 2H), 4.02 - 4.20 (m, 2H), 3.96 (s, 3H), 3.70 - 3.91 (m, 2H), 3.53 - 3.70 (m, 2H), 3.41 - 3.52 (m, 2H), 3.15 - 3.29 (m, 2H), 2.31 - 2.44 (m, 2H). Purity = 98%.

Methyl 2-(4-(3-(benzylamino)propoxy)-3-bromophenyl)-1H-benzo[d]imidazole-6-carboxylate (3.31). Methyl 2-(3-Bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.28**) (0.030 g, 0.071 mmol) and benzylamine (0.023 mL, 0.212 mmol) were reacted according to General procedure F. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.009 g, 0.018 mmol, 26%). LC (Analytical HPLC- General method A): 1.749 min. HRMS (ESI): calcd for $C_{25}H_{25}BrN_3O_3$ $[M+H]^+$ m/z 496.1057, found 496.1085. 1H NMR (400 MHz, Methanol- d_4) δ ppm 8.36 (d, $J = 2.35$ Hz, 1H), 8.30 - 8.34 (m, 1H), 8.12 (dd, $J = 2.35, 8.61$ Hz, 1H), 8.04 (dd, $J = 1.57, 8.61$ Hz, 1H), 7.70 (d, $J = 8.61$ Hz, 1H), 7.45 - 7.56 (m, 5H), 7.26 - 7.32 (m, 2H), 4.33 (t, $J = 5.67$ Hz, 2H), 4.29 (s, 2H), 3.95 (s, 3H), 3.38 (t, $J = 7.43$ Hz, 2H), 2.26 - 2.37 (m, 2H). Purity = 91%.

Dimethyl 2,2'-((propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylate) (3.32). A vial was charged with methyl 2-(3-bromo-4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.28**) (0.030 g, 0.071 mmol), potassium carbonate (0.020 g, 0.142 mmol), sodium iodide (10.6 mg, 0.071 mmol), methyl 2-(3-bromo-4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (0.025 g, 0.071 mmol) and DMF (1.5 mL). The resulting mixture was heated at 90°C for 22 h and then the cooled mixture was diluted with chloroform and the solids were removed by filtration. The filtrate was then concentrated to dryness and the resulting crude residue was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as an amorphous white solid (0.009 g, 0.012 mmol, 17% yield). LC (Analytical HPLC- General method A): 2.156 min. HRMS (ESI): calcd for

$C_{33}H_{27}Br_2N_4O_6$ $[M+H]^+$ m/z 735.0275, found 735.0290. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.41 (d, J = 1.96 Hz, 2H), 8.18 (dd, J = 2.35, 8.61 Hz, 2H), 8.13 - 8.16 (m, 2H), 7.86 (dd, J = 1.76, 8.41 Hz, 2H), 7.67 (d, J = 8.22 Hz, 2H), 7.41 (d, J = 9.00 Hz, 2H), 4.41 (t, J = 6.06 Hz, 4H), 3.88 (s, 6H), 2.33 (td, J = 5.77, 11.93 Hz, 2H). Purity = 89%.

2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(N,N-dimethyl-1H-benzo[d]imidazole-6-carboxamide) (3.33). 2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22) (0.015 g, 0.021 mmol) and dimethylamine (2.367 μ l, 0.047 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.007 g, 9.20 μ mol, 43%). LC (Analytical HPLC- General method A): 1.903 min. HRMS (ESI): calcd for $C_{35}H_{33}Br_2N_6O_4$ $[M+H]^+$ m/z 759.0925, found 759.0917. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.43 (d, J = 2.35 Hz, 2H), 8.19 (dd, J = 2.35, 8.61 Hz, 2H), 7.64 (dd, J = 3.33, 4.89 Hz, 4H), 7.43 (d, J = 8.61 Hz, 2H), 7.27 - 7.33 (m, 2H), 4.42 (t, J = 6.26 Hz, 4H), 2.99 (br s, 12H), 2.29 - 2.40 (m, 2H). Purity > 99%.

2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(N-methyl-1H-benzo[d]imidazole-6-carboxamide) (3.34). 2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22) (0.015 g, 0.021 mmol) and methylamine (1.45 mg, 0.047 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.007 g, 9.56 μ mol, 45%). LC (Analytical HPLC- General method A): 1.895 min. HRMS (ESI): calcd for $C_{33}H_{29}Br_2N_6O_4$ $[M+H]^+$ m/z 731.0612, found 731.0590. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.46 (q, J = 4.30 Hz, 2H), 8.43 (d, J = 2.35 Hz, 2H), 8.20 (dd, J = 2.35, 8.61 Hz, 2H), 8.09 (s, 2H), 7.77 (dd, J = 1.57, 8.61 Hz, 2H), 7.63 (d, J = 8.61 Hz, 2H), 7.42 (d, J = 8.61 Hz, 2H), 4.42 (t, J = 6.06 Hz, 4H), 2.82 (d, J = 4.70 Hz, 6H), 2.29 - 2.38 (m, 2H). Purity = 88%.

(((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-2,6-diyl))bis(morpholinomethanone) (3.35). 2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22) (0.015 g, 0.021 mmol) and morpholine (4.07 μ l, 0.047 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.009 g, 10.66 μ mol, 50%). LC (Analytical HPLC- General

method A): 1.866 min. HRMS (ESI): calcd for C₃₉H₃₇Br₂N₆O₆ [M+H]⁺ *m/z* 843.1136, found 843.1126. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.43 (d, *J* = 2.35 Hz, 2H), 8.19 (dd, *J* = 2.35, 8.61 Hz, 2H), 7.59 - 7.70 (m, 4H), 7.43 (d, *J* = 9.00 Hz, 2H), 7.32 (dd, *J* = 1.17, 8.22 Hz, 2H), 4.34 - 4.47 (m, 4H), 3.50 - 3.56 (m, 16H), 2.28 - 2.40 (m, 2H). Purity > 99%.

2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(N-(2-hydroxyethyl)-1H-benzo[d]imidazole-6-carboxamide) (3.36). 2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22) (0.015 g, 0.021 mmol) and ethanolamine (2.8 μl, 0.047 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the desired product as a white solid (0.007 g, 8.83 μmol, 42%). LC (Analytical HPLC- General method A): 1.830 min. HRMS (ESI): calcd for C₃₅H₃₃Br₂N₆O₆ [M+H]⁺ *m/z* 793.0806, found 793.0808. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.46 (t, *J* = 5.67 Hz, 2H), 8.42 (d, *J* = 2.35 Hz, 2H), 8.20 (dd, *J* = 2.35, 8.61 Hz, 2H), 8.11 (s, 2H), 7.74 - 7.81 (m, 2H), 7.63 (d, *J* = 8.61 Hz, 2H), 7.42 (d, *J* = 9.00 Hz, 2H), 4.42 (t, *J* = 6.26 Hz, 4H), 3.54 (t, *J* = 6.26 Hz, 4H), 3.31 - 3.41 (m, 4H), 2.26 - 2.38 (m, 2H). Purity >99%.

2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(N-ethyl-N-(2-hydroxyethyl)-1H-benzo[d]imidazole-6-carboxamide) (3.37). 2,2'-((Propane-1,3-diylbis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.22) (0.015 g, 0.021 mmol) and 2-(ethylamino)ethanol (4.5 μl, 0.047 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.010 g, 0.012 mmol, 56%). LC (Analytical HPLC- General method A): 1.856 min. HRMS (ESI): calcd for C₃₉H₄₁Br₂N₆O₆ [M+H]⁺ *m/z* 847.1449, found 847.1451. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.41 (d, *J* = 2.35 Hz, 2H), 8.19 (dd, *J* = 2.35, 8.61 Hz, 2H), 7.61 (d, *J* = 8.22 Hz, 2H), 7.41 (d, *J* = 9.00 Hz, 2H), 7.23 (d, *J* = 7.83 Hz, 2H), 6.83 (t, *J* = 9.39 Hz, 2H), 4.41 (t, *J* = 6.06 Hz, 2H), 4.15 - 4.29 (m, 4H), 2.83 - 3.13 (m, 4H), 1.99 - 2.15 (m, 4H), 1.80 - 1.96 (m, 6H), 0.74 - 0.89 (m, 6H). Purity = 88%.

2,2'-(((2,2-Bis((4-(6-carboxy-1H-benzo[d]imidazol-2-yl)phenoxy)methyl)propane-1,3-diylbis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (3.38). A solution of dimethyl 2,2'-(((2,2-bis((4-(6-(methoxycarbonyl)-1H-benzo[d]imidazol-2-yl)phenoxy)methyl)propane-1,3-diylbis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylate) (3.39c) (0.020 g, 0.018 mmol), 1 M NaOH (0.211 mL, 0.211 mmol), methanol (0.21

mL) and THF (0.63 mL) was stirred at 70°C for 1 h. The cooled mixture was concentrated to dryness, the residue was acidified with 1M aqueous HCl and the mixture was again concentrated to dryness. The residue was then suspended in methanol, collected by filtration and washed with ether to give the title compound as a white solid (0.008 g, 7.40 μ mol, 42%). LC (Analytical HPLC- General method A): 1.899 min. HRMS (ESI): calcd for C₆₁H₄₅N₈O₁₂ [M+H]⁺ *m/z* 1081.3151, found 1081.3145. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 12.98 (br s, 4H), 8.27 (d, *J* = 9.00 Hz, 8H), 8.15 (s, 4H), 7.86 - 7.94 (m, 4H), 7.69 (d, *J* = 8.61 Hz, 4H), 7.31 (d, *J* = 9.00 Hz, 8H), 4.55 (br s, 8H). Purity > 99%.

Methyl 2-(4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (3.39a). A 100 mL sealed tube was charged with methyl 3,4-diaminobenzoate (1.857 g, 11.18 mmol), 4-hydroxybenzaldehyde (1.365 g, 11.18 mmol), DMF (47.6 mL) and sodium metabisulfite (2.125 g, 11.18 mmol). The resulting reaction mixture was stirred for 3 h at 80°C. The cooled mixture was filtered and the filter-cake was suspended in water-acetonitrile (1:1) and sonicated for 30 min. The resulting slurry was filtered and the obtained solid was dried in vacuo to give the title compound as a yellow solid (2.057 g, 7.67 mmol, 69%). Purification of a small portion of this material (60 mg) by prep HPLC (Preparative HPLC- General method A) gave the product as a white solid. LC (Analytical HPLC- General method A): 1.505 min. HRMS (ESI): calcd for C₁₅H₁₃N₂O₃ [M+H]⁺ *m/z* 269.0921, found 269.0931. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.16 - 8.20 (m, 1H), 8.05 (d, *J* = 8.61 Hz, 2H), 7.94 (dd, *J* = 1.57, 8.61 Hz, 1H), 7.74 (d, *J* = 8.61 Hz, 1H), 7.01 (d, *J* = 8.61 Hz, 2H), 3.89 (s, 3H). Purity > 99%.

2,2-Bis((tosyloxy)methyl)propane-1,3-diyl bis(4-methylbenzenesulfonate) (3.39b). To a solution of 2,2-bis(hydroxymethyl)propane-1,3-diol (0.500 g, 3.67 mmol) in pyridine (8.9 mL) was added *p*-toluenesulfonyl chloride (3.14 g, 16.45 mmol) and the resulting mixture was stirred at room temperature for 48 h. Cold water was then added and the precipitate was collected by filtration, washed with cold water and dried in vacuo to give the title compound as a white solid (1.781 g, 2.366 mmol, 64%). LC (Analytical HPLC- General method A): 2.270 min. HRMS (ESI): calcd for C₃₃H₃₇O₁₂S₄ [M+H]⁺ *m/z* 753.1162, found 753.1076. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 7.64 (d, *J* = 8.61 Hz, 8H), 7.46 (d, *J* = 8.22 Hz, 8H), 3.80 (s, 8H), 2.43 (s, 12H).

2,2'-(((2,2-Bis((4-(6-methylcarboxy-1H-benzo[d]imidazol-2-yl)phenoxy)methyl)propane-1,3-diyl)bis(oxy))bis(4,1-phenylene))bis(6-methylcarboxy-1H-benzo[d]imidazole) (3.39c). A mixture of methyl 2-(4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.39a**) (0.100 g,

0.373 mmol), 2,2-bis((tosyloxy)methyl)propane-1,3-diyl bis(4-methylbenzenesulfonate) (**3.39b**) (0.070 g, 0.093 mmol) and potassium carbonate (0.052 g, 0.373 mmol) in DMF (1.75 mL) was stirred while heating to reflux for 24 h. The cooled mixture was concentrated to dryness, the residue was suspended in water and the solid obtained was collected by filtration. The crude solid was then taken up in DMF and purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as an amorphous white solid (0.024 g, 0.021 mmol, 23%). LC (Analytical HPLC- General method A): 2.135 min. HRMS (ESI): calcd for C₆₅H₅₃N₈O₁₂ [M+H]⁺ *m/z* 1137.3777, found 1137.3598. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.05 - 8.16 (m, 12H) 7.82 (dd, *J* = 8.4, 1.8 Hz, 4H) 7.62 (d, *J* = 8.6 Hz, 4H) 7.26 (d, *J* = 9.0 Hz, 8H) 4.53 (br s, 8H) 3.87 (s, 12H). Purity = 82%.

Methyl 2-(4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (3.40). To a solution of methyl 2-(4-hydroxyphenyl)-1H-benzo[d]imidazole-6-carboxylate (Scheme 3-7, Step a, **3.39a**; 0.800 g, 2.98 mmol) and cesium carbonate (1.457 g, 4.47 mmol) stirred in acetone (15 mL) was added 1-bromo-3-chloropropane (0.324 mL, 3.28 mmol) and the resulting mixture was heated at 80°C for 16 h. The cooled mixture was then concentrated to dryness and the residue was purified by flash chromatography (40 g cartridge, 0-5% MeOH-dichloromethane) to give the title compound as a white solid (0.340 g, 0.986 mmol, 33%). LC (Analytical HPLC- General method A): 1.824 min. HRMS (ESI): calcd for C₁₈H₁₈ClN₂O₃ [M+H]⁺ *m/z* 345.1000, found 345.1021. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.08 - 8.16 (m, 3H), 7.83 (d, *J* = 7.43 Hz, 1H), 7.53 - 7.71 (m, 1H), 7.15 (d, *J* = 8.61 Hz, 2H), 4.19 (t, *J* = 6.06 Hz, 2H), 3.87 (s, 3H), 3.82 (t, *J* = 6.46 Hz, 2H), 2.21 (quint, *J* = 6.16 Hz, 2H).

Methyl 2-(4-(3-(4-methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (3.41). A mixture of methyl 2-(4-(3-chloropropoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.40**) (0.146 g, 0.423 mmol), 1-methylpiperazine (0.141 mL, 1.270 mmol), cesium carbonate (0.414 g, 1.270 mmol) and sodium iodide (0.063 g, 0.423 mmol) in DMF (10 mL) was heated at 100°C for 20 h. The cooled mixture was diluted with chloroform and the insolubles were removed by filtration. The filtrate was then concentrated to dryness and the residue was purified by flash chromatography [eluting with 0-40% dichloromethane-methanol-NH₄OH (50:45:5)-dichloromethane] to give the title compound as a white solid (0.096 g, 0.235 mmol, 56%). LC (Analytical HPLC- General method A): 1.348 min. HRMS (ESI): calcd for C₂₃H₂₉N₄O₃ [M+H]⁺ *m/z* 409.2234, found 409.2199. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.05 - 8.15 (m, 3H), 7.80

(d, $J = 8.61$ Hz, 1H), 7.49 - 7.69 (m, 2H), 7.11 (d, $J = 9.00$ Hz, 2H), 4.08 (t, $J = 6.26$ Hz, 2H), 3.85 (s, 3H), 2.18 - 2.44 (m, 10H), 2.13 (s, 3H), 1.87 (quint, $J = 6.85$ Hz, 2H). Purity = 97%.

2-(4-(3-(4-Methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylic acid (3.42). A solution of methyl 2-(4-(3-(4-methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylate (**3.41**) (0.196 g, 0.480 mmol) and 1 M sodium hydroxide (1.44 mL, 1.44 mmol) in a mixture of methanol (1.44 mL) and THF (4.32 mL) was stirred with heating at 70°C for 1 h. The cooled mixture was concentrated to dryness and the crude residue obtained was treated with 1 M HCl (1.44 mL, 1.44 mmol). The resulting slurry was filtered, and the filter-cake was washed with water and then dried in vacuo to give the title compound as a solid. It was used as such for the next step without further purification. LC (Analytical HPLC- General method B): 0.992 min. MS (APCI): calcd for $C_{22}H_{26}N_4O_3$ $[M+H]^+$ m/z 395.2, found 395.2.

N-Benzyl-2-(4-(3-(4-methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxamide (3.43). 2-(4-(3-(4-Methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylic acid (**3.42**) (0.025 g, 0.063 mmol) and benzylamine (0.015 mL, 0.133 mmol) were reacted according to General procedure D. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.011 g, 0.023 mmol, 36%). LC (Analytical HPLC- General method A): 1.513 min. HRMS (ESI): calcd for $C_{29}H_{34}N_5O_2$ $[M+H]^+$ m/z 484.2707, found 484.2717. 1H NMR (400 MHz, DMSO- d_6) δ ppm 9.11 (t, $J = 6.06$ Hz, 1H), 8.11 - 8.20 (m, 3H), 7.83 - 7.91 (m, 1H), 7.68 (d, $J = 8.61$ Hz, 1H), 7.30 - 7.38 (m, 4H), 7.21 - 7.29 (m, 1H), 7.18 (d, $J = 9.00$ Hz, 2H), 4.53 (d, $J = 5.87$ Hz, 2H), 4.16 (t, $J = 6.26$ Hz, 2H), 3.22 - 3.49 (m, 4H), 2.83 - 3.11 (m, 6H), 2.78 (br s, 3H), 1.95 - 2.11 (m, 2H). Purity = 90%.

2-(4-(3-(4-Methylpiperazin-1-yl)propoxy)phenyl)-N-(2-morpholinoethyl)-1H-benzo[d]imidazole-6-carboxamide (3.44). 2-(4-(3-(4-Methylpiperazin-1-yl)propoxy)phenyl)-1H-benzo[d]imidazole-6-carboxylic acid (**3.42**) (0.025 g, 0.063 mmol) and 2-morpholinoethanamine (0.017 mL, 0.133 mmol) were reacted according to General Method A. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.009 g, 0.018 mmol, 28%). LC (Analytical HPLC- General method A): 1.013 min. HRMS (ESI): calcd for $C_{28}H_{39}N_6O_3$ $[M+H]^+$ m/z 507.3078, found 507.3070. 1H NMR (400 MHz, DMSO- d_6) δ ppm 9.66 - 9.86 (m, 1H), 8.75 (t, $J = 5.67$ Hz, 1H), 8.07 - 8.18 (m, 3H), 7.77 (d, $J = 9.39$ Hz, 1H), 7.64 (d, $J = 8.22$ Hz, 1H), 7.14 (d, $J = 9.00$ Hz, 2H),

4.13 (t, $J = 6.26$ Hz, 2H), 3.84 - 4.06 (m, 4H), 3.54 - 3.84 (m, 12H), 3.26 - 3.40 (m, 2H), 3.13 (dt, $J = 4.11, 7.53$ Hz, 2H), 2.76 (br s, 3H), 1.93 - 2.06 (m, 2H), 1.14 - 1.28 (m, 2H). Purity = 95%.

2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (3.45). A sealable tube was charged with 3-bromo-4-methoxybenzaldehyde (1.00 g, 4.65 mmol), 3,4-diaminobenzoic acid (0.708 g, 4.65 mmol), anhydrous DMF (20 mL) and sodium metabisulfite (0.884 g, 4.65 mmol), and the resulting mixture was heated at 80°C for 3 h. The cooled mixture was then concentrated to dryness and the crude residue obtained was partitioned with ethyl acetate-water. The organic layer was separated, dried over MgSO₄, filtered and concentrated to afford the crude title compound as a grey solid that was used as such in the next step (1.80 g, 5.18 mmol, 111%). LC (Analytical HPLC- General method A): 1.657 min. HRMS (ESI): calcd for C₁₅H₁₂BrN₂O₃ [M+H]⁺ m/z 347.0026, found 347.0048. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.41 (d, $J = 2.35$ Hz, 1H), 8.20 (dd, $J = 2.15, 8.80$ Hz, 1H), 8.15 (s, 1H), 7.84 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.64 (d, $J = 8.22$ Hz, 1H), 7.34 (d, $J = 9.00$ Hz, 1H), 3.95 (s, 3H). Purity > 99%.

2-(3-Bromo-4-methoxyphenyl)-N,N-diethyl-1H-benzo[d]imidazole-6-carboxamide (3.46). 2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (3.45) (0.050 g, 0.144 mmol) and diethylamine hydrochloride (0.017 g, 0.158 mmol) were reacted according to General procedure E. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.034 g, 0.085 mmol, 59%). LC (Analytical HPLC- General method A): 1.727 min. HRMS (ESI): calcd for C₁₉H₂₁BrN₃O₂ [M+H]⁺ m/z 402.0812, found 402.0830. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 8.42 (d, $J = 1.96$ Hz, 1H), 8.20 (dd, $J = 2.35, 8.61$ Hz, 1H), 7.67 (d, $J = 8.22$ Hz, 1H), 7.59 (s, 1H), 7.37 (d, $J = 8.61$ Hz, 1H), 7.26 (dd, $J = 1.76, 8.41$ Hz, 1H), 3.97 (s, 3H), 3.37 (br s, 4H), 1.13 (br s, 6H). Purity = 93%.

N-Benzyl-2-(3-bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxamide (3.47). 2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (3.45) (0.050 g, 0.144 mmol) and phenylethanamine (0.017 g, 0.158 mmol) were reacted according to General procedure E. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.028 g, 0.064 mmol, 45%). LC (Analytical HPLC- General method A): 1.867 min. HRMS (ESI): calcd for C₂₂H₁₉BrN₃O₂ [M+H]⁺ m/z 436.0655, found 436.0655. ¹H NMR (400 MHz, DMSO-d₆) δ ppm 9.12 (t, $J = 6.06$ Hz, 1H), 8.43 (d, $J = 2.35$ Hz, 1H), 8.21 (dd, $J = 2.35, 8.61$ Hz, 1H), 8.18 (s, 1H), 7.87 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.69 (d, $J =$

8.22 Hz, 1H), 7.29 - 7.41 (m, 5H), 7.19 - 7.29 (m, 1H), 4.53 (d, $J = 6.26$ Hz, 2H), 3.97 (s, 3H). Purity = 98%.

(2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazol-6-yl)(4-methylpiperazin-1-yl)methanone (3.48). 2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (**3.45**) (0.050 g, 0.144 mmol) and 1-methylpiperazine (0.016 g, 0.158 mmol) were reacted according to General procedure E. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.028 g, 0.065 mmol, 45%). LC (Analytical HPLC- General method A): 1.301 min. HRMS (ESI): calcd for $C_{20}H_{22}BrN_4O_2$ $[M+H]^+$ m/z 429.0921, found 429.0946. 1H NMR (400 MHz, DMSO- d_6) δ ppm 9.99 (br s, 1H), 8.41 (d, $J = 2.35$ Hz, 1H), 8.19 (dd, $J = 2.35, 8.61$ Hz, 1H), 7.71 (s, 1H), 7.66 (d, $J = 8.22$ Hz, 1H), 7.28 - 7.36 (m, 2H), 3.94 (s, 3H), 3.21 - 3.50 (m, 6H), 3.02 - 3.14 (m, 2H), 2.82 (s, 3H). Purity > 99%.

2-(3-Bromo-4-methoxyphenyl)-N-(2-morpholinoethyl)-1H-benzo[d]imidazole-6-carboxamide (3.49). 2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (**3.45**) (0.050 g, 0.144 mmol) and 2-morpholinoethanamine (0.021 g, 0.158 mmol) were reacted according to General procedure E. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.030 g, 0.065 mmol, 45%). LC (Analytical HPLC- General method A): 1.415 min. HRMS (ESI): calcd for $C_{21}H_{24}BrN_4O_3$ $[M+H]^+$ m/z 459.1026, found 459.1017. 1H NMR (400 MHz, DMSO- d_6) δ ppm 9.77 (br s, 1H), 8.77 (t, $J = 5.87$ Hz, 1H), 8.42 (d, $J = 2.35$ Hz, 1H), 8.21 (dd, $J = 2.35, 8.61$ Hz, 1H), 8.14 (s, 1H), 7.79 (dd, $J = 1.76, 8.41$ Hz, 1H), 7.67 (d, $J = 8.61$ Hz, 1H), 7.35 (d, $J = 8.61$ Hz, 1H), 3.97 - 4.08 (m, 2H), 3.96 (s, 3H), 3.53 - 3.74 (m, 6H), 3.36 (t, $J = 6.26$ Hz, 2H), 3.17 (br s, 2H). Purity > 99%.

(2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazol-6-yl)(morpholino)methanone (3.50). 2-(3-Bromo-4-methoxyphenyl)-1H-benzo[d]imidazole-6-carboxylic acid (**3.45**) (0.050 g, 0.144 mmol) and morpholine (0.014 g, 0.158 mmol) were reacted according to General procedure E. The crude residue obtained was purified by prep HPLC (Preparative HPLC- General method A) to give the title compound as a white solid (0.038 g, 0.091 mmol, 63%). LC (Analytical HPLC- General method A): 1.536 min. HRMS (ESI): calcd for $C_{19}H_{19}BrN_3O_3$ $[M+H]^+$ m/z 418.0589, found 418.0586. 1H NMR (400 MHz, DMSO- d_6) δ ppm 8.44 (d, $J = 2.35$ Hz, 1H), 8.21 (dd, $J = 2.35,$

8.61 Hz, 1H), 7.68 - 7.75 (m, 2H), 7.33 - 7.42 (m, 2H), 3.97 (s, 3H), 3.41 - 3.71 (m, 8H). Purity > 99%.

Purification of DfrB1 and DfrB4 for kinetic characterization. Histidine-tagged DfrB1 and DfrB4 (using previously reported constructs WT R67 DHFR pQE32(26) and His₆-DfrB4(16), respectively) were overexpressed and purified as described previously(12, 19) with the following modifications. The cell culture volume was 400 mL. During purification on His-Trap HP(12), the imidazole concentration for the plateau was 0.2 M and the elution concentration was 0.5 M. Fractions containing DfrB were identified according to activity assay and analysis on tricine SDS-PAGE(46) then concentrated using an Amicon concentrator (MWCO 3,000, Millipore) for application on a Superose 12 column (1.6 cm × 55 cm) pre-equilibrated with 0.05 M potassium phosphate, pH 8.0. Fractions containing DfrB were identified as above. Sample purity was determined by tricine SDS-PAGE. Concentration was determined using the Bio-Rad protein assay (Bio-Rad, Hercules, CA),

Purification of the human dihydrofolate reductase (hDfr). Overexpression of the hexahistidine-tagged human dihydrofolate reductase (WT His₆-hDHFR) and cell lysis was described previously(47). The lysate was applied on a 5 mL nickel-affinity column (GE Healthcare) with an Äkta fast-protein liquid chromatography (Amersham Biosciences) at 0.5 mL/min. The column was washed with 5 column volumes (CV) of 0.1 M potassium phosphate buffer (pH 8.0) at 1 ml/min. A stepwise gradient of 0.1 M potassium phosphate buffer (pH 8.0) containing increasing concentrations of imidazole (0.02 M, 0.05 M, 0.06 M, 0.08 M, 0.1M, 0.16M and 0.2 M; 2 CV per concentration) was applied. Fractions containing WT His₆-hDHFR were identified according to the standard spectrophotometric DHFR activity assay and analysis on SDS-PAGE. Pooled fractions were concentrated using an Amicon concentrator (MWCO 10,000, Millipore) and further purified on a pre-equilibrated Superose 12 column (1.6 cm × 55 cm) with 0.1 M potassium phosphate buffer (pH 7.5). Fractions containing His₆-hDHFR were identified as above. His₆-hDHFR concentration was determined using the Bio-Rad protein assay.

Construction of the DfrB1 variant for crystallization purposes. To enable removal of the N-terminal 6-histidine tag from DfrB1 following its purification, we inserted an enterokinase cleavage site (DDDDK) into the construct WT R67 DHFR pQE32(26). A forward oligonucleotide primer (5'-
CATGGATCCAAGACTACAAAGACGATGACGACAAGGTATTCCCATCGGACGCCAC-

3'; Sigma-Aldrich (Oakville, ON) was designed to include the protease cleavage site (bold) and the DfrB1 sequence beginning at Val17, the probable chymotrypsin cleavage site, with the *Bam*HI restriction site. A reverse primer removed the 13-residue C-terminal tail (ELGTPGRPAAKLN) that was previously introduced for other purposes(26) and *Hind*III (underlined): 5'-GGGAAGCTTTTAGTTGATGCGTTCAAGCGCC-3'. The PCR product obtained with Phusion High-fidelity DNA polymerase (Thermo Scientific (Waltham, MA, USA)) was digested with *Bam*HI and *Hind*III, isolated on agarose gel and extracted according to the 'EZ-10 Spin Column PCR Products Purification Kit' protocol (Bio Basic, Burlington, ON, Canada). The insert was ligated to *Bam*HI and *Hind*III-digested and dephosphorylated (FastAP Thermosensitive Alkaline Phosphatase, 30 min) pQE32 (Qiagen) with the DNA Ligation Kit Ver.2.1 (Takara, USA). After 30 minutes at 16°C, the DNA ligation solutions were transformed into DH5 α . DNA sequencing confirmed the expected construct *DfrB1-His₆V17 Δ 16/ENTK-pQE32*. The *DfrB1-His₆V17 Δ 16/ENTK-pQE32* was transformed into *E. coli* BL21 pRep4 (Qiagen) for expression.

Purification and Crystallization of DfrB1 soaked with 1. In order to crystallize DfrB1 in complex with **3.1**, an identical protein expression and purification protocol was employed as described in (23). Following expression and purification, the INSF tandem dimer, as previously described(23), was concentrated to 20 mg/mL in 100 mM Tris pH 8.0. Immediately before crystallization, chymotrypsin was added to the sample in a ratio of 1:100 chymotrypsin:INSF, and the protein was diluted to 15 mg/mL using MPD, resulting in a final MPD concentration of 25%. Reservoirs were prepared using 750 μ L of 100 mM sodium phosphate pH 7.6 and 60% MPD in a Greiner 24-well hanging-drop crystallization plate. On a siliconized glass cover slip (Hampton Research), 1.5 μ L of protein were combined with 2.5 μ L of reservoir solution and suspended over the well. The plate was incubated at 277 K and crystals were obtained in a few days. Drops were soaked with 0.2 μ L of a solution of **3.1** in 100% DMSO for a final concentration of 2 mM of **3.1**.

Purification of the DfrB1 variant for crystallization purposes. The crystallization variant DfrB1 (DfrB1-His₆V17 Δ 16/ENTK) was overexpressed and purified as previously described(12, 19), with the modifications noted above for DfrB1 and DfrB4. DfrB1-His₆V17 Δ 16/ENTK was concentrated to 2 mg/ml (24 mg in total) with an Amicon concentrator (MWCO 3,000, Millipore) and it was incubated with enterokinase (13 U) for 20 h at 22°C. The sample was concentrated and further purified on Superose 12 as described above, except that the column was linked at the bottom to His-Trap HP cartridge pre-equilibrated with 0.05 mM potassium phosphate, pH 8.0.

This last step allows removal of uncleaved DfrB1-His₆V17Δ16/ENTK. Fractions containing DfrB1 were identified and purity determined as above. The protein was concentrated 0.05 mM potassium phosphate, pH 8.0 to 13.3 mg/ml for crystallization.

Crystallization of DfrB1-His₆V17Δ16/ENTK was performed as previously described(28). Briefly, the protein concentration was adjusted from 13.3 mg/ml to 10 mg/mL by addition of a final concentration of 25 % MPD. Crystals were obtained using the hanging-drop vapour-diffusion method in a Greiner 24-well hanging-drop crystallization plate. Reservoirs were prepared in Eppendorf tubes with 100mM Tris-Cl pH 8.0 55% MPD. On a siliconized glass cover slip (Hampton Research), 2.0 μL of protein solution were combined with 2.0 μL of the reservoir solution. The plate was incubated at 277 K, and crystals were obtained after 3-4 days. Drops were soaked with 0.4 μL of a solution of **3.3** in 100% DMSO for a final concentration of 12 mM of **3.3** five days prior to crystal harvest.

Data were collected under standard cryogenic conditions using a Rigaku MicroMax-007HF generator equipped with VariMax HF optics and a Saturn 944+ CCD detector. The data were processed using HKL-2000 (Table A2-I)(48). The integrated and scaled data were imported into the CCP4 software suite(49). Iterative rounds of positional and B-factor refinement using Refmac(50) and manual model building using Coot(51) were then performed, using the coordinates from PDB ID 3SFM with all solvent and ligands removed(23) as the starting model. As the 3SFM structure was crystallized in the same space group and with the same cell dimensions, care was taken to ensure that the same reflections were included in the R_{free} test set. During the late stages of refinement, anisotropic B-factor refinement was introduced, and ligand and solvent molecules were incorporated into the model. Coordinates and structure factors were deposited to the PDB as PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’

Kinetic studies. For all kinetic assays, pH of solutions was 7. IC₅₀ and K_i values were determined as previously described(12, 19) except that the reaction volume was 100 μL. The assays were conducted in Corning™ 96-well flat-bottomed polystyrene plates (Thermo Scientific (Waltham, MA, USA)). Initial rates (first 15 % of substrate consumption) were collected with a Beckman DTX 880 plate reader integrated with a BioMek NX automated liquid handling workstation (Beckman Coulter, Brea, CA). The length of compounds (**3.4** to **3.9** and **3.13** to **3.19**) was

calculated with Chemdraw 3D (ChemOffice 2017, PerkinElmer Informatics). K_M^{NADPH} , K_M^{DHF} and k_{cat} were determined as previously described(12).

Building small molecule input database and initial pose generation. All studied molecules were sketched in ChemDraw Professional 16. All atomistic 3D coordinates at pH7 were generated using the Database Wash application in MOE2018(52). We generated the DfrB1:(**3**)₂ based on the 1,4-dioxobutan-2-ol central core of the crystal structure PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’ as the starting point. The structure was prepared using the QuickPrep application with default parameters. The two molecules of **3.3** were built simultaneously in a stepwise manner using a standard structure-based drug design workflow. The generated DfrB1:(**3.3**)₂ complex was used as input structure to dock all remaining molecules.

Building models of protein:inhibitor complexes. Docking of inhibitors was performed using the Dock application in MOE2018(52). One molecule of **3.3** was replaced by a first molecule of the query inhibitor. Substitution of the second molecule of **3.3** by the second molecule of inhibitor followed. The final dimer complex was assembled from each docking run based on the best RMSD compared to **3.3**. As Placement Method, the Triangle Matcher was used with 30 retained poses based on London dG scoring. Refinement was done in Rigid Receptor mode and the best 10 poses from the GBVI/WSA dG scoring function were written to the output database.

Conformational exploration of ligand dimers in the tunnel. Conformational exploration of each inhibitor molecule bound to DfrB1 was performed using the LowModeMD method(53) in MOE2018(52). All inhibitor atoms were kept free while receptor atoms within the active-site tunnel (all residues 4.5 Å from inhibitor molecules) were tethered with 10 kcal/mol starting from 0.25 Å deviation around the original coordinates. All residues 4.5Å from the tethered atoms were fixed and the rest of the system was set to inert. The AMBER10:EHT forcefield was used for LowModeMD with R-Field solvation. The exploration was stopped after 10,000 conformations or after 100 consecutive conformations already present in the database. All conformations with an energy difference of 100 kcal/mol compared to the lowest energy conformation were retained.

RMSD calculated for 3.1. In order to profile similarity between the position of **3.1** in our simulation and in the crystal structure, RMSDs were calculated for each of the 2,992 conformations from the simulation of **3.1**. From the crystal structure, the three additional subunits to create the complete tetramer were generated. Maintaining the inhibitor molecule in the position

obtained from the simulation (i.e. without performing any superimposition), the RMSD over all inhibitor atoms present in the crystal structure were calculated against the inhibitor in each of the four possible symmetry-generated positions from the crystal structure, and the lowest of the four RMSD values was used. In addition, the RMSD was calculated after superimposing the simulated inhibitor molecules onto the crystal structure inhibitor molecule.

Identification of ligand binding mode. The Protein Ligand Interaction Fingerprint (PLIF) application in MOE2018(52) was used to study the average binding mode of each inhibitor bound to DfrB1. The calculation used the default parameters.

3.8 . SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI:

Crystallographic data, Analysis of molecular simulations of binding interactions, Additional kinetic data

3.9 . ABBREVIATION

AMT, aminopterin; Dfr, dihydrofolate reductase; DfrBs, type II Dfr; DHF, dihydrofolate; hDfr, human Dfr; SUL, sulfonamide; MOE, Molecular Operating Environment; MTX, methotrexate; PLIF, protein ligand interaction fingerprint; PMTX, pemetrexed; TMP, trimethoprim.

3.10 . REFERENCES

1. Giguère S, Prescott JF, Dowling PM. Antimicrobial therapy in veterinary medicine. John Wiley & Sons Inc. 2013;47:256-7.
2. Organization WH, WHO. WHO list of critically important antimicrobials for human medicine.
3. Then RL. History and future of antimicrobial diaminopyrimidines. Journal of Chemotherapy. 1993;5(6):361-8.
4. Masters PA, O'Bryan TA, Zurlo J, Miller DQ, Joshi N. Trimethoprim-sulfamethoxazole revisited. Archives of Internal Medicine. 2003;163:402.
5. Cabello FC. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. Environmental microbiology. 2006;8:1137-44.
6. L'Abée-Lund TM, Sørum H. Class 1 integrons mediate antibiotic resistance in the fish pathogen *Aeromonas salmonicida* worldwide. Microbial drug resistance. 2001;7:263-72.
7. Yassin AK, Gong J, Kelly P, Lu G, Guardabassi L, Wei L, et al. Antimicrobial resistance in clinical *Escherichia coli* isolates from poultry and livestock, China. PLoS One. 2017;12(9):e0185326.
8. Kadlec K, Kehrenberg C, Schwarz S. Molecular basis of resistance to trimethoprim, chloramphenicol and sulphonamides in *Bordetella bronchiseptica*. Journal of Antimicrobial Chemotherapy. 2005;56:485-90.
9. Sunde M. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. Journal of Antimicrobial Chemotherapy. 2005;56:1019-24.
10. Skold O. Resistance to trimethoprim and sulfonamides. Veterinary Research. 2001;32:261-73.
11. Stapleton PJ, Lundon DJ, McWade R, Scanlon N, Hannan MM, O'Kelly F, et al. Antibiotic resistance patterns of *Escherichia coli* urinary isolates and comparison with antibiotic consumption data over 10 years, 2005-2014. Irish Journal of Medical Science. 2017;186(3):733-41.

12. Toulouse JL, Abraham SMJ, Kadnikova N, Bastien D, Gauchot V, Schmitzer AR, et al. Investigation of classical organic and ionic liquid cosolvents for early-stage screening in fragment-based inhibitor design with unrelated bacterial and human dihydrofolate reductases. *Assay and Drug Development Technologies*. 2017;15:141-53.
13. Broad DF, Smith JT. Classification of trimethoprim-resistant dihydrofolate reductases mediated by R-plasmids using isoelectric focussing. *European Journal of Biochemistry*. 1982;125:617-22.
14. Howell EE. Searching sequence space: Two different approaches to dihydrofolate reductase catalysis. *ChemBioChem*. 2005;6:590-600.
15. Huovinen P. Trimethoprim resistance. *Antimicrobial Agents and Chemotherapy*. 1987;14:51-6.
16. Toulouse JL, Edens TJ, Alejandre L, Manges AR, Pelletier JN. Integron-associated DfrB4, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*. 2017;61:1-5.
17. Alonso H, Gready JE. Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends in Microbiology*. 2006;14:236-42.
18. Ho JMW, Juurlink DN. Considerations when prescribing trimethoprim-sulfamethoxazole. *Canadian Medical Association Journal*. 2011;183:1851-8.
19. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, et al. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of Medicinal Chemistry*. 2012;55:3182-92.
20. Duff MR, Chopra S, Strader MB, Agarwal PK, Howell EE. Tales of dihydrofolate binding to R67 dihydrofolate reductase. *Biochemistry*. 2016;55:133-45.
21. Fling ME, Elwell LP. Protein expression in *Escherichia coli* minicells containing recombinant plasmids specifying trimethoprim-resistant dihydrofolate reductases. *Journal of Bacteriology*. 1980;141:779-85.
22. Narayana N. High-resolution structure of a plasmid-encoded dihydrofolate reductase: Pentagonal network of water molecules in the D2-symmetric active site. *Acta Crystallographica Section D Structural Biology*. 2006;62:695-706.

23. Yachnin BJ, Colin DY, Volpato JP, Ebert M, Pelletier JN, Berghuis AM. Novel crystallization conditions for tandem variant R67 DHFR yield a wild-type crystal structure. *Acta Crystallographica Section F Structural Biology Communications*. 2011;67:1316-22.
24. Matthews DA, Smith SL, Baccanari DP, Burchall JJ, Oatley SJ, Kraut J. Crystal structure of a novel trimethoprim-resistant dihydrofolate reductase specified in *Escherichia coli* by R-plasmid R67. *Biochemistry*. 1986;25(15):4194-204.
25. Carrillo N, Ceccarelli EA, Roveri OA. Usefulness of kinetic enzyme parameters in biotechnological practice. *Biotechnology and Genetic Engineering Reviews*. 2010;27:367-382.
26. Schmitzer AR, Lépine F, Pelletier JN. Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: Creating alternative functional configurations. *Protein Engineering, Design and Selection*. 2004;17:809-19.
27. Jackson M, Chopra S, Smiley RD, Maynord PON, Rosowsky A, London RE, et al. Calorimetric studies of ligand binding in R67 dihydrofolate reductase. *Biochemistry*. 2005;44:12420-33.
28. Krahn JM, Jackson MR, DeRose EF, Howell EE, London RE. Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry*. 2007;46:14878-88.
29. Kamath G, Howell EE, Agarwal PK. The tail wagging the dog: Insights into catalysis in R67 dihydrofolate reductase. *Biochemistry*. 2010;49:9078-88.
30. Assaraf YG. Molecular basis of antifolate resistance. *Cancer and Metastasis Reviews*. 2007;26(1):153-81.
31. Lele AC, Mishra DA, Kamil TK, Bhakta S, Degani MS. Repositioning of DHFR Inhibitors. *Current Topics in Medicinal Chemistry*. 2016;16(19):2125-43.
32. Piper JR, McCaleb GS, Montgomery JA, Kisliuk RL, Gaumont Y, Sirotiak FM. Syntheses and antifolate activity of 5-methyl-5-deaza analogues of aminopterin, methotrexate, folic acid, and N10-methylfolic acid. *Journal of Medicinal Chemistry*. 1986;29(6):1080-7.
33. Kisliuk RL. Deaza analogs of folic acid as antitumor agents. *Current Pharmaceutical Design*. 2003;9(31):2615-25.
34. Chopra S, Dooling RM, Horner CG, Howell EE. A balancing act between net uptake of water during dihydrofolate binding and net release of water upon NADPH binding in R67 dihydrofolate reductase. *Journal of Biological Chemistry*. 2008;283:4690-8.

35. Park H, Zhuang P, Nichols R, Howell EE. Mechanistic studies of R67 dihydrofolate reductase. Effects of pH and an H62C mutation. *Journal of Biological Chemistry*. 1997;272:2252-8.
36. Bradrick TD, Beechem JM, Howell EE. Unusual binding stoichiometries and cooperativity are observed during binary and ternary complex formation in the single active pore of R67 dihydrofolate reductase, a D2 symmetric protein. *Biochemistry*. 1996;35:11414-24.
37. Lu Y, Liu Y, Xu Z, Li H, Liu H, Zhu W. Halogen bonding for rational drug design and new drug discovery. *Expert Opinion on Drug Discovery*. 2012;7:375-83.
38. Parisini E, Mentrangolo P, Pilati T, Resnati G, Terraneo G. Halogen bonding in halocarbon-protein complexes: a structural survey. *Chemical Society Reviews*. 2011;40:2267-78.
39. Ballatore C, Huryn DM, Smith AB, 3rd. Carboxylic acid (bio)isosteres in drug design. *ChemMedChem*. 2013;8(3):385-95.
40. Segel IH. *Enzyme kinetics: behavior and analysis of rapid equilibrium and steady state enzyme systems*: Wiley; 1975.
41. Shoichet BK. Screening in a spirit haunted world. *Drug Discovery Today*. 2006;11(13-14):607-15.
42. Hicks SN, Smiley RD, Stinnett LG, Minor KH, Howell EE. Role of Lys-32 residues in R67 dihydrofolate reductase probed by asymmetric mutations. *Journal of Biological Chemistry*. 2004;279:46995-7002.
43. Blakley RL. Crystalline dihydropteroylglutamic acid. *Nature*. 1960;188:231-2.
44. Mayence A, Pietka A, Collins MS, Cushion MT, Tekwani BL, Huang TL, et al. Novel bisbenzimidazoles with antileishmanial effectiveness. *Bioorganic and Medicinal Chemistry Letters*. 2008;18:2658-61.
45. Cappoen D, Forge D, Vercammen F, Mathys V, Kiass M, Roupie V, et al. Biological evaluation of bisbenzaldehydes against four *Mycobacterium* species. *European Journal of Medicinal Chemistry*. 2013;63:731-8.
46. Schagger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*. 1987;166:368-79.

47. Volpato JP, Fossati E, Pelletier JN. Increasing methotrexate resistance by combination of active-site mutations in human dihydrofolate reductase. *Journal of Molecular Biology*. 2007;373:599-611.
48. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*. 1997;276:307-26.
49. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D Structural Biology*. 2011;67(Pt 4):235-42.
50. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D Structural Biology*. 1997;53(Pt 3):240-55.
51. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallographica Section D Structural Biology*. 2010;66(Pt 4):486-501.
52. ULC CCG. *Molecular Operating Environment (MOE)*. 2013.08.
53. Labute P. LowModeMD-Implicit low mode velocity filtering applied to conformational search of macrocycles and protein loops. *Journal of Chemical Information and Modeling*. 2010;50(5):792-800.

Préface au chapitre 4

Le triméthoprim (TMP) est un antibiotique utilisé dans le traitement des infections des voies urinaires et respiratoires ainsi que de manière prophylactique chez les animaux d'élevage. L'utilisation générale et à grande échelle du TMP accélère l'incidence de la résistance à celui-ci. Les dihydrofolates réductases de type II (DfrB) émergentes et largement ignorées, constituent une source de résistance au TMP. Ces DfrBs, structurellement et évolutivement distinctes des Dfr chromosomiques, sont retrouvés sur des plasmides. À ce jour, seule la variante DfrB1 a bien été caractérisée. Nous rapportons ici, l'inhibition de DfrB1 avec les inhibiteurs de type bisubstrat de la 6-hydroxyméthyl-7,8-dihydroptérine pyrophosphokinase (HPPK), enzyme aussi retrouvée dans la voie de biosynthèse du folate. L'homologie de séquence élevée parmi les membres de la famille DfrB suggère qu'elles pourraient être aussi inhibées par les molécules de type bisubstrats. Nous avons développé un protocole simple et efficace, de purification par précipitation thermique pour la DfrB1 et cinq autres DfrBs. Leurs constantes de vitesse catalytique ($k_{cat} = 0,2 - 0,41 \text{ s}^{-1}$) et leurs constantes de Michaelis-Menten ($K_M = 1,3 - 5,8 \text{ }\mu\text{M}$) pour le substrat et le cofacteur sont conservées. Les DfrBs sont toutes insensibles aux fortes concentrations de TMP ($K_i = 0,38 - 1,3 \text{ mM}$). Il est important de noter qu'elles sont toutes inhibées par les inhibiteurs de type bisubstrat, ce qui concorde avec la conservation élevée de leur noyau catalytique. Nos résultats indiquent que l'optimisation des inhibiteurs de type bisubstrats à double cible pourrait cibler HPPK ainsi que toute la famille DfrB.

Pour ce manuscrit en préparation, ma contribution, sous la supervision du Dr. Joelle Pelletier, dans cet article est l'idée de l'article, l'écriture de la première ébauche, la création des figures et des tableaux, la purification ainsi que les analyses cinétiques des protéines. Daniel Deon, Marc Gagnon, Edward Ruediger et Kévin Saint-Jacques, sous la supervision du Dr. Anne Marinier, ont synthétisé l'inhibiteur **4.6**. Delphine Forge, sous la supervision de Jean Jacques Vanden Eynde, a synthétisé l'inhibiteur **4.7**. Genbin Shi, sous la supervision du Dr. Xinhua Ji du National Cancer Institute, a synthétisé les inhibiteurs de type bisubstrat, **4.1** à **4.5**. Ceux-ci prévoient tester les molécules **4.3** à **4.5** pour l'inhibition de l'enzyme HPPK ainsi que compléter le protocole de synthèse retrouvé dans la section « *Materials and methods* » de l'Annexe 3.

Certaines molécules de ce chapitre ont été présentées précédemment. La molécule **4.1** a été présentée comme étant la molécule **1.10** au chapitre 1 alors que les molécules **4.6** et **4.7** proviennent du chapitre 3 et elles se nomment **3.21b** et **3.12** respectivement.

Chapitre 4 – « *Inhibition of an emergent family of intrinsically trimethoprim-resistant type II dihydrofolate reductases with dual target inhibitors* »

4.1 . Article 3

Jacynthe L. Toulouse^{1,2}, Genbin Shi³, Daniel Deon⁴, Marc Gagnon⁴, Edward Ruediger⁴, Kévin Saint-Jacques⁵, Delphine Forge⁶, Jean Jacques Vanden Eynde⁶, Anne Marinier⁴, Xinhua Ji³ and Joelle N. Pelletier (joelle.pelletier@umontreal.ca)^{1,2,7,*}

¹Département de biochimie et de chimie, Université de Montréal, Qc, Canada

²PROTEO, the Québec Network for Research on Protein, Function, Engineering and Applications

³Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702, USA

⁴Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Qc, Canada

⁵Département de chimie, Université de Sherbrooke, Qc, Canada

⁶Laboratoire de chimie organique, Université de Mons, Belgium,

*Corresponding author, Phone :

514-343-2124, Fax : 514-343-7586, E-mail : joelle.pelletier@umontreal.ca

4.2 . ABSTRACT

Trimethoprim (TMP) is an antibiotic employed in the treatment of human urinary and respiratory tract infections as well as prophylactically in livestock. The worldwide, extensive use of TMP is accelerating the incidence of TMP resistance. A source of TMP resistance are the emergent and largely unheeded type II dihydrofolate reductases (DfrBs), a family of plasmid-borne Dfrs that are structurally and evolutionarily unrelated to chromosomal Dfrs. Only the DfrB1 variant has been well characterized to date. Here, we report the inhibition of DfrB1 with bisubstrate inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), also an antibiotic target of the folate pathway. The high homology shared by DfrB family members led us to suspect that they may be similarly inhibited. We developed a straightforward heat-precipitation protocol to effectively purify DfrB1 and five heterologously expressed homologs. Their catalytic rate constants ($k_{\text{cat}} = 0.2 - 0.41 \text{ s}^{-1}$) and productive binding ($K_{\text{M}} = 1.3 - 5.8 \text{ }\mu\text{M}$) for substrate and cofactor are conserved and all are insensitive to high TMP concentrations ($K_{\text{i}} = 0.38 - 1.3 \text{ mM}$). Importantly, all are inhibited by the bisubstrate inhibitors, consistent with the high conservation of their catalytic core. Our results indicate that efforts in optimizing the dual target, bisubstrate inhibitors will be rewarded by targeting HPPK and the entire DfrB family.

KEYWORDS

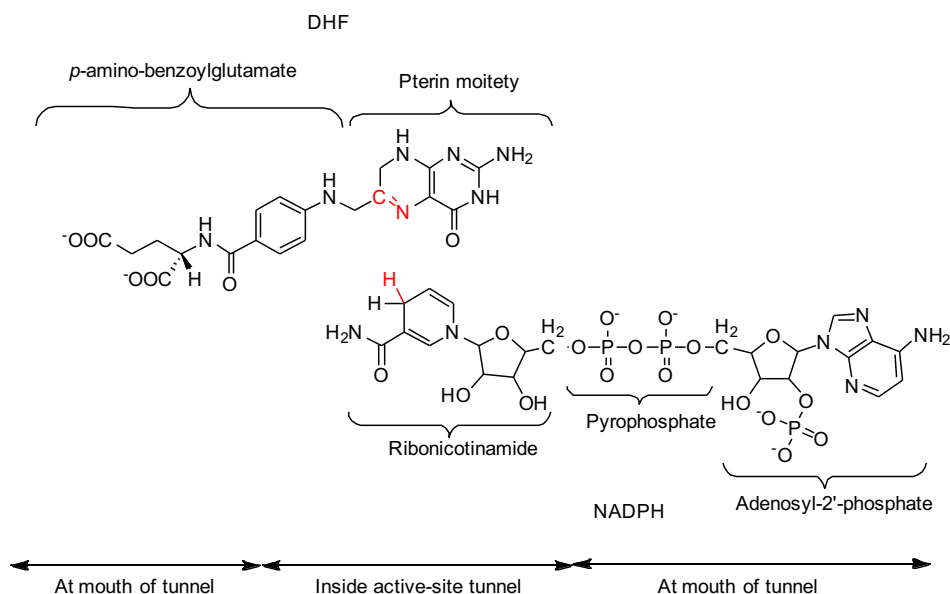
Type II Dihydrofolate reductase, trimethoprim resistance, bisubstrate inhibitor, dual inhibitor.

4.3 . INTRODUCTION

Dihydrofolate reductases (Dfr) have long been targets of the antibiotic therapy as a result of their essential role in the folate pathway. Trimethoprim (TMP) is an antibiotic that selectively inhibits bacterial chromosomal Dfr with little effect on mammalian Dfrs(1). TMP is often used in combination with sulfamethoxazole (SUL) as a front-line treatment against known or unknown infections from aerobic bacteria and occasionally against protozoa(2). The TMP-SUL combination is categorized by the World Health Organization as highly important due to its effectiveness in treatment of human urinary and respiratory tract infections(2-4) and to reduce mortality in people living with HIV/AIDS(5). SUL acts synergistically with TMP, targeting a second enzyme in the microbial folate pathway that is lacking in humans(6). The simultaneous

inhibition of two targets with this combined antibiotic regime has been demonstrated to delay onset of antibiotic resistance(7). Nonetheless, the spread of resistance to TMP in humans(8) as well as livestock(2, 9-11), whether in the context of TMP administered alone or with SUL, is accelerating(8). It is of particular concern that TMP resistance has been shown to be transferable from livestock to humans(2, 12).

Microbial resistance to TMP may result either from mutations in their chromosomal Dfr or as a result of acquiring the evolutionarily and structurally unrelated type II Dfrs (DfrBs)(13-15). DfrBs are plasmid-borne and have been identified in TMP-resistant Gram-negative bacteria, although there is little knowledge of their incidence in clinical samples (3, 10, 11, 15-18). *DfrB* genes were first identified in TMP-resistant wastewater or clinical samples where the cause of this resistance by those genes was not established(3, 10, 11, 15-18). The high TMP resistance of four DfrBs (DfrB1 to DfrB4) has been ascertained *in vitro* (19-21), but no such information is available for the more recently reported DfrB5 to DfrB9.



Scheme 4-1. Depiction of the relative binding orientation of the substrate dihydrofolate (DHF) and the reducing cofactor NADPH within the DfrB1 active-site tunnel.

Atoms involved in the hydride transfer reaction are in red.

To this day, the DfrB family of dihydrofolate reductases remains poorly studied. DfrB1, the only well-characterized DfrB, is a homotetrameric protein constituted of four β -barrel (SH3 domain) protomers contributing equally to the symmetrical, hourglass-shaped active-site cavity

(Figure 4-1A)(13, 14, 21-26). It is unrelated to the monomeric, Rossmann fold-containing bacterial chromosomal Dfrs(14). DfrBs share 74% to 98% (Table A3-I) sequence identity where the β -barrel core of DfrB1 is highly conserved (89% to 98%) (Figure 4-1B). Most of the sequence variation occurs in the unstructured *N*-termini, which are referred to as entropic bristles(27), demonstrated with DfrB1 that interactions between *N*-termini from adjacent protomers to contribute weakly to the stability of the homotetramer(27).

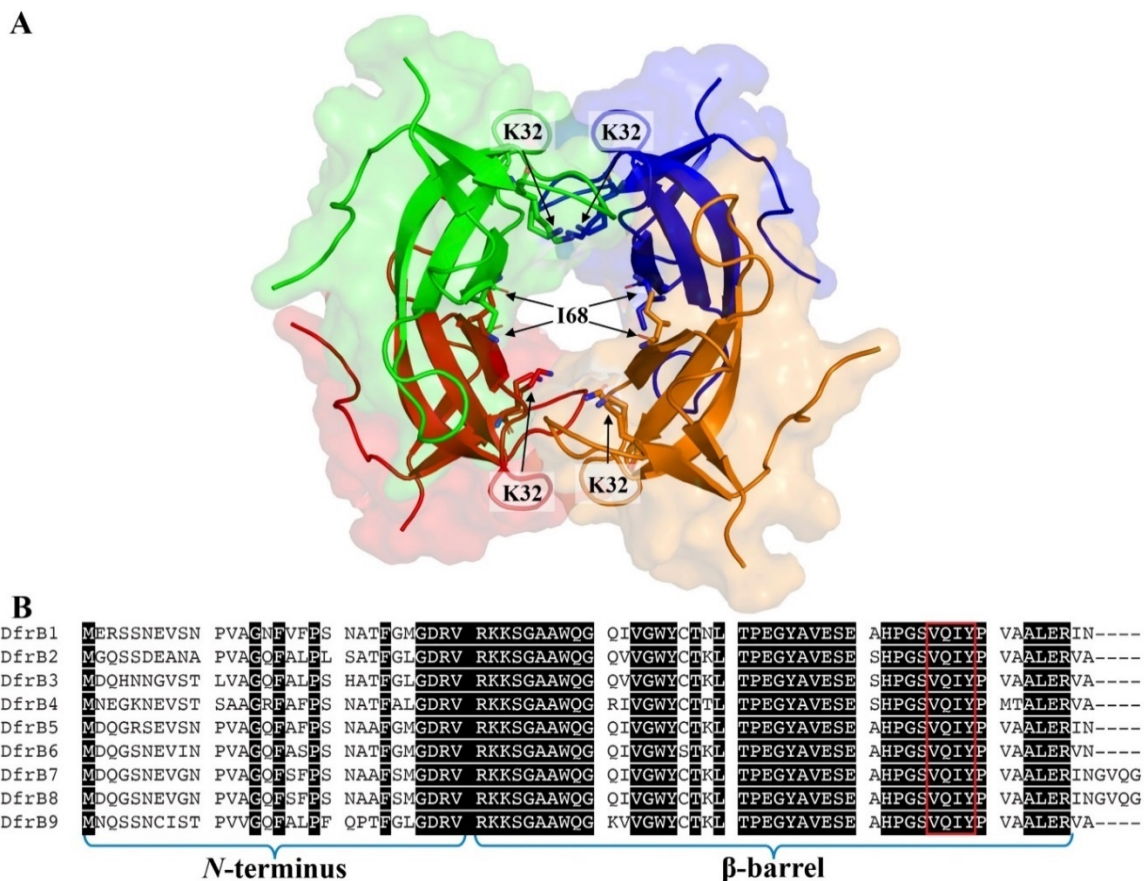


Figure 4-1. The homotetrameric DfrB1 structure (PDB: 2RK1) and the DfrB family sequence alignment.

A) The first 20 unstructured *N*-terminal residues from each protomer are not represented, and each protomer is colored differently. For clarity, only I68 of the Val66-Gln67-Ile68-Tyr69 (VQIY) region and K32 are represented in sticks on each protomer. B) The red rectangle highlights the conserved VQIY residues that line the active site tunnel. The *N*-terminus and the β -barrel core are identified.

DfrBs are intrinsically TMP-resistant and are emergent sources of antibiotic resistance. Their anticipated health-threatening activity calls for the development of new inhibitors. We

have reported the first selective, although weak, inhibitors of the hexahistidine-tagged version of DfrB1 (His₆-DfrB1). These bisbenzimidazoles offer K_i of 1.7 – 12.0 μM (*Toulouse J.L. 2019 submitted*)(28).

Here, we report the inhibition of DfrB1 with inhibitors of *E. coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), an enzyme that also belongs to the bacterial folate pathway, thus establishing these as being dual target inhibitors. They are based on substructures of folic acid and NADPH, qualifying them as bisubstrate inhibitors. Furthermore, we demonstrate with two of the bisubstrate inhibitors their broad inhibitory effect as they also inhibit all DfrB family members assayed. That study was enabled by determination, for the first time, that DfrB members representing the entire extent of sequence diversity exhibit Dfr activity and low sensitivity to TMP, comparable to DfrB1. The potency of the bisubstrate inhibitors toward all DfrBs was comparable to that of the previously-reported bisbenzimidazole inhibitors of DfrB1, confirming the interest in further developing these dual target inhibitors to combat TMP resistance.

4.4 . RESULTS AND DISCUSSION

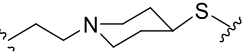
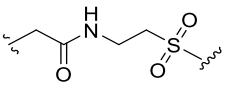
Dual inhibitors, also named hybrid molecules, generally consist of two linked pharmacophores having two distinct biological targets(29). The main asset of these dual target inhibitors is the administration of one drug molecule instead of two individual drugs(29). The folate pathway has long been the focus of dual inhibitors, the Dfr-thymidylate synthase pair being the focus of intense research(30). In mammals, the folate pathway starts with the intake of dietary folate whereas the bacteria are folate autotrophs(31). Bacterial folate biosynthesis depends, among other enzymes, on HPPK which constitutes a potential antibiotic target as it has no mammalian homolog(32). The bisubstrate molecules **4.1** to **4.5** were designed to mimic both the pterin moiety of DHF and the adenosine moiety of NADPH, and are thus designated as bisubstrates (Table 4-I). Bisubstrate molecules **4.1** and **4.2** inhibit HPPK with K_i of 2.55 and 3.1 μM, respectively(32, 33).

The tetrameric tunnel of DfrB1 is lined with the conserved VQIY sequence that is four-fold represented, forming a single symmetrical active site. The VQIY region interacts sterically and through hydrogen bonds with the pterin moiety of the substrate, DHF(34). Binding studies have indicated that the adenosine moiety contributes to the binding of NADPH with DfrB1

whereas structural studies located the solvent-exposed adenosine moiety at the mouth of the active site tunnel(26, 34). As a result, the pterin and adenosine moieties possess two distinct binding sites (Scheme)(34). Because they are constituted of both a pterin and an adenosine moiety, the HPPK bisubstrate inhibitors have the potential to also inhibit DfrB1.

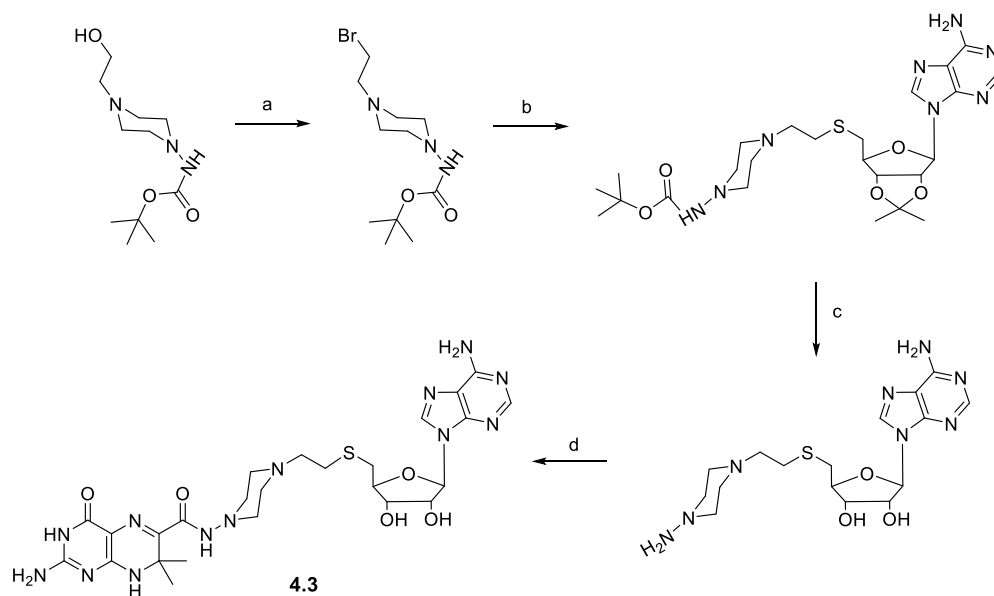
HPPK bisubstrate inhibitors **4.1** to **4.2** were tested for inhibition of His₆-DfrB1. We determined that molecules **4.1** and **4.2** inhibit His₆-DfrB1 with K_i of 20 μM and 18 μM, respectively (Table 4-I). Three novel bisubstrate molecules (**4.3** to **4.5**) (Schemes 4-2 to 4-4) were also assayed for inhibition of His₆-DfrB1 (Table 4-II). Molecules **4.3** to **4.5** inhibit with K_i of 10 - 17 μM, similarly to inhibitors **4.1** and **4.2**, further confirming the potential of this class of bisubstrate molecules as inhibitors of His₆-DfrB1. The potency of inhibitors **4.1** to **4.5** is comparable to the known bisbenzimidazole inhibitors of His₆-DfrB1, **4.6** and **4.7** (K_i = 3.5 and 7.4 μM, respectively) (*Toulouse J.L. 2019* submitted). These results confirm that **4.1** and **4.2** are dual HPPK and His₆-DfrB1 inhibitors and that the potency of **4.1** to **4.5** against His₆-DfrB1 is similar to the previously reported bisbenzimidazole inhibitors **4.6** and **4.7** (Scheme 4-5).

Table 4-I. Inhibition of His₆-DfrB1 and HPPK with bisubstrate inhibitors 4.1 and 4.2.

Inhibitor	X	K _i (μM)	
		His ₆ -DfrB1 ^a	HPPK ^b
4.1		20 ± 3	2.55
4.2		18 ± 4	3.1

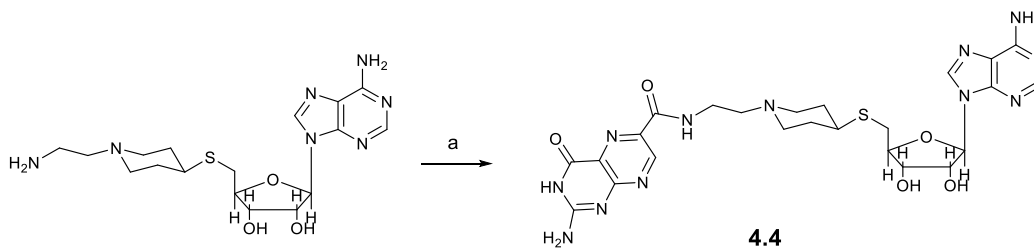
^a K_i values were calculated according to the experimentally determined IC₅₀ (Table A3-II). Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

^b Values taken from (35).



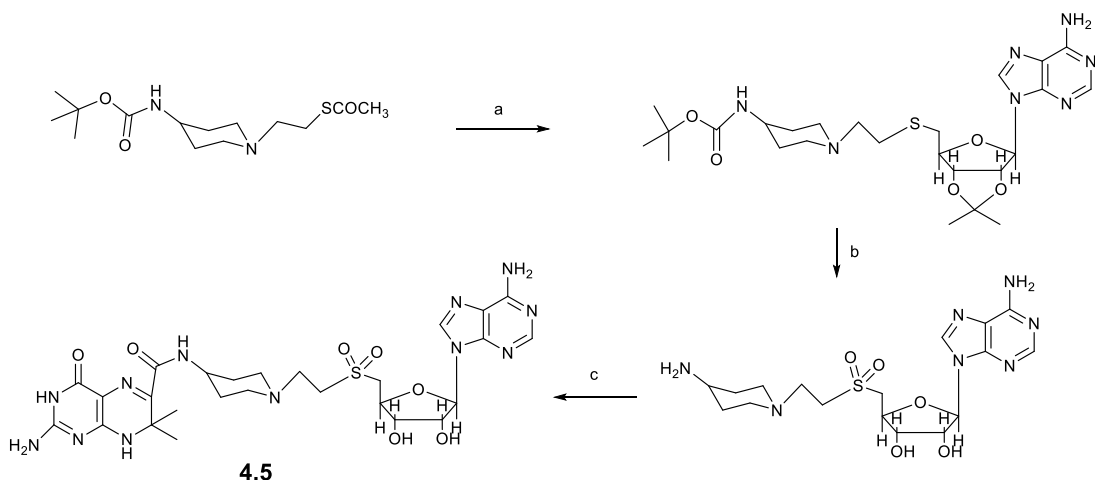
Scheme 4-2. Synthesis of bisubstrate inhibitor 4.3.

Reagents and conditions: a) PPh₃, CBr₄, CH₂Cl. b) NaOCH₃, 2',3'-O-isopropylideneadenosine-5'-thioacetate. c) CF₃COOH, CH₂Cl₂, d) HATU base, 6-carboxy-7,7-dimethyl-7,8-dihydropterin, DIPEA, DMF.



Scheme 4-3. Synthesis of bisubstrate inhibitor 4.4.

Reagents and conditions: a) HATU base, 2-amino-4-oxo-3,4-dihydropteridine-6-carboxylic acid, DIPEA, DMF.



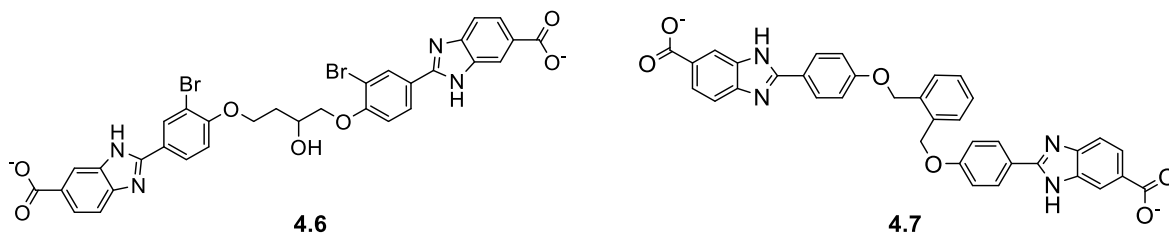
Scheme 4-4. Synthesis of bisubstrate inhibitor 4.5.

Reagents and conditions: a) NaOCH₃, MeOH, 2',3'-O-isopropylideneadenosine-5'-thioacetate. b) Oxone, TFA, CH₂Cl₂ c) HATU base, 6-amino-7,7-dimethyl-7,8-dihydropterin, DIPEA, DMF.

Table 4-II. Inhibition of His₆-DfrB1 with bisubstrate inhibitors 4.3 to 4.5.

Inhibitor	X	R ₁	K _i (μM) ^a
4.3		-CH ₃	16 ± 2
4.4		-H	10 ± 1
4.5		-CH ₃	17 ± 4

^a K_i values were calculated according to the experimentally determined IC₅₀ (Table A3-II). Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.



Scheme 4-5. Structures of 4.6 and 4.7.

Molecules 4.6 and 4.7 are known bisbenzimidazole inhibitors of His₆-DfrB1. (Toulouse J.L. 2019 submitted)

The high sequence identity of 74% - 98% of DfrBs (Table A3-I) suggested that His₆-DfrB1 inhibitors may inhibit the broader DfrB family. To verify this hypothesis, we selected six among the nine members of DfrB family to represent the greatest pairwise sequence differences: DfrB1 to DfrB5, and DfrB7. Variation is essentially confined to the *N*- and *C*-termini and to the turns in the β -barrel core, the core being highly conserved and constituting the active site (Figure 4-1A). For production of these proteins, we sought to verify the potential impact of an *N*-terminal hexahistidine-tag on inhibition. We thus turned to previously reported symmetrical bisbenzimidazole inhibitors **4.6** and **4.7**. They had been assayed against the *N*-terminally, hexahistidine-tagged His₆-DfrB1 harbouring additional non-native *C*-terminal residues, with K_i **4.6** = 2.6 μ M and K_i **4.7** = 7.4 μ M (Toulouse J.L. 2019 submitted). Comparison with inhibition of DfrB1 bearing its native termini yielded \approx 8-fold weaker inhibition, with K_i **4.6** = 18 μ M and K_i **4.7** = 62 μ M (Table 4-III), demonstrating the contribution of the non-native terminal additions to inhibitor binding. The impact of the non-native terminal additions on inhibition was further verified with two representative bisubstrate inhibitors, **4.1** and **4.3**, as we observed a 3- to 8.4-fold decrease in inhibition in the absence of the terminal additions (Table 4-III). This observation was unexpected since the symmetrical bisbenzimidazole inhibitors bear little resemblance to the bisubstrate inhibitors. To eliminate any contribution of the hexahistidine tag or the non-native *C*-terminal residues to inhibition, all variants were produced in their non-tagged, native form for further investigation.

Table 4-III. Inhibition of DfrB1 (without His₆-tag) with bisubstrate and bisbenzimidazole inhibitors 4.1, 4.3, 4.6 and 4.7.

Inhibitor:	K _i (μM) ^a			
	4.1	4.3	4.6	4.7
	130 ± 98	46 ± 5	18 ± 6	62 ± 6

^a K_i values were calculated according to the experimentally determined IC₅₀ (Table A3-III). Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

To access highly purified DfrB variants, we capitalized on the remarkable thermostability of DfrB1 and DfrB2, perhaps best illustrated by maintenance of activity upon boiling for 20 min(36-38). We purified the DfrBs using heat-precipitation of crude lysates. The temperature was selected to maximise the ratio between the folded, soluble DfrB (in solution) and the unfolded, aggregated proteins in the lysate. Treatment at 65°C followed by a size-exclusion purification step maximized yield at the expense of purity, which was only of 67% to 89% for DfrB2, DfrB3, DfrB5 and DfrB7 (data not shown). An additional heat treatment at 75°C after the size-exclusion purification enhanced the purity to ≥ 95 % (Table A3-IV; Figure A3-1). DfrB1 and DfrB4 were purified with one sole heat treatment at 75°C before the size-exclusion column purification step leading to final purities of 95 % (Table A3-IV; Figure A3-1). This one-day purification method provided 6.2 to 12.4 mg (3.7 to 8.5 U) of all DfrBs from 200 mL-scale expressions in *E. coli* (Table A3-IV). We note that the migration of the DfrBs upon SDS-PAGE does not directly reflect their molecular masses, which were confirmed by mass spectrometry (Figure A3-1, Table A3-V).

Prior to addressing their inhibition, it was essential to verify that the DfrB family members do indeed have Dfr activity. To our knowledge, the Dfr activity was verified only for DfrB1 to DfrB4(19-21). The Dfr activity was presumed for the more recent DfrB5 to DfrB9 based on their TMP-resistant-sample origin and the high sequence similarity with DfrB1. The main residues implicated in catalysis and the recognition of the substrate and cofactor are Ile68 (I68) of the VQIY region and Lys32 (K32)(34). They are conserved among the DfrB family, suggesting their possible Dfr activity.

All DfrBs assayed were active Dfrs (Table 4-IV). The DHF productive affinities were nearly identical for all DfrBs (K_m^{DHF} within 1 to 1.4-fold) and some variation was observed for

NADPH (K_m^{NADPH} : 1 to 4.5-fold variation). Those values are also similar to the *E. coli* chromosomal Dfr (*EcDfr*) ($K_m^{\text{DHF}} = 1.2 \mu\text{M}$ and $K_m^{\text{NADPH}} = 0.94 \mu\text{M}$)(39), as befits enzymes that compete for the same cellular resources.

Table 4-IV. Kinetic constants for the dihydrofolate reductase activity of DfrBs.

Dfrs	K_m^{DHF} (μM)	K_m^{NADPH} (μM)	$k_{\text{cat}}^{\text{DHF}}$ (s^{-1})	$k_{\text{cat}}^{\text{DHF}} / K_m^{\text{DHF}}$
<i>EcDfr</i>	1.2 ^a	0.94 ^a	12.0 ^{b,c}	10
DfrB1	4.4 ± 0.5	5.8 ± 0.8	0.32 ± 0.006	0.073
DfrB2	2.3 ± 0.2	3.3 ± 0.3	0.41 ± 0.01	0.18
DfrB3	3.0 ± 0.5	1.3 ± 0.3	0.20 ± 0.01	0.067
DfrB4	4.6 ± 0.7	3.5 ± 0.5	0.28 ± 0.01	0.061
DfrB5	3.2 ± 0.4	3.1 ± 0.2	0.32 ± 0.01	0.10
DfrB7	3.4 ± 0.4	2.8 ± 0.3	0.24 ± 0.007	0.071

^a Values taken from (39).

^b Value taken from (40).

^c k_{cat} value calculated for the global reaction.

The $k_{\text{cat}}^{\text{DHF}}$ of all DfrBs tested (Table 4-IV), determined by varying DHF, within 2-fold of each other ($k_{\text{cat}}^{\text{DHF}} = 0.20$ to 0.41 s^{-1}) and showed excellent agreement with the $k_{\text{cat}}^{\text{NADPH}}$, determined by varying NADPH (Table A3-VI). DfrBs are designated as poor catalysts relative to *EcDfr* ($k_{\text{cat}} = 12.0 \text{ s}^{-1}$)(40). This is reflected in their catalytic efficiency ($k_{\text{cat}}^{\text{DHF}} / K_m^{\text{DHF}}$) which is nearly two orders of magnitude lower than that of *EcDfr* (Table 4-IV). The high sequence similarity of DfrB6 and DfrB9 with their closest homolog among the assayed DfrBs (91% and 85%, respectively; Table A3-I) suggests that their catalytic properties, and those of closely related homologs that may be identified in the future, lie within this range.

Table 4-V. Inhibition of DfrBs with TMP and inhibitors 4.1, 4.3, 4.6 and 4.7.

Inhibitor:	K_i (μM) ^a				
	TMP	4.1	4.3	4.6	4.7
DfrB1	$(0.38 \pm 0.09) \times 10^3$	130 ± 98^b	46 ± 5^b	18 ± 6^b	62 ± 6^b
DfrB2	$(0.81 \pm 0.08) \times 10^3$	49 ± 3	52 ± 14	10 ± 2	29 ± 1
DfrB3	$(0.45 \pm 0.10) \times 10^3$	12 ± 2	17 ± 2	2.4 ± 0.4	6.4 ± 2.1
DfrB4	$(0.50 \pm 0.07) \times 10^3$	35 ± 4	51 ± 1	13 ± 1	20 ± 2
DfrB5	$(1.3 \pm 0.30) \times 10^3$	49 ± 9	42 ± 1	7.4 ± 0.1	29 ± 3
DfrB7	$(0.66 \pm 0.11) \times 10^3$	52 ± 32	37 ± 10	7.1 ± 2.2	19 ± 7

^a K_i values were calculated according to experimentally determined IC_{50} (Table A3-III and Table A3-VII). Values are given as the average \pm standard deviation from the mean of at least triplicates of two independent experiments.

^b Values from Table 3.

The K_i for TMP for *EcDfr* is 20 pM indicating strong inhibition(41). *E. coli* cannot overcome the TMP inhibition by increasing the copy level of *EcDfr* since the overexpression of this enzyme results in cellular toxicity(42). The *dfrB* genes are plasmid-borne; their copy level is unknown. Only the K_i^{TMP} for DfrB2 has been reported(20). We determined that the K_i^{TMP} for DfrB1 to DfrB5 and DfrB7 (0.38 mM to 1.3 mM) are seven to eight orders of magnitude higher than for *EcDfr*, confirming that the DfrB family displays high TMP resistance. Our data demonstrate that the family of DfrBs possess the activity and tolerance to TMP that should support bacterial proliferation in the presence of high concentrations of TMP and thus provide a survival advantage to bacteria carrying *dfrB* genes(13-15).

Considering the similarities between all DfrBs, we hypothesized that they can be similarly inhibited. To verify this hypothesis, we tested bisubstrate inhibitors **4.1** and **4.3** against DfrB2 to DfrB5 and DfrB7 (Table 4-V). Bisubstrates **4.1** and **4.3** inhibit the DfrBs with K_i **4.1** = 12 – 130 μM and K_i **4.3** = 17 – 52 μM . This demonstrates that bisubstrate **4.1** is a dual HPPK and DfrB family inhibitor. We further verified that the reported DfrB1 bisbenzimidazole inhibitors **4.6** and **4.7** inhibit all DfrBs with K_i in the low μM range (K_i **4.6** = 2.4 - 17.5 μM and K_i **4.7** = 6.4 - 61.7 μM) (Table 4-V). Overall, bisbenzimidazole inhibitors **4.6** and **4.7** offer slightly better affinity than bisubstrate inhibitors **4.1** and **4.3** (K_i values differ 2-7-fold for any

given DfrB). Our results demonstrate the promise held by the bisubstrate inhibitors, given their dual-target advantage.

Short molecular dynamic studies of molecules **4.6** and **4.7** with DfrB1 have revealed three binding hotspots: the K32 network (Lys32, Gly35 and Ala36), the YTT cluster (Tyr46, Thr48 and Thr51) and the VQIY region (*Toulouse J.L. 2019* submitted). These regions are conserved among DfrBs, consistent with the similar inhibition we observe here. Our results with two inhibitor classes, the symmetrical bisbenzimidazoles and the bisubstrates, indicate that early-stage drug development targeting DfrBs can be conducted on a single DfrB that will adequately represent the family. DHF, NADPH, and inhibitors **4.6** and **4.7** share two common binding hotspots that are conserved among DfrBs: K32 network and VQIY region. Inhibitors **4.1** and **4.3** are based on the DHF and NADPH binding moieties, suggesting that these inhibitors should bind to those same hotspots. These features are useful knowledge when it comes for designing an antibiotic targeting the DfrBs. Antibiotic resistance arises, in most cases, from mutation of the antibiotic target. Targeting the substrate binding residues of DfrBs could slow down the rise of antibiotic resistance since these residues are required for this essential reaction.

4.5 . CONCLUSION

A heat-precipitation method was developed to rapidly procure highly purified DfrB1, DfrB2, DfrB3, DfrB4, DfrB5 and DfrB7. Kinetic characterization confirmed their Dfr activity and their low sensitivity to TMP. We report the first inhibitors with broad effect, targeting all of the tested DfrBs. Further optimization of affinity and additional properties of bisubstrate inhibitors **4.1** and **4.3**, and symmetrical bisbenzimidazole inhibitors **4.6** and **4.7**, could lead to an antibiotic targeting the entire DfrB family. In particular, parallel optimization of bisubstrate inhibitors **4.1** and **4.3** as HPPK inhibitors would result in a dual antibiotic mimicking the SUL-TMP combination while slowing down the emergence of TMP resistance without the inconvenience of co-administering two antibiotics.

4.6 . SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI: Additional enzyme kinetics data, DfrB purification characterization, supplemental materials and methods.

4.7 . ABBREVIATIONS

Dfr, dihydrofolate reductase; DfrBs, type II Dfr; DHF, dihydrofolate; *EcDfr*, *E. coli* chromosomal Dfr; His₆-DfrB1, hexahistidine-tagged DfrB1; HPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase; K32 network, Lys32, Gly35 and Ala36; SUL, sulfomethoxazole; TMP, trimethoprim; VQIY, Val66-Gln67-Ile68-Tyr69; YTT cluster, Tyr46, Thr48 and Thr51.

4.8 . REFERENCES

1. Toulouse JL, Abraham SMJ, Kadnikova N, Bastien D, Gauchot V, Schmitzer AR, et al. Investigation of classical organic and ionic liquid cosolvents for early-stage screening in fragment-based inhibitor design with unrelated bacterial and human dihydrofolate reductases. *Assay and Drug Development Technologies*. 2017;15:141-53.
2. Giguère S, Prescott JF, Dowling PM. *Antimicrobial therapy in veterinary medicine*. John Wiley & Sons Inc. 2013;47:256-7.
3. Masters PA, O'Bryan TA, Zurlo J, Miller DQ, Joshi N. Trimethoprim-sulfamethoxazole revisited. *Archives of Internal Medicine*. 2003;163:402.
4. Organization WH, WHO. WHO list of critically important antimicrobials for human medicine.
5. Suthar AB, Vitoria MA, Nagata JM, Anglaret X, Mbori-Ngacha D, Sued O, et al. Cotrimoxazole prophylaxis in adults, including pregnant women, with HIV: a systematic review and meta-analysis. *Lancet HIV*. 2015;2(4):e137-50.
6. Debabov D. Antibiotic resistance: Origins, mechanisms, approaches to counter. *Applied Biochemistry and Microbiology*. 2013;49:665-71.
7. Giannella M, Trearichi EM, Giacobbe DR, De Rosa FG, Bassetti M, Bartoloni A, et al. Effect of combination therapy containing a high-dose carbapenem on mortality in patients with

carbapenem-resistant *Klebsiella pneumoniae* bloodstream infection. *International Journal of Antimicrobial Agents*. 2018;51(2):244-8.

8. Stapleton PJ, Landon DJ, McWade R, Scanlon N, Hannan MM, O'Kelly F, et al. Antibiotic resistance patterns of *Escherichia coli* urinary isolates and comparison with antibiotic consumption data over 10 years, 2005-2014. *Irish Journal of Medical Science*. 2017;186(3):733-41.

9. Yassin AK, Gong J, Kelly P, Lu G, Guardabassi L, Wei L, et al. Antimicrobial resistance in clinical *Escherichia coli* isolates from poultry and livestock, China. *PLoS One*. 2017;12(9):e0185326.

10. Kadlec K, Kehrenberg C, Schwarz S. Molecular basis of resistance to trimethoprim, chloramphenicol and sulphonamides in *Bordetella bronchiseptica*. *Journal of Antimicrobial Chemotherapy*. 2005;56:485-90.

11. Sunde M. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. *Journal of Antimicrobial Chemotherapy*. 2005;56:1019-24.

12. Skold O. Resistance to trimethoprim and sulfonamides. *Veterinary Research*. 2001;32:261-73.

13. Broad DF, Smith JT. Classification of trimethoprim-resistant dihydrofolate reductases mediated by R-plasmids using isoelectric focussing. *European Journal of Biochemistry*. 1982;125:617-22.

14. Howell EE. Searching sequence space: Two different approaches to dihydrofolate reductase catalysis. *ChemBioChem*. 2005;6:590-600.

15. Huovinen P. Trimethoprim resistance. *Antimicrob Agents Chemother*. 1987;14:51-6.

16. Alonso H, Gready JE. Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends Microbiology*. 2006;14:236-42.

17. Ho JMW, Juurlink DN. Considerations when prescribing trimethoprim-sulfamethoxazole. *Canadian Medical Association Journal*. 2011;183:1851-8.

18. L'Abée-Lund TM, Sørum H. Class 1 integrons mediate antibiotic resistance in the fish pathogen *Aeromonas salmonicida* worldwide. *Microbial drug resistance*. 2001;7:263-72.

19. Broad DF, Smith JT. Classification of trimethoprim-resistant dihydrofolate reductases mediated by R-plasmids using isoelectric focussing. *European Journal of Biochemistry*. 1982;125:617-22.
20. Amyes SG, Smith JT. The purification and properties of the trimethoprim-resistant dihydrofolate reductase mediated by the R-factor, R388. *European Journal of Biochemistry*. 1976;61(2):597-603.
21. Toulouse JL, Edens TJ, Alejandre L, Manges AR, Pelletier JN. Integron-associated DfrB4, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*. 2017;61:1-5.
22. Krahn JM, Jackson MR, DeRose EF, Howell EE, London RE. Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry*. 2007;46:14878-88.
23. Duff MR, Chopra S, Strader MB, Agarwal PK, Howell EE. Tales of dihydrofolate binding to R67 dihydrofolate reductase. *Biochemistry*. 2016;55:133-45.
24. Fling ME, Elwell LP. Protein expression in *Escherichia coli* minicells containing recombinant plasmids specifying trimethoprim-resistant dihydrofolate reductases. *Journal of Bacteriology*. 1980;141:779-85.
25. Yachnin BJ, Colin DY, Volpato JP, Ebert M, Pelletier JN, Berghuis AM. Novel crystallization conditions for tandem variant R67 DHFR yield a wild-type crystal structure. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 2011;67:1316-22.
26. Jackson M, Chopra S, Smiley RD, Maynord PON, Rosowsky A, London RE, et al. Calorimetric studies of ligand binding in R67 dihydrofolate reductase. *Biochemistry*. 2005;44:12420-33.
27. Bhojane PP, Duff MR, Bafna K, Agarwal P, Stanley C, Howell EE. Small Angle Neutron Scattering Studies of R67 Dihydrofolate Reductase, a Tetrameric Protein with Intrinsically Disordered N-Termini. *Biochemistry*. 2017;56:5886-99.
28. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, et al. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of Medicinal Chemistry*. 2012;55:3182-92.

29. Rahman A, Zaman K. *Topics in Anti-Cancer Research*: Bentham Science Publishers; 2014.
30. Arooj M, Sakkiah S, Cao G, Lee KW. An innovative strategy for dual inhibitor design and its application in dual inhibition of human thymidylate synthase and dihydrofolate reductase enzymes. *PLoS One*. 2013;8(4):e60470.
31. Wright DL, Anderson AC. Antifolate agents: a patent review (2006 - 2010). *Expert Opinion on Therapeutic Patents*. 2011;21:1293-308.
32. Shi G, Shaw G, Liang YH, Subburaman P, Li Y, Wu Y, et al. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New design with improved properties. *Bioorganic and Medicinal Chemistry*. 2012;20:47-57.
33. Shi G, Shaw G, Li Y, Wu Y, Yan H, Ji X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New lead exhibits a distinct binding mode. *Bioorganic and Medicinal Chemistry*. 2012;20:4303-9.
34. Krahn JM, Jackson MR, DeRose EF, Howell EE, London RE. Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry*. 2007;46:14878-88.
35. Shaw GX, Li Y, Shi G, Wu Y, Cherry S, Needle D, et al. Structural enzymology and inhibition of the bi-functional folate pathway enzyme HPPK-DHPS from the biowarfare agent *Francisella tularensis*. *FEBS Journal*. 2014;281(18):4123-37.
36. Ebert MC, Morley KL, Volpato JP, Schmitzer AR, Pelletier JN. Asymmetric mutations in the tetrameric R67 dihydrofolate reductase reveal high tolerance to active-site substitutions. *Protein Science*. 2015;24(4):495-507.
37. Zhuang P, Yin M, Holland JC, Peterson CB, Howell EE. Artificial duplication of the R67 dihydrofolate reductase gene to create protein asymmetry. Effects on protein activity and folding. *Journal of Biological Chemistry*. 1993;268(30):22672-9.
38. Zolig JW, Hanggi UJ, Zachau HG. Isolation of a small DNA fragment carrying the gene for a dihydrofolate reductase from a trimethoprim resistance factor. *Molecular Genetics and Genomics*. 1978;164(1):15-29.
39. Howell EE, Booth C, Farnum M, Kraut J, Warren MS. A second-site mutation at phenylalanine-137 that increases catalytic efficiency in the mutant aspartate-27 to serine *Escherichia coli* dihydrofolate reductase. *Biochemistry*. 1990;29(37):8561-9.

40. Fierke CA, Johnson KA, Benkovic SJ. Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*. *Biochemistry*. 1987;26(13):4085-92.
41. Stone SR, Morrison JF. Mechanism of inhibition of dihydrofolate reductases from bacterial and vertebrate sources by various classes of folate analogues. *Biochimica et Biophysica Acta*. 1986;869(3):275-85.
42. Bhattacharyya S, Bershtein S, Yan J, Argun T, Gilson AI, Trauger SA, et al. Transient protein-protein interactions perturb *Escherichia coli* metabolome and cause gene dosage toxicity. *Elife*. 2016;5.

Chapitre 5 - Développement d'un criblage à haut débit pour une librairie d'extraits d'invertébrés marins.

5.1 . Introduction

Les inhibiteurs rapportés au chapitres 3 et 4 possèdent des K_i dans l'ordre du bas μM . Ces valeurs sont trop élevées pour que ces inhibiteurs soient considérés comme tête de série. Afin de découvrir des nouvelles classes de molécules susceptibles d'offrir une inhibition plus puissante soit l'ordre du nM, le criblage à haut débit d'une librairie de sources naturelles a été développé.

À ces fins, une librairie d'extraits biologiques marins (Prof. Raymond Andersen, Université de la Colombie-Britannique) a été utilisée pour un criblage *in vitro* sur la DfrB1(67, 69, 71). Cette librairie de près de 2 000 extraits contient plusieurs composés présents en quantités variables. Lorsqu'un extrait est identifié comme contenant un inhibiteur potentiel, cet extrait est fractionné pour séparer les composés constituant cet extrait selon une propriété telle la taille, la charge ou l'hydrophobie, isolant des sous-groupes de molécules constituant cet extrait. Ces fractions peuvent ensuite être testées pour identifier celles qui inhibent, pour être de nouveau fractionnées et testées jusqu'à l'isolation d'un composé actif(67, 69, 71).

5.2 . Méthodes

Les méthodes détaillées et le matériel pour le criblage sont décrits à l'annexe A4. Le criblage consistait en des tests d'activité standards de la DfrB1 tels qu'entrepris aux Chapitres 3 et 4 de cette thèse, avec l'ajout des extraits et suivi de la diminution de l'absorbance à 340 nm. Brièvement, les extraits d'invertébrés marins furent dégelés immédiatement avant l'utilisation. Ils ont été testés à des concentrations variant de 10 à 420 $\mu\text{g/ml}$, en estimant que chaque puits contient de 2 à 5 mg d'extrait d'invertébrés et que de 0,5 à 5% de cette masse est constituée de composés potentiels. Chaque puits a été solubilisé dans le volume de DMSO requis pour effectuer cinq essais par puits et si nécessaire, un autre essai. Pour certains extraits, la solubilisation était facilement réalisable dans ce volume minimal de DMSO alors que d'autres

extraits possédaient des particules insolubles. Afin d'éliminer les particules insolubles, les plaques furent centrifugées.

Le criblage fut entrepris en plaques de 96 puits à l'aide d'une station de pipetage automatisée BioMek NX^P. La station de pipetage prépare une colonne de huit puits où les essais enzymatiques seront effectués. La plaque est ensuite transférée dans une lectrice à plaque où des mesures d'absorbance à 340 nm en fonction de la durée de la réaction sont obtenues. Ce cycle est répété pour les 12 colonnes de la plaque. Le pipetage fut programmé de sorte à ne toucher que la phase liquide des extraits. Malgré ces précautions, certains extraits précipitaient durant les heures suivant leur solubilisation ou lors de l'essai. Dans ces cas, le signal d'absorbance de l'essai cinétique (décrit ci-dessous) affichait une saturation.

Afin d'évaluer la qualité du criblage, le facteur Z' a été calculé selon l'équation (1)(72, 73, 75):

$$Z' = 1 - \frac{(3\sigma_{CTL+} - 3\sigma_{CTL-})}{|\bar{\chi}_{CTL+} - \bar{\chi}_{CTL-}|} \quad (1)$$

où σ_{CTL+} et σ_{CTL-} sont les écarts types des contrôles positifs et des contrôles négatifs respectivement, alors que $\bar{\chi}_{CTL+}$ et $\bar{\chi}_{CTL-}$ sont les moyennes des contrôles positifs et des contrôles négatifs respectivement. Lorsque la différence des écarts types ($3\sigma_{CTL+} - 3\sigma_{CTL-}$) tend vers zéro, donc avec peu de variation, et/ou que le terme $|\bar{\chi}_s - \bar{\chi}_{CTL}|$ tend vers l'infini, donc avec une grande différence entre les données et les contrôles, le facteur Z' s'approche de la valeur de 1. On dit alors que l'essai est idéal(73). Pour une valeur se situant entre 1 et 0,5 inclusivement ($1 > Z' \geq 0,5$), l'essai de criblage est considéré excellent(73). Une valeur entre 0,5 et 0 indique que la variation des signaux est superposée sur la variation des contrôles et ce type d'essai repose sur une réponse qualitative de type oui/non(73).

Lors d'essais d'inhibition d'une librairie, il est attendu que la majorité des molécules seront inefficaces ou auront peu d'effet. Leurs moyennes et leurs écarts-types seront près de ceux du contrôle positif(73). L'inverse est aussi applicable pour le criblage d'une librairie à la recherche d'activateurs, leur moyenne et écart type seront plus près de ceux des contrôles négatifs(73).

5.3 . Résultats

5.3.1 . Détermination du facteur Z'

Le facteur Z' a été déterminé pour le criblage de la librairie des extraits d'invertébrés marins avec l'enzyme DfrB1. Le contrôle positif détermine l'activité maximale de la DfrB1(73, 75). Le contrôle négatif est défini comme le plus faible signal résultant de l'essai sans l'enzyme. Chaque colonne d'une plaque de 96 puits comporte soit des contrôles négatifs ou des contrôles positifs en alternance. Ce protocole permet également de confirmer la stabilité de l'enzyme au cours du criblage. La valeur calculée du facteur Z' est de 0,62 (Figure 5-1). L'essai est considéré comme excellent(71). La variation entre les contrôles positifs est de $\pm 10\%$ et entre les contrôles négatifs de $\pm 3\%$. Une régression linéaire des moyennes des activités enzymatiques relatives par colonne démontre une pente correspondant à une perte d'activité de 1,2% entre chaque cycle de lecture d'absorbance (1 colonne). À la fin du criblage (12 cycles de lecture), l'activité enzymatique résiduelle serait en moyenne 85,6%. Parmi les 48 contrôles positifs, un seul possède une activité enzymatique faible (60,6%). Il est probable que ceci soit attribuable à une erreur technique telle le pipetage d'une bulle d'air.

L'activité enzymatique dépend également de la stabilité des substrats au cours du criblage. En effet, si ceux-ci se dégradent, l'activité enzymatique s'en verra réduite. Ainsi, la stabilité du DHF et du NADPH a été évaluée selon leur absorbance dans chaque colonne, la première colonne servant de référence au temps zéro (Figure 5-1). Une régression linéaire des moyennes des absorbances relatives par colonne indique que l'absorbance diminue de 1,9% par colonne, donc d'environ 20% à la douzième colonne de la plaque. Sachant que la concentration initiale des substrats est de 50 μM et qu'elle doit être $\geq 25 \mu\text{M}$ pour que la DfrB1 soit saturée, cette diminution n'aura pas d'impact significatif sur l'activité observée. Ainsi, ces conditions sont excellentes pour procéder au criblage d'extraits d'invertébrés marins.

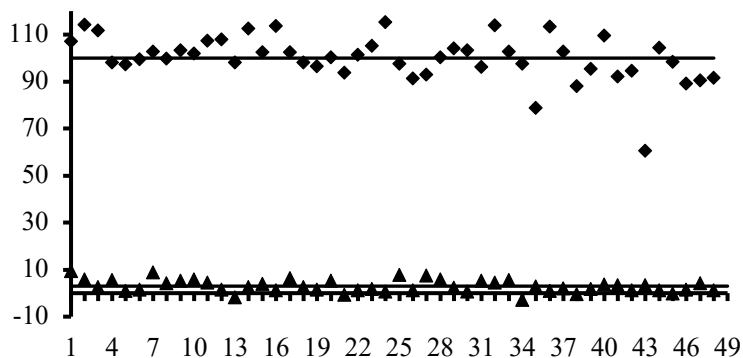


Figure 5-1. Détermination du facteur Z' pour l'essai d'activité dihydrofolate réductase avec la DfrB1.

Les 48 contrôles positifs (CTL+) et 48 contrôles négatifs (CTL-; sans DfrB1) contiennent 10% de DMSO. La moyenne des contrôles négatifs est de $3 \pm 3\%$ (-0.0005 ± 0.0004 abs/min) alors que celle pour les contrôles positifs est de $100 \pm 10\%$ (-0.0166 ± 0.0016 abs/min). La valeur du facteur Z' est de 0,62.

5.3.2 . Criblage des extraits d'invertébrés marins

Les données résultant du criblage de deux plaques (plaques A et B) de 80 extraits d'invertébrés (80 échantillons + 16 contrôles positifs = 96 puits) ont été traitées afin de transformer l'absorbance par min en pourcentage d'activité enzymatique, où la moyenne des contrôles positifs correspond à 100% d'activité. Les plaques furent analysées séparément et les activités relatives (%) normalisées relativement à la moyenne de celles-ci (X_s) et à leur écart type afin d'observer une distribution de ces activités relatives (Figure 5-2). Sachant que la DfrB1 est sensible aux variations de pH, chaque puits a été testé à la fin de l'essai, à l'aide de papier indicateur de pH, pour confirmer le maintien d'un pH de 7. Des résultats interprétables furent obtenus pour 37 et 63 puits des plaques A et B, respectivement. Les 60 autres puits testés ont démontré soit une absorbance de départ trop élevée ou de la précipitation et ils furent exclus des analyses.

Il serait possible de diluer ces extraits d'invertébrés dans un plus grand volume de DMSO afin de diminuer l'absorbance due aux extraits ou d'éviter leur précipitation lors de l'essai. Cependant, cela diminuerait leur effet sur l'activité de la DfrB1. De plus, si les composés des extraits d'invertébrés ne sont pas solubles à faible concentration (10 à 420 $\mu\text{g/ml}$), ceux-ci ne sont pas intéressants puisqu'ils ne respectent pas le critère de solubilité de la Règle de cinq

de Lipinski ni de la Règle de trois(81). Ainsi, près de 63% des échantillons des plaques A et B étaient utilisables pour le criblage.

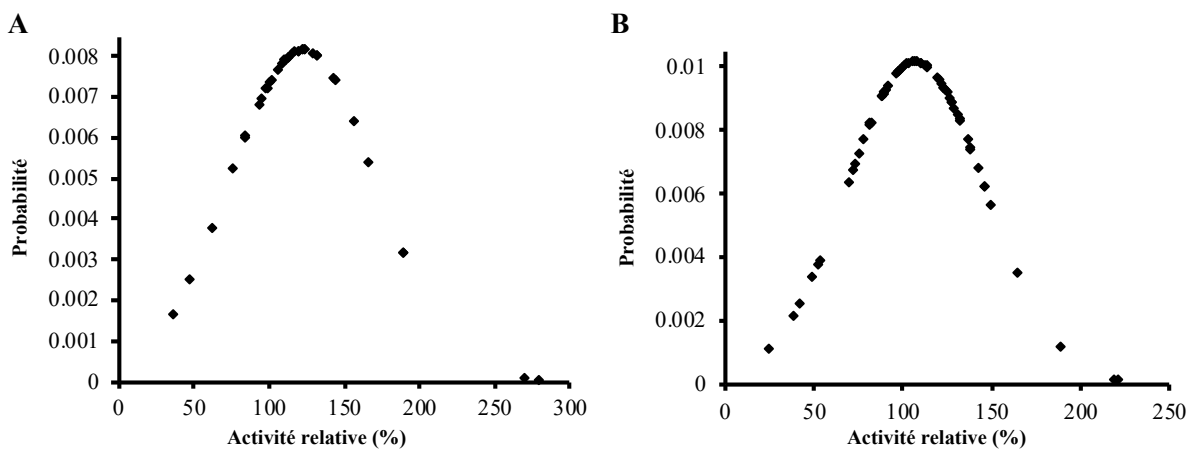


Figure 5-2. Distribution normale de l'activité relative (%) de la DfrB1 mesurée avec les extraits d'invertébrés solubilisés en DMSO pour les plaques A et B.

Pour une distribution normale, le seuil de discrimination entre des molécules effectrices ou des molécules inefficaces peut être déterminé de façon empirique ou statistique. La règle utilisée par *Zhang et al.* basée sur les contrôles positifs est couramment utilisée(67, 69, 71). Cette règle stipule que le seuil de sélection se situe à trois fois l'écart type ($3 \times \sigma_{CTL+}$) par rapport à la moyenne avec un intervalle de confiance de 99,73%(73). Selon cette méthode, le seuil de sélection se situe à trois fois la valeur d'écart type des contrôles positifs ($3 \times \sigma_{CTL+}$; $3 \times 10\%$). Les données de la Figure 5-1 peuvent être utilisées à titre d'exemple comme des données provenant d'échantillons et le seuil de sélection de celles-ci devrait être de 70% ($100\% - 3 \times 10\%$). Un seul parmi les 48 essais enzymatiques effectués (60,6%) se trouve en deçà de ce seuil. Par la méthode de *Zhang et al.*, trois puits ont été identifiés comme positifs pour la plaque A alors que pour la plaque B, il y en a sept. Pour diminuer le nombre de faux positifs, tous les puits devraient être testés une deuxième fois au cours d'une seconde expérience indépendante.

Cette méthode pourrait être aussi utilisée pour identifier des activateurs donc, en cherchant des molécules qui augmentent l'activité enzymatique. Dans ce cas, le seuil de sélection serait de 130% ($100\% + 3 \times 10\%$).

Les puits positifs peuvent contenir soit un inhibiteur fort en faible concentration ou un faible inhibiteur en quantité importante. Après plusieurs cycles de fractionnement, lorsqu'il y aura seulement un composé par échantillon, il sera alors possible de déterminer le potentiel de

celui-ci et ainsi, classer ces inhibiteurs selon leur puissance. Différentes considérations peuvent compliquer l'analyse de ce type d'essai. Par exemple, il est possible que certains puits contiennent à la fois un inhibiteur et un activateur menant donc à peu d'effet mesurable sur l'activité globale de l'enzyme. Dans ce cas-ci, ces puits sont dits des faux négatifs. Un second cas où nous pourrions obtenir un faux positif peut être illustré par la présence d'un composé dénaturant protéique menant à une activité enzymatique réduite, d'où la nécessité d'effectuer des essais subséquents confirmant ces inhibiteurs. Tous ces désavantages sont souvent rencontrés dans le contexte d'un criblage à haut débit.

5.4 . Conclusion

La stabilité de la DfrB1 et celle des substrats ont démontré que le criblage à haut débit est réalisable. La faible solubilité de ces extraits d'invertébrés dans les conditions aqueuses de réalisation de l'essai enzymatique diminue toutefois le nombre d'échantillons pouvant être testés. Ceci réduit donc la probabilité de découvrir un inhibiteur de la DfrB1. Malgré l'identification de quelques extraits diminuant l'activité de la DfrB1, ces extraits nécessitent plus d'analyse afin de confirmer qu'il ne s'agit pas de faux positifs. La méthode étant maintenant établie, il sera possible de cribler d'autres extraits d'invertébrés. Malheureusement, par manque de temps, le criblage de la librairie de 2 000 extraits n'a pas été complété.

Chapitre 6 - Discussion et conclusion

La résistance aux antibiotiques survient naturellement par mutation génétique chez les microorganismes. Cependant, une accélération de ce processus causée par l'utilisation inadéquate des antibiotiques est observée. Cette utilisation inadéquate se produit, entre autres, en milieu clinique où l'antibiotique est prescrit lors d'une infection virale où il n'est pas indiqué alors qu'en milieu d'élevage, les antibiotiques sont utilisés à des fins prophylactiques(23). Le transfert des souches résistantes aux antibiotiques peut se produire soit de l'animal à l'humain, soit d'un humain à un autre(19, 21). Les coûts de soins de santé associés aux patients ayant une infection résistante aux antibiotiques sont supérieurs aux coûts des patients ayant une infection sans résistance aux antibiotiques. Cette différence entre les coûts de soins de santé est due à la longue durée de la maladie, aux tests supplémentaires ainsi qu'à l'utilisation de médicaments plus coûteux(23). Ainsi, la résistance aux antibiotiques est une préoccupation majeure en santé publique mondiale.

Il est estimé que l'appendicectomie, la césarienne ou de simples chirurgies ne seront plus aussi anodines(23). Plusieurs autres progrès médicaux dépendent des antibiotiques tels que la chimiothérapie et la transplantation d'organes(82). Il devient impératif de développer de nouveaux antibiotiques et ce, en maintenant les recommandations de l'OMS afin de ralentir le développement de résistances aux antibiotiques.

6.1 . La prévalence et l'émergence des gènes *dfrBs*

Comme mentionné au chapitre 1, il existe plusieurs mécanismes de résistance chez les microorganismes pour contrer l'effet d'un antibiotique. Un de ces mécanismes consiste à muter le gène de la cible de l'antibiotique(4). C'est le mécanisme dominant lors de la résistance au TMP où le gène *dfrA*, une version mutée du gène *dfr* chromosomal, est retrouvé sur des plasmides transférables horizontalement entre microorganismes(6, 83). Cette observation provient de résultats d'amplification par PCR avec des amorces complémentaires aux gènes *dfrAs*; les gènes *dfrBs* ne sont pas détectés puisqu'on n'utilise pas d'amorces qui y sont complémentaires (83, 84). Au chapitre 2, il a été souligné que les gènes *dfrBs* ont été répertoriés avec des amorces complémentaires aux intégrons, mais cette pratique est exceptionnelle et résulte donc en la sous-estimation de leur prévalence qui demeure inconnue. Afin d'obtenir une évaluation adéquate de leur prévalence, la recherche directe des gènes *dfrBs* devrait être faite non seulement sur les

échantillons cliniques résistants au TMP, mais aussi sur les échantillons alimentaires en provenance de bétail et d'aquaculture. Ainsi, il serait possible de déterminer la prévalence et l'émergence des gènes *dfrBs* en milieu clinique et en milieu alimentaire, mais également d'identifier le plasmide porteur des gènes. L'identification de ce plasmide permettrait de définir les autres gènes de résistance entourant les gènes *dfrBs* telle qu'obtenue avec le gène *dfrB4* retrouvé dans un plasmide multi-résistant.

Il a été aussi avancé, au chapitre 2, que l'incorporation du gène *dfrB4* dans un plasmide multi-résistant est récente. Cette affirmation est basée sur les résultats de la recherche de l'alignement basique local ou *Basic Local Alignment Search Tool* (BLAST) démontrant, jusqu'à récemment, l'existence du plasmide multi-résistant majoritairement sans le gène *dfrB4*. Il existe toutefois une autre preuve indiquant sa récente incorporation : l'élément génétique voisin du gène *dfrB4* est le site *attI* (Figure 2-1). Il faut se rappeler que le site *attI* est l'emplacement principal pour l'insertion des nouvelles cassettes de gènes. D'ailleurs, le gène *dfrB4* est retrouvé adjacent au gène *SulI* (Figure 2-1), gène exprimant une résistance aux antibiotiques prescrits en combinaison avec le TMP, les sulfamides. Cette grande proximité entre les gènes *SulI* et *dfrB4* augmente la possibilité que ceux-ci soient excisés puis épissés ensemble d'un plasmide multi-résistant à un autre. Bref, cette émergence est inquiétante dans l'optique de l'augmentation de la résistance au TMP.

6.2 . La résistance au TMP transposée aux DfrBs

Il a aussi été démontré, au chapitre 2, que la DfrB4 peut contribuer à la résistance au TMP dans l'échantillon clinique. Cet échantillon clinique possède une concentration minimale inhibitrice au TMP (MIC^{TMP}) de $> 16 \mu\text{g/mL}$, ce qui est supérieur au seuil de détermination de résistance du TMP ($> 2 \mu\text{g/mL}$)(85). L'échantillon clinique peut contenir plusieurs mécanismes de résistance au TMP telle la *dfrB4* et la distinction n'a pas pu être établie. Toutefois, la résistance au TMP apportée par la DfrB4 dans les conditions expérimentales *in vitro* a été évaluée à $> 600 \mu\text{g/mL}$, donc supérieure à la limite de solubilité du TMP, suggérant sa contribution à la résistance clinique. Cette valeur n'est pas démesurée, sachant que la MIC^{TMP} pour la DfrB2 est de $3\ 000 \mu\text{g/mL}$ (86).

Dans ces essais, un système artificiel d'expression de la His₆-DfrB4 fut utilisé, qui pourrait surestimer la MIC^{TMP} . Ce système artificiel a été utilisé puisque l'accès à l'échantillon clinique

n'a pas été possible et que le niveau d'expression basale de la DfrB4 de cet échantillon clinique n'est pas connu. La valeur MIC^{TMP} de His₆-DfrB4 est près de 350 fois plus élevée que la concentration plasmatique du TMP de 1,72 µg/mL(87). Ceci indique qu'il faut augmenter la posologie de plus de 350 fois afin de contrer la résistance causée par système artificiel d'expression de la His₆-DfrB4 alors que plus de neuf fois la posologie serait nécessaire pour contrer la résistance dans l'échantillon clinique (> 16 µg/mL).

La contribution dangereuse de la DfrB4 dans la résistance au TMP, montrée au chapitre 2, est alarmante. Les valeurs de MIC^{TMP} pour la DfrB2 ainsi que pour la His₆-DfrB4 soutiennent une résistance au TMP au niveau cellulaire(88, 89). Il a été rapporté, aux chapitres 3 et 4, la similarité des K_i^{TMP} des DfrBs étudiées démontrant leur très faible sensibilité au TMP en milieu *in vitro*. En transposant les valeurs de K_i^{TMP} aux valeurs de MIC^{TMP}, nous pouvons avancer que les bactéries possédant une des DfrBs peuvent être dangereusement résistantes au TMP. En combinant les résultats de leur émergence et de leur dangerosité, nous pouvons prévoir leur perniciosité, préjudiciable pour la santé.

6.3 . Les inconvénients présagés des inhibiteurs

La thérapie combinatoire avec plusieurs médicaments permet de ralentir l'émergence de bactéries résistantes, mais elle possède plusieurs inconvénients majeurs(90). Une approche alternative utilisée pour contrer la prise de plusieurs médicaments est le développement d'un seul inhibiteur qui atteint deux cibles thérapeutiques, un inhibiteur double. Le chapitre 4 a démontré que la molécule de type bisubstrat **4.1** est un inhibiteur double ciblant à la fois les DfrBs et la HPPK. Ce type d'inhibiteur permet de diminuer le nombre d'antibiotiques nécessaires pour guérir une infection tout en ralentissant le développement d'une résistance. Quant aux inhibiteurs bisbenzimidazoles n'ayant pas d'autres cibles connues que les DfrBs, ils possèdent une plus forte affinité pour les DfrBs tout en affectant peu la hDfr, démontré aux chapitres 3 et 4. De plus, ils apportent une diversité structurelle dans la librairie d'inhibiteurs déjà connus de la DfrB1(44). Toutefois, les parties ptérine et adénosine des inhibiteurs de type bisubstrat ainsi que la partie benzimidazole des inhibiteurs bisbenzimidazoles laissent présager une faible sélectivité.

D'abord, la partie ptérine des inhibiteurs de type bisubstrat est retrouvée dans les substrats des enzymes de la voie de biosynthèse du folate(5, 13). L'inhibition des enzymes de la voie de biosynthèse du folate exclusives aux procaryotes pourrait augmenter la valeur des inhibiteurs de

type bisubstrat en agissant sur un large spectre. Malgré cet avantage, il est possible que les enzymes eucaryotes de la voie de biosynthèse du folate soient aussi inhibées. Par ailleurs, il est envisageable que les nouveaux inhibiteurs bisbenzimidazoles des DfrBs optimisés inhibent aussi les enzymes eucaryotes de la voie de biosynthèse du folate dont la hDfr. Des inhibiteurs prometteurs de la hDfr ont été rapportés contenant la partie benzimidazole dans l'optique de développer un médicament anticancer(91). Il serait inacceptable pour un antibiotique d'être peu sélectif en affectant les enzymes du patient.

Lors de l'optimisation des inhibiteurs des DfrBs, leur sélectivité devra être considérée telle qu'effectuée au chapitre 3 où la sélectivité des inhibiteurs bisbenzimidazoles de la DfrB1 a été vérifiée en les testant en parallèle avec la hDfr. Les inhibiteurs des DfrBs peuvent être aussi testés sur la *EcDfr* puisque son inhibition augmenterait la qualité de l'antibiotique par l'élimination du TMP. Un antibiotique approprié aurait peu d'effet sur la hDfr tout en inhibant favorablement la *EcDfr*. Il est à noter que l'antibiotique ciblant les DfrBs sans inhiber la *EcDfr* devra être co-administré avec le TMP.

Les inhibiteurs bisbenzimidazoles des DfrBs sont constitués de l'hétérocycle benzimidazole retrouvé aussi dans la structure de plusieurs agents thérapeutiques ayant des propriétés antivirale, anticancer ou encore anti-inflammatoire(92). Cet hétérocycle est un analogue aux purines telle que l'adénine, qui est incluse dans la partie adénosine des inhibiteurs de type bisubstrat. La présence du benzimidazole ainsi que de l'adénosine peut affecter la sélectivité des inhibiteurs des DfrBs. Effectivement, l'adénosine, et par conséquent l'adénine, se retrouve, entre autres, dans les cofacteurs NADPH (Figure 1-2) et flavine adénine dinucléotide (FAD) (Figure 6-1) ainsi que dans toutes molécules contenant des ribonucléosides dont l'ADN et l'ARN. L'omniprésence des cofacteurs NADPH et FAD fait en sorte que leurs analogues tels que les inhibiteurs des DfrBs peuvent lier les enzymes de la classe des oxydoréductases dont les protéines de la chaîne respiratoire mitochondriale. Il faut se rappeler que l'inhibition de la chaîne respiratoire mitochondriale interrompt la formation d'ATP. De plus, l'inhibition des protéines interagissant avec l'ADN et l'ARN pourrait engager de graves conséquences au niveau génétique de la cellule. Toutes ces effets néfastes présagés devront être pris en considération une fois une tête de série identifiée. Des tests *in vivo* sur les cellules humaines avec cette tête de série inhibant les DfrBs pourront déterminer son effet sur les enzymes de la voie de biosynthèse du folate, mais également au niveau génétique de la cellule. Des modifications de cette tête de série pourront être effectués

afin de diminuer les effets néfastes, s'il y a lieu. De plus, une affinité élevée, soit dans le nanomolaire ou picomolaire, de la tête de série pour les DfrBs pourrait diminuer les liaisons non-spécifiques et donc, les effets néfastes.

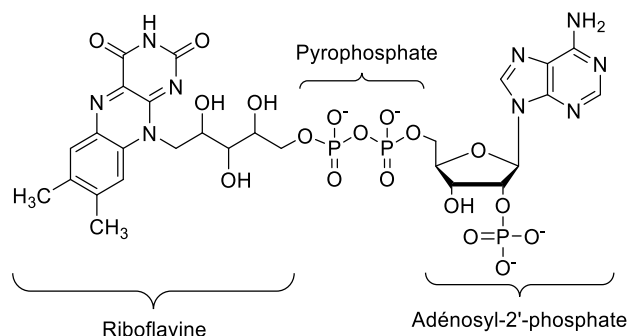


Figure 6-1. Structure du FAD; les parties importantes sont identifiées.

6.4 . L'optimisation des inhibiteurs de type bisubstrat

Nous avons présenté au chapitre 3 une étude de liaison sur les inhibiteurs à base de benzimidazoles avec la DfrB1. Nos expériences ont été appuyées par de courtes simulations de dynamique moléculaire (LowModeMD) effectuées sur la structure cristalline de la DfrB1. La méthode LowModeMD a été utilisée pour pallier les limites de la structure cristalline partielle de l'inhibiteur **3.3** avec la DfrB1 (Figure 6-2). L'hypothèse émise est que ceci provient du dynamisme de l'inhibiteur exposé au solvant tel que vu avec l'extrémité 4'-aminobenzoylglutamate du DHF (Figure 1-2)(45). Il a été rapporté que ce dynamisme est important dans la catalyse de l'enzyme puisque le changement de liaison entre les deux K32 situées à la même ouverture du tunnel favorise la formation de l'état de transition(43). Les simulations suggèrent que les K32 lient les carboxylates de l'inhibiteur **3.3** en alternance, similairement à l'extrémité 4'-aminobenzoylglutamate du DHF.

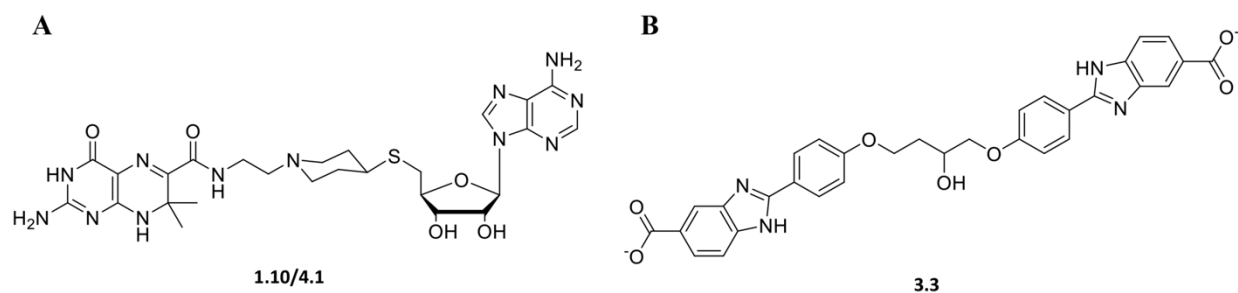


Figure 6-2. Structures de l'inhibiteur de type bisubstrat 1.10/4.1 et de l'inhibiteur non-symétrique de type bisbenzimidazole 3.3.

Quant au développement d'une tête de série, l'optimisation des inhibiteurs de type bisubstrat peut être guidée par la connaissance des points de liaison avec la DfrB1. Afin d'identifier ces points de liaison, nous avons tenté de cristalliser la DfrB1 avec **4.1** (Figure 6-2), sans succès. Une hypothèse a donc été émise stipulant que la molécule **4.1** lie la DfrB1 en coopérativité avec le NADPH. En effet, la molécule **4.1** possède la partie ptérine du DHF qui est connu pour lier en coopérativité avec la partie nicotinamide du NADPH au centre du site active de la DfrB1(45). L'ajout du NADPH et de la molécule **4.1** au cristal de la DfrB1 a conduit à une densité électronique observée dans le site actif associée à deux molécules de NADPH. Ceci peut découler du fait que la DfrB1 possède une meilleure affinité pour deux molécules de NADPH (K_{d1} 2.5 μ M et K_{d2} 96 μ M) que pour l'inhibiteur ($K_i = 134 \mu$ M)(44). Un ajustement des concentrations de NADPH et de la molécule **4.1** pourrait permettre la cristallisation de ceux-ci avec la DfrB1. Cependant, une fois la cristallisation réussie, il se peut que celle-ci ne soit pas suffisante pour observer la liaison de la molécule **4.1** avec la DfrB1, tel qu'est le cas des structures cristallines incomplètes obtenues au chapitre 3 et précédemment (45).

Pour contourner la possibilité d'obtenir des structures cristallines incomplètes, il serait préférable de procéder par les méthodes d'amarrage moléculaire et de LowModeMD basées sur les résultats du mode de liaison obtenus *in vitro*. La caractérisation du mode de liaison *in vitro* validerait le type d'inhibition par le graphique de Dixon, la coopérativité de liaison ainsi que le nombre de molécules de **4.1** par site actif par le graphique de Hill. En effet, tout comme le NADPH, le DHF ou les inhibiteurs de type benzimidazole, plus qu'une molécule de **4.1** pourrait se lier par tétramère de DfrB, avec ou sans coopérativité de liaison. Ces résultats du mode de liaison guideraient l'amarrage moléculaire de la molécule **4.1** sur la DfrB. Le modèle de départ sera nécessairement la structure cristalline de DfrB1•NADP⁺•ptérine^{DHF} (PDB 2RK1)(45). La présence du NADP⁺ et la ptérine^{DHF} est importante afin de fixer une des deux molécules pour l'amarrage moléculaire de l'inhibiteur puisque celle-ci ne peut être déterminée pour plus qu'une molécule à la fois.

Dans le premier cas, où la molécule **4.1** possède une coopérativité de liaison avec le NADPH et qu'elle est compétitive, la ptérine^{DHF} sera éliminée de la structure cristalline DfrB1•NADP⁺•ptérine^{DHF} générant le modèle DfrB1•NADP⁺ avec un site de liaison libre. L'amarrage moléculaire de la molécule **4.1** sera exécutée puis soumis à une simulation LowModeMD telle que réalisée au chapitre 3. Dans le deuxième cas, où deux molécules **4.1** se

lient par site actif, le NAPD^+ sera enlevé de la structure cristalline afin que le site actif soit occupé que par la ptérine^{DHF} résultant du modèle DfrB1•ptérine^{DHF}. La ptérine^{DHF} mimera la présence d'une molécule **4.1**. L'amarrage moléculaire ainsi que la LowModeMD seront effectués pareillement au premier cas présenté. L'amarrage moléculaire de la deuxième molécule **4.1** sera répété avec le modèle DfrB1•**4.1**, par le retrait de la ptérine^{DHF} du modèle. Dans le cas où seulement une molécule de **4.1** se lie à la DfrB1 et ce, peu importe le type d'inhibition, l'amarrage moléculaire ainsi que la LowModeMD seront accomplis avec la structure cristalline de l'apo-DfrB1. Les conformations de la molécule **4.1** engendrées telles que présentés au chapitre 3, indiqueront les points de liaison entre celle-ci et la DfrB1, permettant de guider l'optimisation de cet inhibiteur.

Les structures cristallines de la DfrB1, dont celles présentées au chapitre 3, comportent seulement son cœur de baril β puisque la cristallisation de celle-ci nécessite le clivage post-traductionnel du terminus amine(45). La ressemblance entre les cœurs de baril β des DfrBs démontrée au chapitre 4 suggère que seule la structure cristalline de la DfrB1 est suffisante pour les études structurales avec la molécule **4.1**.

6.5 . L'optimisation des inhibiteurs de type bisbenzimidazole

L'inhibiteur **3.12** est considéré comme un des meilleurs inhibiteurs de la DfrB1, toutes librairies confondues. Cet inhibiteur sélectif et compétitif possède un faible K_i ($7.4 \pm 2.7 \mu\text{M}$) et forme un complexe binaire DfrB1•**3.12**, contrairement aux autres inhibiteurs bisbenzimidazole qui se lient en paire à l'enzyme. De plus, cette molécule **3.12** inhibe toutes les DfrBs avec des constantes d'inhibition dans le bas μM (Tableau 4-V). Malgré ceci, l'affinité de l'inhibiteur **3.12** n'est pas suffisante pour servir de tête de série où une affinité de l'ordre du nM est visée.

L'étude de sa liaison avec la DfrB1 exposée au chapitre 3 indique la possibilité de modifier certaines parties de cet inhibiteur afin d'y augmenter son affinité. En effet, le fragment acide 1H-benzimidazole-5-carboxylique est d'une grande importance dans la liaison des inhibiteurs bisbenzimidazoles avec la DfrB1 puisque son retrait abolit l'inhibition. Les carboxylates sont à éviter dans les médicaments étant donné le peu de sélectivité causée par leur reconnaissance dans différentes voies de biosynthèse de la cellule. Des groupements fonctionnels alternatifs aux carboxylates (Figure 6-3; **6.1**) procurant ou non des donneurs et accepteurs d'hydrogène ont été introduits dans plusieurs molécules au chapitre 3 (Figure 6-3; groupements **6.2** à **6.5**). Malheureusement, ces modifications n'ont pas permis de maintenir le caractère inhibiteur.

Toutefois, il existe plusieurs bioisostères rapportés pour les carboxylates comme le tétrazole (6.6), l'acide sulfonique (6.7), l'acide hydroxamique (6.8) et le 1,2,4-oxadiazole (6.9)(93, 94). En remplaçant les carboxylates par un de ses bioisostères, les nouvelles molécules pourraient inhiber les DfrBs et ce, avec une affinité semblable ou plus élevée(93).

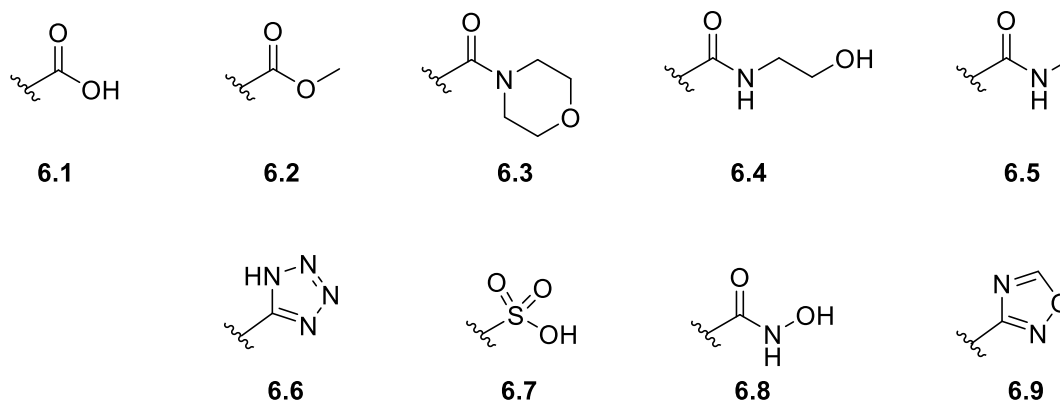


Figure 6-3. Structures des groupements fonctionnels alternatifs aux carboxylates (6.1).
La partie supérieure représente les groupements fonctionnels testés sur la DfrB1 (6.2 à 6.5). La partie inférieure montre des exemples de bioisostères : le tétrazole (6.6), l'acide sulfonique (6.7), l'acide hydroxamide (6.8) et le 1,2,4-oxadiazole (6.9).

Il a été aussi présenté au chapitre 3 que les changements dans l'espaceur des inhibiteurs bisbenzimidazoles sont tolérés étant donné le large tunnel composant le site actif. L'introduction de groupements donneurs et accepteurs d'hydrogène dans l'espaceur du 3.12 pourrait augmenter le nombre de points de liaison avec les résidus de la région VQIY. Les chaînes principales des résidus de cette région interagissent avec les ligands connus de la DfrB1. De plus, la conformation allongée du 3.12 dans le site actif suggère une possibilité de modifier l'espaceur (Figure 3-6). Il est possible de maintenir cette conformation en introduisant une rigidité, comme avec l'analogue *cis* de la molécule 3.10 (Figure 6-4). Nous notons que les résultats des courtes simulations stipulent que la double liaison *trans* du 3.10 ne participe pas à la liaison avec la DfrB1. Ainsi, cette nouvelle double liaison *cis* introduite dans l'espaceur ne servirait qu'à maintenir la conformation allongée du 3.12 sans toutefois contribuer directement à la liaison avec la DfrB1. Bref, l'optimisation de l'espaceur et des carboxylates de l'inhibiteur 3.12 pourrait augmenter l'affinité.

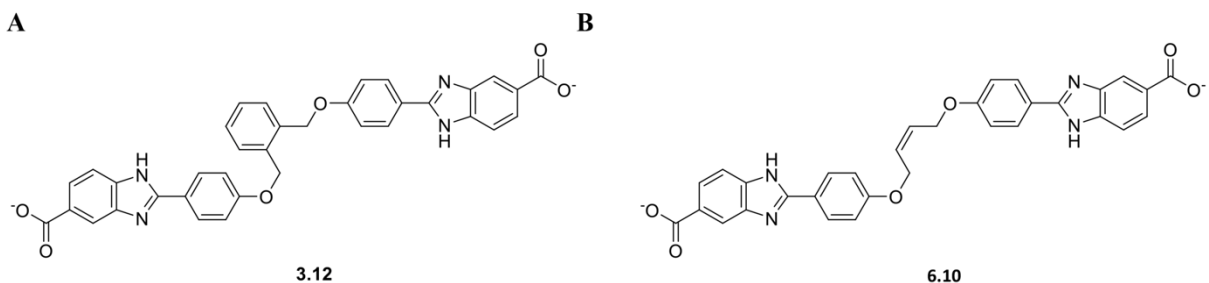


Figure 6-4. Structures de l'inhibiteur 3.12 et de la molécule 5.10.

Comme expliqué à la section précédente, la DfrB1 peut être utilisée comme modèle structurel pour les DfrBs sachant que seul le baril β se retrouve dans la structure cristalline et qu'il est fortement conservé auprès des DfrBs. Ainsi, l'optimisation des inhibiteurs basée sur la structure de la DfrB1 pourra fournir des inhibiteurs potentiels aux DfrBs.

6.6 . La recherche d'inhibiteurs allostériques ciblant l'interface protéine-protéine

Le volumineux site actif des DfrBs, facilement accessible au solvant, rend difficile la découverte de ligands ayant une forte affinité. Afin d'augmenter la probabilité d'inhiber les DfrBs, une autre stratégie a été entamée en collaboration avec le Dr. David Ryan Koes de l'Université de Pittsburg. Lorea Alejandre, étudiante au doctorat au laboratoire Pelletier, poursuit la recherche d'inhibiteurs allostériques en ciblant l'interface dimère-dimère. Ces inhibiteurs allostériques mimeront l'interaction des résidus importants dans l'interface protéine-protéine des monomères assemblés en tétramère DfrB. Un premier criblage *in silico* de bibliothèques de molécules fut effectué par le laboratoire du Dr. Koes qui se concentre sur le développement de nouveaux algorithmes dans le criblage d'inhibiteurs.

Cibler les interfaces protéine-protéine exige une connaissance de la structure de la protéine ainsi que de cette interface ce qui a été étudiée pour la DfrB1. Comme expliqué au chapitre 1, la multimérisation de la DfrB1 se fait en deux étapes distinctes: deux monomères s'associent afin de former un dimère puis deux dimères se lient pour créer la forme active de la DfrB1, soit le tétramère(16, 40). L'association des monomères requière l'entrelacement de leurs brins β alors que les dimères utilisent les H62 situées sur les boucles non-structurées. Malgré le manque d'informations structurelles pour les DfrBs, leur similarité de séquence près des H62 montrée au

chapitre 4 suppose qu'elles partageraient la même l'interface dimère-dimère. En utilisant la DfrB1 comme modèle, il est espéré que les inhibiteurs découverts lieront aussi les autres DfrBs. Ainsi, le criblage *in silico* de la DfrB1 suivi par les essais enzymatiques *in vitro* pourront mener à la découverte d'inhibiteurs allostériques pour les DfrBs.

Un désavantage possible de ces inhibiteurs allostériques est leur taille volumineuse requise pour suffisamment couvrir et prévenir la formation de l'interaction protéine-protéine. En effet, les interfaces protéine-protéine sont généralement larges (1 000 – 2 000 Å²) et plates comparées aux cavités de liaison des inhibiteurs habituels (300 – 500 Å²)(95). Il est suggéré que l'interface protéine-protéine doit être au minimum de 500 Å² pour maintenir l'intégrité de l'assemblage du dimère(95).

Les résultats préliminaires du criblage *in silico* de la DfrB1 démontrent des molécules avec un enrichissement de groupements hétérocycliques à cinq membres contenant le carbone ou l'azote. Le groupement prépondérant est le triazole.

Les 12 composés possédant le meilleur score *in silico* ont été sélectionnés pour les essais enzymatiques *in vitro* avec la DfrB1 par Lorea Alejandre. La faible solubilité aqueuse de dix parmi ces composés (20-500 µM dans les conditions expérimentales *in vitro*) a limité les concentrations utilisées pour la plupart des composés. Aucune des 12 molécules n'inhibait la DfrB1. Ceci permet de conclure que le criblage *in silico* n'a pas mené à de meilleurs inhibiteurs que ceux précédemment découverts dont ceux du chapitre 3(18, 44). Ainsi pour le prochain criblage *in silico*, une attention particulière à la solubilité prédite des molécules sera de mise.

6.7 . Conclusion

Le TMP demeure un antibiotique important en milieu clinique ou d'élevage. Cet antibiotique est une possibilité envisageable pour traiter les infections aux MRSA et *Mycobacterium tuberculosis* multi-résistant(9, 96). Globalement, cette thèse apporte un regard sur l'importance des DfrBs dans la résistance au TMP. Les résultats de cette thèse contribuent des connaissances sur la prévalence des gènes *dfrBs* ainsi que la cinétique et l'inhibition des enzymes DfrBs. Le portrait de la prévalence des gènes *dfrBs* émergents reste incomplet et nécessite d'être évalué dans plus d'une librairie d'échantillons en provenance de milieu clinique ou d'élevage. La dangerosité des DfrBs ne fait plus aucun doute avec leur faible sensibilité au TMP et ce, malgré leur faible capacité à catalyser la réduction du DHF. Cette thèse propose des pistes de solution

pour inhiber les DfrBs afin de pallier la résistance au TMP. La compréhension de l'inhibition de la DfrB1 par différents types d'inhibiteurs présentée pourrait être transférable aux autres DfrBs et ainsi, développer un antibiotique permettant de vaincre la résistance au TMP causée par les DfrBs.

Bibliographie

1. Davies J, Davies D. Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews*. 2010;74:417-33.
2. Landecker H. Antibiotic Resistance and the Biology of History. *Body & Society*. 2016;22:19-52.
3. Martínez JL. Effect of antibiotics on bacterial populations: a multi-hierarchical selection process. *F1000Research*. 2017;6:51.
4. Debabov D. Antibiotic resistance: Origins, mechanisms, approaches to counter. *Applied Biochemistry and Microbiology*. 2013;49:665-71.
5. Bermingham A, Derrick JP. The folic acid biosynthesis pathway in bacteria: Evaluation of potential for antibacterial drug discovery. *BioEssays*. 2002;24:637-48.
6. Huovinen P, Sundstrom L, Swedberg G, Skold O. Trimethoprim and sulfonamide resistance. *Antimicrobial Agents and Chemotherapy*. 1995;39:279-89.
7. Masters PA, O'Bryan TA, Zurlo J, Miller DQ, Joshi N. Trimethoprim-Sulfamethoxazole Revisited. *Archives of Internal Medicine*. 2003;163:402.
8. Close SJ, McBurney CR, Garvin CG, Chen DC, Martin SJ. Trimethoprim-sulfamethoxazole activity and pharmacodynamics against glycopeptide-intermediate *Staphylococcus aureus*. *Pharmacotherapy*. 2002;22(8):983-9.
9. Palomino JC, Martin A. The potential role of trimethoprim-sulfamethoxazole in the treatment of drug-resistant tuberculosis. *Future Microbiology*. 2016;11(4):539-47.
10. Huovinen P. Trimethoprim resistance. *Antimicrobial Agents and Chemotherapy*. 1987;31:1451-6.
11. Canada Go. Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) 2015 Annual Report. Public Health Agency of Canada,. 2017:Guelph, Ontario.
12. Organization WH, WHO. WHO list of critically important antimicrobials for human medicine (WHO CIA list) 2016 [Available from: <http://www.who.int/foodsafety/publications/cia2016fr2.pdf?ua=1>].
13. Wright DL, Anderson AC. Antifolate agents: a patent review (2006 - 2010). *Expert Opinion on Therapeutic Patents*. 2011;21:1293-308.

14. Dougherty TJ, Pucci MJ. Antibiotic Discovery and Development: Springer US; 2011.
15. Hawser S, Lociuro S, Islam K. Dihydrofolate reductase inhibitors as antibacterial agents. *Biochemical Pharmacology*. 2006;71:941-8.
16. Howell EE. Searching sequence space: Two different approaches to dihydrofolate reductase catalysis. *ChemBioChem*. 2005;6:590-600.
17. Hecht D, Tran J, Fogel GB. Structural-based analysis of dihydrofolate reductase evolution. *Molecular Phylogenetics and Evolution*. 2011;61:212-30.
18. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, et al. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of Medicinal Chemistry*. 2012;55:3182-92.
19. Cabello FC. Heavy use of prophylactic antibiotics in aquaculture: A growing problem for human and animal health and for the environment. *Environmental Microbiology*. 2006;8:1137-44.
20. Luepke KH, Suda KJ, Boucher H, Russo RL, Bonney MW, Hunt TD, et al. Past, Present, and Future of Antibacterial Economics: Increasing Bacterial Resistance, Limited Antibiotic Pipeline, and Societal Implications. *Pharmacotherapy*. 2017;37:71-84.
21. Vieira AR, Collignon P, Aarestrup FM, McEwen SA, Hendriksen RS, Hald T, et al. Association Between Antimicrobial Resistance in Escherichia coli Isolates from Food Animals and Blood Stream Isolates from Humans in Europe: An Ecological Study. *Foodborne Pathogens and Disease*. 2011;8:1295-301.
22. Fleming MP, Datta N, Grüneberg RN. Trimethoprim Resistance determined by R Factors. *British Medical Journal*. 1972;1:726-8.
23. Organization WH, WHO. 2018 [Available from: <http://www.who.int>].
24. Khardori N. Antibiotics - past, present, and future. *Medical Clinics of North America*. 2006;90:1049-76.
25. Reece LJ, Nichols R, Howell EE, Ogden RC. Construction of a Synthetic Gene for an R-Plasmid-Encoded Dihydrofolate Reductase and Studies on the Role of the N-Terminus in the Protein. *Biochemistry*. 1991;30:10895-904.
26. Alonso H, Gready JE. Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends in Microbiology*. 2006;14:236-42.

27. Recchia GD, Hall RM. Gene cassettes: A new class of mobile element. *Microbiology*. 1995;141:3015-27.
28. Gillings M, Boucher Y, Labbate M, Holmes A, Krishnan S, Holley M, et al. The evolution of class 1 integrons and the rise of antibiotic resistance. *Journal of Bacteriology*. 2008;190:5095-100.
29. Gillings MR. Integrons: past, present, and future. *Microbiology and molecular biology reviews*. 2014;78:257-77.
30. Schnell JR, Dyson HJ, Wright PE. Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct*. 2004;33:119-40.
31. Cody V, Pace J, Makin J, Piraino J, Queener SF, Rosowsky A. Correlations of inhibitor kinetics for *Pneumocystis jirovecii* and human dihydrofolate reductase with structural data for human active site mutant enzyme complexes. *Biochemistry*. 2009;48:1702-11.
32. Duff MR, Chopra S, Strader MB, Agarwal PK, Howell EE. Tales of Dihydrofolate Binding to R67 Dihydrofolate Reductase. *Biochemistry*. 2016;55:133-45.
33. Fling ME, Elwell LP. Protein expression in *Escherichia coli* minicells containing recombinant plasmids specifying trimethoprim-resistant dihydrofolate reductases. *Journal of Bacteriology*. 1980;141:779-85.
34. Narayana N. High-resolution structure of a plasmid-encoded dihydrofolate reductase: Pentagonal network of water molecules in the D2-symmetric active site. *Acta Crystallographica Section D: Biological Crystallography*. 2006;62:695-706.
35. Yachnin BJ, Colin DY, Volpato JP, Ebert M, Pelletier JN, Berghuis AM. Novel crystallization conditions for tandem variant R67 DHFR yield a wild-type crystal structure. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 2011;67:1316-22.
36. Bhojane PP, Duff MR, Bafna K, Agarwal P, Stanley C, Howell EE. Small Angle Neutron Scattering Studies of R67 Dihydrofolate Reductase, a Tetrameric Protein with Intrinsically Disordered N-Termini. *Biochemistry*. 2017;56:5886-99.
37. Martinez MA, Pezo V, Marlière P, Wain-Hobson S. Exploring the functional robustness of an enzyme by in vitro evolution. *The EMBO journal*. 1996;15:1203-10.
38. Mayer BJ. SH3 domains: complexity in moderation. *Journal of cell science*. 2001;114:1253-63.

39. Park H, Zhuang P, Nichols R, Howell EE. Mechanistic studies of R67 dihydrofolate reductase. Effects of pH and an H62C mutation. *Journal of Biological Chemistry*. 1997;272:2252-8.
40. Schmitzer AR, Lépine F, Pelletier JN. Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: Creating alternative functional configurations. *Protein Engineering, Design and Selection*. 2004;17:809-19.
41. Carrillo N, Ceccarelli EA, Roveri OA. Usefulness of kinetic enzyme parameters in biotechnological practice. *Biotechnology and Genetic Engineering Reviews*. 2010;27:367-82.
42. Broad DF, Smith JT. Classification of trimethoprim-resistant dihydrofolate reductases mediated by R-plasmids using isoelectric focussing. *European Journal of Biochemistry*. 1982;125:617-22.
43. Kamath G, Howell EE, Agarwal PK. The tail wagging the dog: Insights into catalysis in R67 dihydrofolate reductase. *Biochemistry*. 2010;49:9078-88.
44. Jackson M, Chopra S, Smiley RD, Maynord PON, Rosowsky A, London RE, et al. Calorimetric studies of ligand binding in R67 dihydrofolate reductase. *Biochemistry*. 2005;44:12420-33.
45. Krahn JM, Jackson MR, DeRose EF, Howell EE, London RE. Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry*. 2007;46:14878-88.
46. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. *British journal of pharmacology*. 2011;162:1239-49.
47. Katiyar C, Gupta A, Kanjilal S, Katiyar S. Drug discovery from plant sources: An integrated approach. *Ayu*. 2012;33(1):10-9.
48. Bajorath J. Computer-aided drug discovery. *F1000Res*. 2015;4.
49. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem*. 2014;14(16):1923-38.
50. Lamoree B, Hubbard Roderick E. Current perspectives in fragment-based lead discovery (FBLD). *Essays In Biochemistry*. 2017;61:453-64.
51. Mello JdFre, Gomes RA, Vital-Fujii DG, Ferreira GM, Trossini GHG. Fragment-based drug discovery as alternative strategy to the drug development for neglected diseases. *Chemical Biology & Drug Design*. 2017;38:42-9.

52. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*. 1996;274:1531-4.
53. Joseph-McCarthy D, Campbell AJ, Kern G, Moustakas D. Fragment-based lead discovery and design. *Journal of Chemical Information and Modeling*. 2014;54:693-704.
54. Hajduk PJ, Sheppard G, Nettlesheim DG, Olejniczak ET, Shuker SB, Meadows RP, et al. Discovery of Potent Nonpeptide Inhibitors of Stromelysin Using SAR by NMR. *Journal of the American Chemical Society*. 1997;119(25):5818-27.
55. Hajduk PJ. SAR by NMR: putting the pieces together. *Mol Interv*. 2006;6(5):266-72.
56. Roughley S, Wright L, Brough P, Massey A, Hubbard RE. Hsp90 inhibitors and drugs from fragment and virtual screening. *Top Curr Chem*. 2012;317:61-82.
57. Burlison JA, Neckers L, Smith AB, Maxwell A, Blagg BSJ. Novobiocin: Redesigning a DNA gyrase inhibitor for selective inhibition of Hsp90. *Journal of the American Chemical Society*. 2006;128:15529-36.
58. Cartwright PD, Helfinger RG, Howell JJ, Siepmann KK. Introducing an acute pain service. *Anaesthesia*. 1991;46:188-91.
59. Tamma PD, Cosgrove SE, Maragakis LL. Combination therapy for treatment of infections with gram-negative bacteria. *Clinical Microbiology Reviews*. 2012;25:450-70.
60. Liljefors T, Krogsgaard-Larsen P, Madsen U. *Textbook of Drug Design and Discovery*. Third edition ed: CRC Press; 2003.
61. Shi G, Shaw G, Liang YH, Subburaman P, Li Y, Wu Y, et al. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New design with improved properties. *Bioorganic and Medicinal Chemistry*. 2012;20:47-57.
62. Shi G, Shaw G, Li Y, Wu Y, Yan H, Ji X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New lead exhibits a distinct binding mode. *Bioorganic and Medicinal Chemistry*. 2012;20:4303-9.
63. Shi G, Blaszczyk J, Ji X, Yan H. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: Synthesis and biochemical and crystallographic studies. *Journal of Medicinal Chemistry*. 2001;44(9):1364-71.
64. Blaszczyk J, Shi G, Yan H, Ji X. Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. *Structure*. 2000;8(10):1049-58.

65. Harvey AL. Natural products in drug discovery. *Drug Discov Today*. 2008;13(19-20):894-901.
66. Chandra S, Lata H, Varma A, SpringerLink (Service en ligne). *Biotechnology for Medicinal Plants : Micropropagation and Improvement*. Berlin, Heidelberg: Springer Berlin Heidelberg : Imprint: Springer; 2013. 1 ressource en ligne. p.
67. Tarling CA, Woods K, Zhang R, Brastianos HC, Brayer GD, Andersen RJ, et al. The search for novel human pancreatic alpha-amylase inhibitors: high-throughput screening of terrestrial and marine natural product extracts. *Chembiochem*. 2008;9(3):433-8.
68. Raviña E. *The evolution of drug discovery : from traditional medicines to modern drugs*. Weinheim: Wiley-VCH; 2011.
69. Tysoe C, Williams LK, Keyzers R, Nguyen NT, Tarling C, Wicki J, et al. Potent Human alpha-Amylase Inhibition by the beta-Defensin-like Protein Helianthamide. *ACS Cent Sci*. 2016;2(3):154-61.
70. Smart AL, Gaisford S, Basit AW. Oral peptide and protein delivery: intestinal obstacles and commercial prospects. *Expert Opin Drug Deliv*. 2014;11(8):1323-35.
71. Zoraghi R, Worrall L, See RH, Strangman W, Popplewell WL, Gong H, et al. Methicillin-resistant *Staphylococcus aureus* (MRSA) pyruvate kinase as a target for bis-indole alkaloids with antibacterial activities. *Journal of Biological Chemistry*. 2011;286(52):44716-25.
72. Mayr LM, Bojanic D. Novel trends in high-throughput screening. *Curr Opin Pharmacol*. 2009;9(5):580-8.
73. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen*. 1999;4(2):67-73.
74. Bittker JA, Ross NT. *High Throughput Screening Methods: Evolution and Refinement*: Royal Society of Chemistry; 2016.
75. Seethala R, Fernandes P. *Handbook of Drug Screening*: Taylor & Francis; 2001.
76. Schneider G. Automating drug discovery. *Nat Rev Drug Discov*. 2018;17(2):97-113.
77. Barlow RS, Pemberton JM, Desmarchelier PM, Gobius KS. Isolation and characterization of integron-containing bacteria without antibiotic selection. *Antimicrobial Agents and Chemotherapy*. 2004;48:838-42.

78. Kadlec K, Kehrenberg C, Schwarz S. Molecular basis of resistance to trimethoprim, chloramphenicol and sulphonamides in *Bordetella bronchiseptica*. *Journal of Antimicrobial Chemotherapy*. 2005;56:485-90.
79. Sunde M. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. *Journal of Antimicrobial Chemotherapy*. 2005;56:1019-24.
80. Tennstedt T, Szczepanowski R, Braun S, Puhler A, Schluter A. Occurrence of integron-associated resistance gene cassettes located on antibiotic resistance plasmids isolated from a wastewater treatment plant. *FEMS Microbiol Ecol*. 2003;45(3):239-52.
81. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol*. 2004;1(4):337-41.
82. The evolving threat of antimicrobial resistance, (2012).
83. Kehrenberg C, Schwarz S. *dfrA20*, A novel trimethoprim resistance gene from *Pasteurella multocida*. *Antimicrobial Agents Chemotherapy*. 2005;49(1):414-7.
84. Hu LF, Chang X, Ye Y, Wang ZX, Shao YB, Shi W, et al. *Stenotrophomonas maltophilia* resistance to trimethoprim/sulfamethoxazole mediated by acquisition of *sul* and *dfrA* genes in a plasmid-mediated class 1 integron. *International Journal of Antimicrobial Agents*. 2011;37(3):230-4.
85. Testing ECoAS. International MIC distribution 2018 [Available from: <https://mic.eucast.org/Eucast2/regShow.jsp?Id=4125>].
86. Amyes SG, Smith JT. R-factor trimethoprim resistance mechanism: an insusceptible target site. *Biochemical and Biophysical Research Communications*. 1974;58(2):412-8.
87. Kremers P, Duvivier J, Heusghem C. Pharmacokinetic studies of co-trimoxazole in man after single and repeated doses. *The Journal of Clinical Pharmacology*. 1974;14(2):112-7.
88. Toulouse JL, Edens TJ, Alejaldre L, Manges AR, Pelletier JN. Integron-associated *DfrB4*, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*. 2017;61:1-5.
89. Amyes SG, Smith JT. The purification and properties of the trimethoprim-resistant dihydrofolate reductase mediated by the R-factor, R388. *European Journal of Biochemistry*. 1976;61(2):597-603.

90. Tamma PD, Cosgrove SE, Maragakis LL. Combination therapy for treatment of infections with gram-negative bacteria. *Clinical Microbiology Reviews*. 2012;25(3):450-70.
91. Singla P, Luxami V, Paul K. Triazine-benzimidazole hybrids: anticancer activity, DNA interaction and dihydrofolate reductase inhibitors. *Bioorganic & Medicinal Chemistry*. 2015;23(8):1691-700.
92. Akhtar W, Khan MF, Verma G, Shaquiquzzaman M, Rizvi MA, Mehdi SH, et al. Therapeutic evolution of benzimidazole derivatives in the last quinquennial period. *European Journal of Medicinal Chemistry*. 2017;126:705-53.
93. Beyett TS, Gan X, Reilly SM, Chang L, Gomez AV, Saltiel AR, et al. Carboxylic Acid Derivatives of Amlexanox Display Enhanced Potency toward TBK1 and IKKepsilon and Reveal Mechanisms for Selective Inhibition. *Molecular Pharmacology*. 2018;94(4):1210-9.
94. Lima LM, Barreiro EJ. Bioisosterism: a useful strategy for molecular modification and drug design. *Current Medicinal Chemistry*. 2005;12(1):23-49.
95. Lipinski CA. Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. *Advanced Drug Delivery Reviews*. 2016;101:34-41.
96. Grim SA, Rapp RP, Martin CA, Evans ME. Trimethoprim-sulfamethoxazole as a viable treatment option for infections caused by methicillin-resistant *Staphylococcus aureus*. *Pharmacotherapy*. 2005;25(2):253-64.

Annexe 1 - Matériel supplémentaire du chapitre 2

Table A1-I. DfrB Gene cassettes obtained from Genbank as templates for *in silico* screening.

Gene name	GenBank accession #	Organism	Ref
<i>dfrB1</i>	U36276	<i>Escherichia coli</i>	Unpub ^a
<i>dfrB2</i>	U12441	<i>Escherichia coli</i>	(1)
<i>dfrB3</i>	AY123252.1	<i>Escherichia coli</i>	Unpub ^a
<i>dfrB4</i>	AY968808	<i>Klebsiella pneumoniae</i>	Unpub ^a
	KP314737.1	<i>Escherichia coli</i>	(2)
<i>dfrB5</i>	AY943084.1	<i>Pseudomonas aeruginosa</i>	(3)
<i>dfrB6</i>	DQ274503.1	<i>Salmonella enterica subsp. enterica serovar Infantis</i>	(4)
<i>dfrB7</i>	GU295656.1	<i>Aeromonas hydrophila</i>	(5)

^a Unpublished

ANNEXE 1 - REFERENCES

1. Partridge SR, Brown HJ, Hall RM. 2002. Characterization and movement of the class 1 integron known as Tn2521 and Tn1405. *Antimicrobial Agents and Chemotherapy* 46:1288–1294.
2. Delport TC, Harcourt RG, Beaumont LJ, Webster KN, Power ML. 2015. Molecular detection of antibiotic-resistance determinants in *Escherichia coli* isolated from the endangered Australian sea lion (*Neophoca cinerea*). *Journal of Wildlife Diseases* 51:555–563.
3. Lolans K, Queenan AM, Bush K, Sahud A, Quinn JP. 2005. First nosocomial outbreak of *Pseudomonas aeruginosa* producing an integron-borne metallo- β -lactamase (VIM-2) in the United States. *Antimicrobial agents and chemotherapy* 49:3538–3540.
4. Levings RS, Lightfoot D, Elbourne LDH, Djordjevic SP, Hall RM. 2006. New integron-associated gene cassette encoding a trimethoprim-resistant DfrB-type dihydrofolate reductase. *Antimicrobial Agents and Chemotherapy* 50:2863–2865.
5. Xu H, Broersma K, Miao V, Davies J. 2011. Class 1 and class 2 integrons in multidrug-resistant gram-negative bacteria isolated from the Salmon River, British Columbia. *Canadian Journal of Microbiology* 57:460–467.

Annexe 2 - Matériel supplémentaire du chapitre 3

Table A2-I. Crystallographic data for DfrB1-3.1 and DfrB1-3.3 (PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’).

	DfrB1-3.3 (PDB ID ^a)	DfrB1-3.1(PDB ID ^a)
Data collection statistics		
Space group	I4 ₁ 22	I4 ₁ 22
Number of Molecules per Asymmetric Unit	1	1
a=b (Å)	67.617	67.492
c (Å)	51.977	51.791
$\alpha = \beta = \gamma$ (°)	90	90
Wavelength (Å)	1.54	0.98
Resolution Range (Å) ^b	50-1.75 (1.78-1.75)	50-1.40 (1.43-1.40)
Completeness (%) ^b	94.3 (48.7)	100.0 (100.0)
Redundancy ^b	79.2 (1.5)	20.1 (18.5)
R _{merge} (%) ^b	20.4 (>100)	7.6 (>100)
R _{meas} (%) ^b	20.5 (>100)	ND ^c
R _{pim} (%) ^b	2.0 (>100)	ND ^c
CC _{1/2} ^b	(0.697)	ND ^c
CC* ^b	(0.907)	ND ^c
Refinement statistics		
Total number of reflections (reflections in R-free set)	5681 (319)	11,543 (624)
R _{factor} (%)	17.85	13.858
R _{free} (5% free test set) (%)	20.82	15.75
Number of atoms	515	562
Protein	429	459
Water	60	72
Buffer components	18	23
Inhibitor	8	8
RMSD		
Bond length (Å)	0.020	0.027
Bond angle (°)	2.061	2.493
Average atomic B-Factor (Å ²)	21.063	16.414
Protein (Å ²)	19.200	14.268
Water (Å ²)	33.637	26.536
Buffer components (Å ²)	70.420	36.917
Inhibitor (Å ²)	29.145	12.133
Ramachandran Plot	53	55
Residues in Favoured Positions	53	55
Residues in Allowed Positions	0	0
Residues in Disallowed Positions	0	0

^a ‘Authors will release the atomic coordinates and experimental data upon article publication.’

^b Items in parentheses refer to the highest resolution shell.

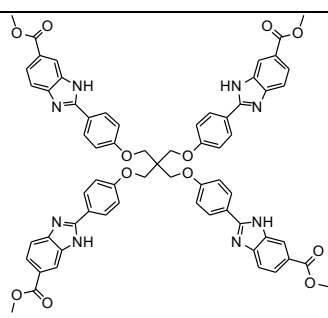
^c Not determined; the version of HKL-2000 used to process these data did not report these statistics.

Table A2-III. Type of interactions between inhibitor segments and amino acids of DfrB1 determined using the Protein Ligand Interaction Fingerprints (PLIF) module in MOE (CCG).

Inhibitors are listed in the order of their appearance in the Results section. H-bonds: D: side-chain donor, A: side-chain acceptor, d: backbone donor, a: backbone acceptor. I: ionic interaction. R: arene interaction.

	Moiety	[...]	K32	K33	S34	G35	A36	[...]	Y46	C47	T48	N49	L50	T51	[...]	S65	V66	Q67	I68	Y69	[...]
3.3	Terminal carboxylates		DI		dD	d	d		D		D										
	Benzimidazole																Rd	A	a	R	
	Phenoxy																	D			
	Central core																a	AD	ad		
3.1	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																ad		d	R	
	Phenoxy																	D	dR	R	
	Central core																a	A	ad		
3.2	Terminal carboxylates		DI		dD	d	d		D		D										
	Benzimidazole																aR		a	R	
	Phenoxy																d		R		
	Central core																		R		
3.4	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																	dR		a	R
	Phenoxy																		D	R	
	Central core																				
3.5	Terminal carboxylates		DI		dD	d	d		D		D							d			
	Benzimidazole															R		ad		a	R
	Phenoxy																d	DR	Rd		
	Central core																				
3.6	Terminal carboxylates		DI		D	d	d		D		D										
	Benzimidazole															R	adR		a	R	
	Phenoxy																	D	R		
	Central core																				
3.7	Terminal carboxylates		DI		D	d	d		D		D							d		Ra	R
	Benzimidazole																			R	
	Phenoxy																			R	
	Central core																				
3.10	Terminal carboxylates		DI			d	d		D		D								D	aR	R
	Benzimidazole																		D	R	
	Phenoxy																				
	Central core																				
3.11	Terminal carboxylates		DI	I	dD	d	d		D		D										
	Benzimidazole																		AR	d	R
	Phenoxy																	d	DR	R	
	Central core																R		R		
3.15	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																				
	Phenoxy																R	R			R
	Central core																				
3.2	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																R	R		ad	R
	Phenoxy																		d		
	Central core																a	AD	d		
3.21b	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																	dRa			R
	Phenoxy																A	a		Ra	
	Central core																	a	AD	D	
3.22	Terminal carboxylates		DI			d	d		D		D										
	Benzimidazole																	aRd	D	a	R
	Phenoxy																	a	A		
	Central core																				
3.38	Terminal carboxylates		DI		D	d	d		D		D										
	Benzimidazole																	aR		a	R
	Phenoxy																		DR	R	R
	Central core																		A		
3.12	Terminal carboxylates		DI		Dd	d	d		D		D										
	Benzimidazole																	ad	A	aR	R
	Phenoxy																		D	R	
	Central core																		R		R

Table A2-IV. The methylester of 3.38 does not inhibit DfrB1.

Compound	Structure	IC ₅₀ (μM) ^a
3.39c		>50

^a Compound **3.39c** was tested at the highest possible concentration, considering constraints due to precipitation.

Table A2-V. IC₅₀ (μM) of different inhibitors for DfrB1 and DfrB4.^a

	IC ₅₀ (μM)				
	3.1	3.3	3.12	3.21b	3.38
DfrB1	130 ± 11 ^b	113 ± 13	240 ± 87	83 ± 22	35 ± 9
DfrB4	68.9 ± 22.0	76 ± 18	210 ± 25	73 ± 13	37.3 ± 9.6

^a Values are given as the average ± standard deviation from the mean of at least duplicate of triplicate.

^b Values taken from Table 2 in (1).

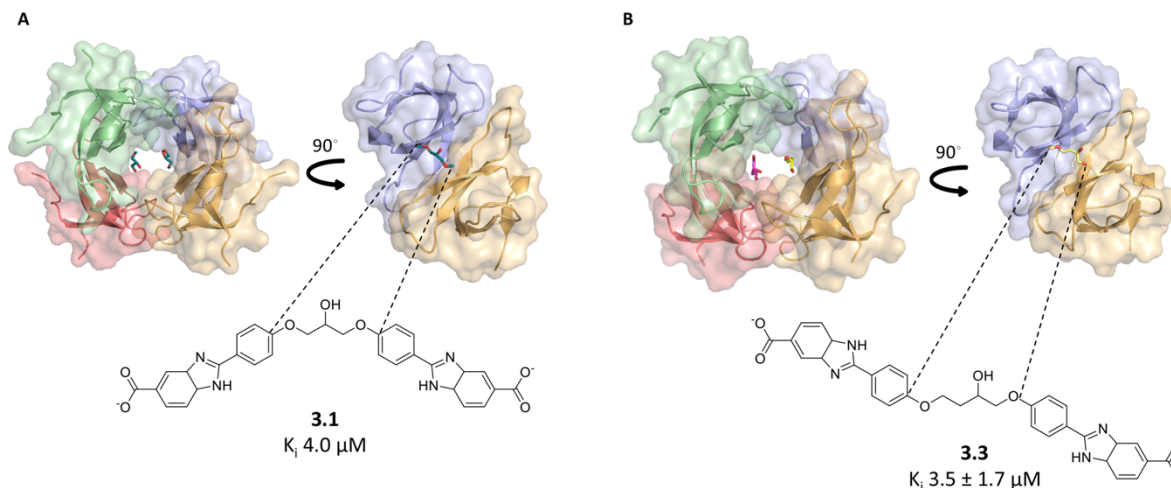


Figure A2-1. Crystal structures of 3.1 PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’ (left) and of 3.3 PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’ (right).

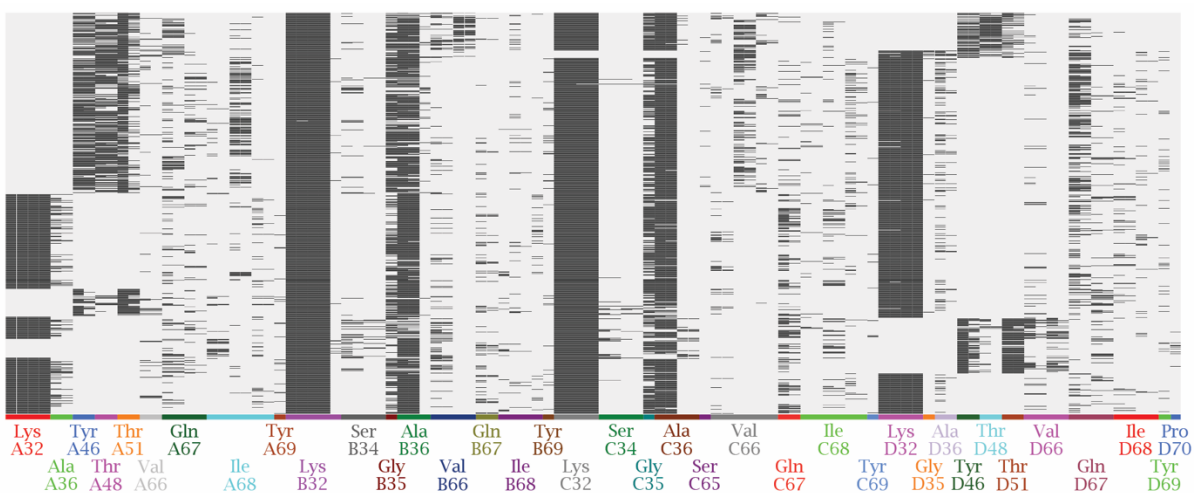


Figure A2-2. Matrix of the binding residues of DfrB1 with 3.3.

Each row represents one pose and each column, a specific interaction with a residue. For example, each of four Lys32 in the tetramer (Lys32A, Lys32B, Lys32C and Lys32D) can form 4 different interactions with the terminal carboxylate: (side-chain H-bond1, side-chain H-bond2, side-chain ionic interaction1 and side-chain ionic interaction2). The letters A to D indicate the protomer implicated in binding. Each interaction is scored as a black mark.

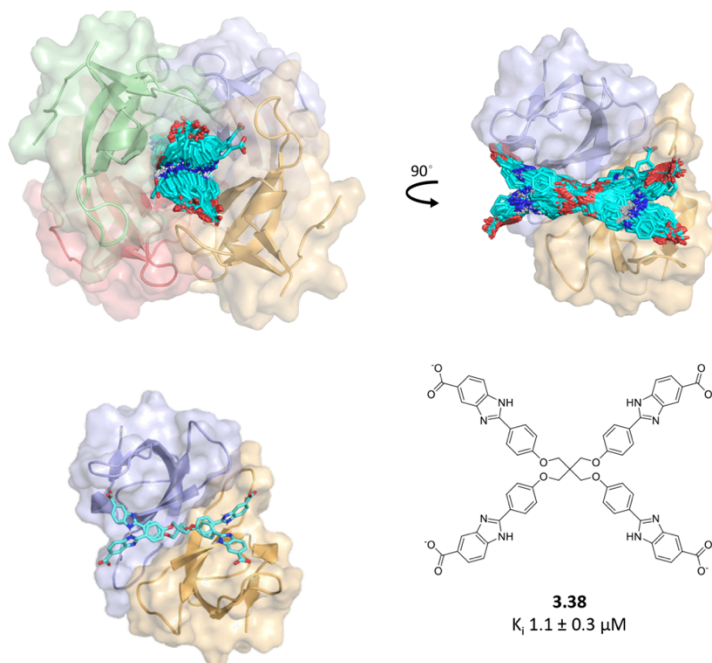


Figure A2-3. *In silico* reconstruction of 3.38 in the active site of DfrB1 based on the crystal structure PDB ID ‘Authors will release the atomic coordinates and experimental data upon article publication.’.

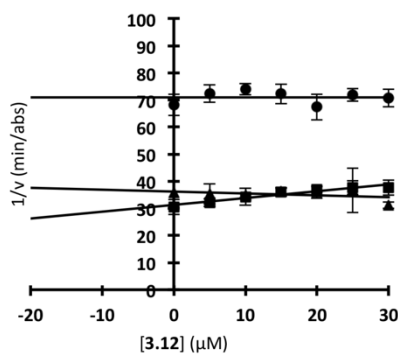


Figure A2-4. Dixon plot of 3.12 with DfrB1.

The reciprocal velocity of DfrB1 as a function of inhibitor concentration determines the type of inhibition. NADPH was held constant at 50 μM and DHF was 5 μM ($\sim 2 \times K_M^{\text{DHF}}$; ●) 25 μM ($13 \times K_M^{\text{DHF}}$; ▲) or 164 μM ($86 \times K_M^{\text{DHF}}$; ■). Values are given as the mean \pm standard deviation for triplicate results.

ANNEXE 2 - REFERENCE

1. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, et al. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of medicinal chemistry*. 2012;55:3182-92.

Annexe 3 - Matériel supplémentaire du chapitre 4

Table A3-I. Protein sequence identity and query coverage (in brackets) of DfrBs with the Expected value (in bold) ^a.

	DfrB1	DfrB2	DfrB3	DfrB4	DfrB5	DfrB6	DfrB7	DfrB9
DfrB1	100%	98% (79%) 6e-49	79% (98%) 9e-49	78% (98%) 7e-48	88% (100%) 7e-54	87% (100%) 3e-52	88% (100%) 2e-53	78% (98%) 9e-48
DfrB2		100%	86% (100%) 8e-50	74% (100%) 6e-46	81% (98%) 8e-49	83% (98%) 1e-48	79% (98%) 2e-48	85% (100%) 2e-50
DfrB3			100%	79% (100%) 2e-48	83% (98%) 5e-50	84% (98%) 3e-49	81% (98%) 6e-49	85% (100%) 1e-50
DfrB4				100%	78% (98%) 4e-49	78% (98%) 7e-47	77% (98%) 1e-48	74% (100%) 3e-46
DfrB5					100%	91% (100%) 2e-54	94% (100%) 2e-56	78% (98%) 1e-47
DfrB6						100%	91% (100%) 5e-54	81% (98%) 8e-48
DfrB7							100%	77% (93%) 3e-47
DfrB9								100%

^a The expected value (e) represents the probability of randomly matching two different sequences. The lower the e-value, the more significant the match.

Table A3-II. IC₅₀ of His₆-DfrB1 with TMP and inhibitors 4.1 to 4.5.

Inhibitor	IC ₅₀ (μM) ^a				
	4.1	4.2	4.3	4.4	4.5
His ₆ -DfrB1	650 ± 87	580 ± 140	520 ± 62	330 ± 19	550 ± 120

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

Table A3-III. IC₅₀ of DfrB1 (untagged) with TMP and inhibitors 4.1, 4.3, 4.6 and 4.7.

Inhibitor	TMP	IC ₅₀ (μM) ^a			
		4.1	4.3	4.6	4.7
DfrB1	(12 ± 3) × 10 ³	1300 ± 940	440 ± 47	170 ± 55	590 ± 62

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

Table A3-IV. Total activity, yield and purity of DfrB purifications. ^a

	Total activity (U/200 mL) ^{b,c}	Yield (mg/200 mL) ^c	Purity (%)
DfrB1	3.7	9.0	95
DfrB2	8.5	11.0	> 99
DfrB3	4.4	7.8	95
DfrB4	1.9	6.2	95
DfrB5	8.2	12.4	99
DfrB7 ^c	5.4	9.7	96

^a DfrB1 and DfrB4 were purified by treatment at 75°C followed by a size-exclusion purification step whereas all other DfrBs were purified by treatment at 65°C followed by a size-exclusion purification step, then by a second treatment at 75°C.

^b 1 unit (U) is defined as the concentration of substrate converted to product per minute. (μM/min).

^c Per 200 mL of expression medium.

Table A3-V. Theoretical and measured masses of purified DfrBs.

DfrBs	Theoretical mass (Da) ^a	Measured mass (Da) ^b
DfrB1	8,446.4	8,446.4
DfrB2 ^c	8,064.0	8,063.1
DfrB3	8,328.4	8,328.3
DfrB4	8,410.4	8,410.1
DfrB5	8,386.5	8,386.3
DfrB7	8,701.8	8,701.3

^a Calculated from amino acid sequence with ExPASy Bioinformatics Resource Portal.

^b Determined by mass spectroscopy.

^c Missing the initiator methionine.

Table A3-VI. $k_{\text{cat}}^{\text{NADPH}}$ for the dihydrofolate reductase activity of DfrBs.

DfrB homolog	$k_{\text{cat}}^{\text{NADPH}}$ (s ⁻¹) ^a
DfrB1	0.32 ± 0.01
DfrB2	0.35 ± 0.01
DfrB3	0.18 ± 0.01
DfrB4	0.22 ± 0.01
DfrB5	0.28 ± 0.01
DfrB7	0.26 ± 0.01

^a k_{cat} value calculated for the global reaction. Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

Table A3-VII. IC₅₀ of DfrBs with TMP and inhibitors 4.1, 4.3, 4.6 and 4.7.

Inhibitor	IC ₅₀ (μM) ^a				
	TMP	4.1	4.3	4.6	4.7
DfrB2	(13 ± 1) × 10 ³	790 ± 45	840 ± 230	170 ± 33	460 ± 14
DfrB3	(18 ± 4) × 10 ³	480 ± 63	660 ± 68	93 ± 14	250 ± 81
DfrB4	(16 ± 2) × 10 ³	540 ± 57	780 ± 15	190 ± 20	310 ± 35
DfrB5	(22 ± 5) × 10 ³	850 ± 39	730 ± 12	130 ± 2	500 ± 54
DfrB7	(12 ± 2) × 10 ³	980 ± 610	710 ± 190	130 ± 41	360 ± 130

^a Values are given as the average ± standard deviation from the mean of at least triplicates of two independent experiments.

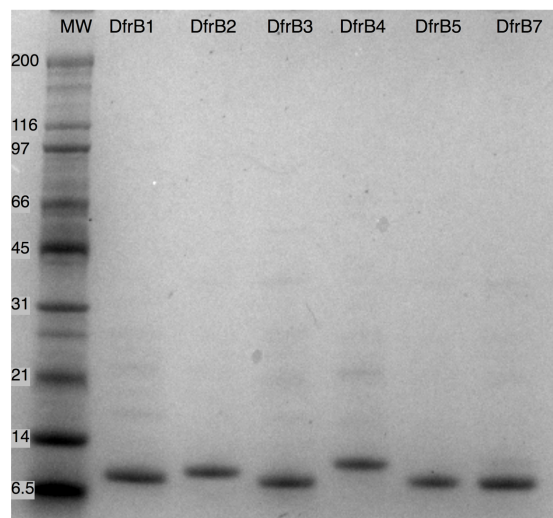


Figure A3-1. 12% SDS-PAGE gel of purified DfrBs.

MATERIALS AND METHODS

Phusion Green High-Fidelity DNA Polymerase was purchased from Thermo Scientific (Waltham, MA, USA). Unless otherwise mentioned, restriction enzymes were purchased from New England Biolabs Canada (NEB) and oligonucleotide primers were synthesized by Alpha DNA, S.E.N.C. (Montreal, QC). Ligations were performed with the Takara DNA Ligation Kit Ver.2.1. Kanamycin

was purchased from Teknova (Ontario, Canada) and TMP from Sigma Aldrich (Ontario, Canada). DNA sequencing was performed at the IRIC Genomic Platform at Université de Montréal.

Subcloning of DfrBs.

His₆-DfrB1 and His₆-DfrB4 were obtained as previously reported.(*Toulouse J.L. 2019* submitted) The genetic sequences of *dfrB2*, *dfrB3*, *dfrB5* and *dfrB7* were purchased from Bio Basic (Ontario, Canada) according to their respective GenBank accession numbers.(1) *DfrB* genes were obtained in pUC57 and subcloned into pET24 (Qiagen) for expression, as follows. A *dfrB* forward oligonucleotide primer (5'- AAAAACCGCGGCTAGCAAAGAGGAGAAATA-3') and two reverse oligonucleotide primers (DfrB2, DfrB5 and DfrB7: 5'- ATCGGATCCCGTCAGCTAATTAAGCTTTTA-3'; DfrB3: 5'- GACGGGCCCCTCAGCTAATTAAGCTTTTA-3') including the *HindIII* restriction site (underlined) served to PCR amplify each gene with Phusion Green High-Fidelity DNA Polymerase. The PCR products were digested with *NdeI* and *HindIII* and gel extracted using the Monarch DNA gel extraction kit from NEB. The inserts were ligated into *NdeI* and *HindIII*-digested pET24 for 30 minutes at 16°C. The DNA ligation products were transformed into competent *E. coli* DH5α by heat shock. The expected sequences of the final constructs were confirmed by DNA sequencing and transformed into *E. coli* BL21(DE3) (Qiagen) for expression.

To create the *dfrB4-pET24* construct, *His₆-DfrB4-pET24*(1) was used as a template for PCR amplification with a forward oligonucleotide primer (5'- AAAAAAACACATATGAATGAAGGAAAAAATGAGGTCAG-3') including a *NdeI* restriction site (underlined) and a reverse oligonucleotide primer (5'- AGCTAATTAAGCTTTTATCAGGCCACCC-3'; Sigma-Aldrich (Oakville, ON)) with a *HindIII* restriction site (underlined). The PCR product was purified and subcloned into pET24 and transformed as described above. The DNA sequence of the *dfrB4-pET24* construct was confirmed by sequencing.

To generate the *dfrB1-pET24* construct, a PCR method was used to remove the hexahistidine tag and the non-native13-residue C-terminal tail (ELGTPGRPAAKLN) from the WT R67 DHFR pQE32 construct.(2) A *NdeI* restriction site (underlined) was added using a forward oligonucleotide primer (5'- AAAAACCGCGGCTAGCAAAGAGGAGAAATAGCATATGAGAGGATCTCACCATCACC

ATC-3'; Sigma-Aldrich (Oakville, ON) and a *HindIII* restriction site (underlined) with a reverse oligonucleotide primer (5'- GGGAAGCTTTTAGTTGATGCGTTCAAGCGCC-3'; Sigma-Aldrich (Oakville, ON)) for compatibility with the *NdeI* and *HindIII*-digested pET24 plasmid. The PCR product was digested with *NdeI* and *HindIII* (Thermo Scientific, Waltham, MA) and purified using the PureLink Quick Gel Extraction and PCR Purification Combo Kit (Invitrogen, Carlsbad, CA). The digested fragment was subcloned into pET24 and transformed as described above. The sequence of the final construct was confirmed by DNA sequencing.

DfrB purification

Recombinant DfrBs were overexpressed in *E. coli* BL21 (DE3) using ZYP-5052 autoinduction medium.(3) Bacteria were propagated in 200 mL for two hours at 37°C then at 22°C for 16 hours to allow overexpression (DfrB1: 3 × 200 mL; DfrB2, DfrB3, DfrB4 and DfrB5: 2 × 200 mL; DfrB7: 1 × 200 mL). The cells were harvested and lysed as previously described.(2) Lysates were incubated at 65°C for 10 minutes (or at 75°C for DfrB1 and DfrB4) then chilled on ice for 10 minutes and centrifuged (10 minutes at 16,000 × g). The supernatants were concentrated to 1 mL using an Amicon concentrator (MWCO 3,000, Millipore) and injected on a Superose 12 column (1.6 cm × 55 cm) pre-equilibrated with 0.05 mM potassium phosphate buffer, pH 7.0. DfrB-containing fractions were identified according to their activity using the standard spectrophotometric Dfr activity assay. After this step, DfrB2, DfrB3, DfrB5 or DfrB7 were heated a second time at 75°C for 10 min and then chilled on ice for 10 min, followed by a centrifugation as above. Purity was determined by resolution on 12% SDS-PAGE. Protein concentration was determined using the Bio-Rad protein assay (Bio-Rad, Hercules, CA). Mass of the DfrBs was verified by mass spectrometry (Regional Mass Spectrometry Centre at Université de Montréal).

DfrB kinetics

The productive affinity (K_m), turnover number (k_{cat}) and the half maximal inhibitory concentration (IC_{50}) were determined as previously described.(4) The inhibition constant (K_i) was calculated by applying the Cheng-Prusoff equation(5) which relates IC_{50} to K_i .

Inhibitor synthesis

General procedure A. The selected adenosine derivative or piperidin derivative solubilized in methanol (MeOH) was added dropwise to a solution of $NaOCH_3/MeOH$ under nitrogen and stirred

at room temperature for 10 min. The solution was transferred to the previously synthesized adenosine derivative or piperidine derivative, and then, heated at 60 °C overnight. The reaction mixture was concentrated by rotary evaporation and extracted by CH₂Cl₂. Then, the concentrated solution was purified by flash chromatography.

General procedure B. The selected adenosine derivative was then added to a mixture of dihydropterin derivative, *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) and *N,N*-Diisopropylethylamine (DIPEA) in anhydrous dimethylformamide (DMF). After 2h, the solvent was evaporated under high vacuum and the reaction residue was purified by HPLC.

2-amino-N-(2-(4-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)piperidin-1-yl)ethyl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (1): Synthesis of **1** was done as described previously(6).

2-amino-N-(2-((2-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)sulfonyl)ethyl)amino)-2-oxoethyl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (2): Synthesis of **2** was done as described previously(6).

2-amino-N-(4-(2-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)ethyl)piperazin-1-yl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (3): To a solution of tert-butyl (4-(2-hydroxyethyl)piperazin-1-yl)carbamate in 20 mL dichloromethane (CH₂Cl₂) was added triphenylphosphine (PPh₃) and tetrabromomethane (CBr₄). After stirring at 0 °C for 10 min, the reaction mixture was concentrated by rotary evaporation. The reaction mixture contained the tert-butyl (4-(2-bromoethyl)piperazin-1-yl)carbamate and the compound was pure enough to proceed to the next step. 2',3'-*O*-isopropylideneadenosine-5'-thioacetate solubilized in MeOH and tert-butyl (4-(2-bromoethyl)piperazin-1-yl)carbamate were reacted according to General procedure A to result to tert-butyl (4-(2-(((3aS,4S,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)thio)ethyl)piperazin-1-yl)carbamate. The compound was pure enough to proceed to the next step. MS (ESI-MS) calculated for C₂₄H₃₈N₈O₅S (MH⁺): 551.27; found: 551.30. Trifluoroacetic acid (TFA) was added dropwise at -20 °C to a

solution of tert-butyl (4-(2-(((3aS,4S,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)thio)ethyl)piperazin-1-yl)carbamate in CH₂Cl₂. Then, the mixture was stirred at room temperature overnight. After the reaction was finished, the solvent was removed under vacuum, and the residue was purified by silica gel chromatography, to afford (2R,3R,4S,5S)-2-(6-amino-9H-purin-9-yl)-5-(((2-(4-aminopiperazin-1-yl)ethyl)thio)methyl)tetrahydrofuran-3,4-diol. The compound was pure enough to proceed to the next step. The compound **(2R,3R,4S,5S)-2-(6-amino-9H-purin-9-yl)-5-(((2-(4-aminopiperazin-1-yl)ethyl)thio)methyl)tetrahydrofuran-3,4-diol** and 6-carboxy-7,7-dimethyl-7,8-dihydropterin were reacted according to General procedure B to provide the final product **2-amino-N-(4-(2-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)ethyl)piperazin-1-yl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (5)**. MS (ESI-MS) calculated for C₂₅H₃₅N₁₃O₅S (MH⁺): 630.26; found: 630.30.

2-amino-N-(2-(4-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)piperidin-1-yl)ethyl)-4-oxo-3,4-dihydropteridine-6-carboxamide (4): (2R,3R,4S,5S)-2-(6-amino-9H-purin-9-yl)-5-(((1-(2-aminoethyl)piperidin-4-yl)thio)methyl)tetrahydrofuran-3,4-diol was synthesis as described previously(7). 6-pteridinecarboxylic acid and (2R,3R,4S,5S)-2-(6-amino-9H-purin-9-yl)-5-(((1-(2-aminoethyl)piperidin-4-yl)thio)methyl)tetrahydrofuran-3,4-diol were reacted according to General procedure B to result to provide **2-amino-N-(2-(4-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)piperidin-1-yl)ethyl)-4-oxo-3,4-dihydropteridine-6-carboxamide (4)** MS (ESI-MS) calculated for C₂₄H₃₀N₁₂O₅S (MH⁺): 599.22; found: 599.30.

2-amino-N-(1-(2-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)sulfonyl)ethyl)piperidin-4-yl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (5): S-(2-(4-((tert-butoxycarbonyl)amino)piperidin-1-yl)ethyl) ethanethioate solubilized and 2',3'-O-isopropylidene-5'-O-tosyladenosine were reacted according to General procedure A to result to tert-butyl (1-(2-(((3aS,4S,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)thio)ethyl)piperidin-4-yl)carbamate. This compound was oxidized by oxone in MeOH and the BOC group was removed by TFA/CH₂Cl₂ to yield to (2R,3R,4S,5S)-2-(6-amino-9H-purin-

9-yl)-5-(((2-(4-aminopiperidin-1-yl)ethyl)sulfonyl)methyl)tetrahydrofuran-3,4-diol.
(2R,3R,4S,5S)-2-(6-amino-9H-purin-9-yl)-5-(((2-(4-aminopiperidin-1-yl)ethyl)sulfonyl)methyl)tetrahydrofuran-3,4-diol and 6-carboxy-7,7-dimethyl-7,8-dihydropterin were reacted according to General procedure A to result to provide **2-amino-N-(1-(2-(((2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)sulfonyl)ethyl)piperidin-4-yl)-7,7-dimethyl-4-oxo-3,4,7,8-tetrahydropteridine-6-carboxamide (5)** MS (ESI-MS) calculated for C₂₆H₃₆N₁₂O₇S (MH⁺): 661.26; found: 661.30.

2,2'-(((2-hydroxybutane-1,4-diyl)bis(oxy))bis(3-bromo-4,1-phenylene))bis(1H-benzo[d]imidazole-6-carboxylic acid) (**6**) and 2,2'-(((1,2-phenylenebis(methylene))bis(oxy))bis(4,1-phenylene))bis(1H-benzo[d]imidazole-5-carboxylic acid) (**7**) were synthesized as previously described (*Toulouse J.L. 2019* submitted).

ANNEXE 3- REFERENCES

1. Toulouse JL, Edens TJ, Alejaldre L, Manges AR, Pelletier JN. Integron-associated DfrB4, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*. 2017;61:1-5.
2. Schmitzer AR, Lépine F, Pelletier JN. Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: Creating alternative functional configurations. *Protein Engineering, Design and Selection*. 2004;17:809-19.
3. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification*. 2005;41(1):207-34.
4. Toulouse JL, Abraham SMJ, Kadnikova N, Bastien D, Gauchot V, Schmitzer AR, et al. Investigation of classical organic and ionic liquid cosolvents for early-stage screening in fragment-based inhibitor design with unrelated bacterial and human dihydrofolate reductases. *Assay and Drug Development Technologies*. 2017;15:141-53.
5. Cheng Y, Prusoff WH. Relationship between the inhibition constant and the concentration of inhibitor which causes 50 percent inhibition of an enzymatic reaction. *Biochemical Pharmacology*. 1973;22:3099-108.
6. Shi G, Shaw G, Li Y, Wu Y, Yan H, Ji X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New lead exhibits a distinct binding mode. *Bioorganic and Medicinal Chemistry*. 2012;20:4303-9.
7. Shi G, Shaw G, Liang YH, Subburaman P, Li Y, Wu Y, et al. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New design with improved properties. *Bioorganic and Medicinal Chemistry*. 2012;20:47-57.

Annexe 4 - Procédure expérimentale pour la solubilisation et le criblage de la librairie des extraits d'invertébrés.

Solubilisation des extraits d'invertébrés marins

La librairie des extraits d'invertébrés (fournie par Prof. Raymond Andersen, Université de la Colombie-Britannique) est distribuée dans des plaques de 96 puits. Les plaques sont entreposées à -20°C et ce, entre chaque expérience. Avant la solubilisation des extraits, les plaques sont dégelées à la température ambiante pendant une vingtaine de minutes. La solubilisation des extraits est effectuée par la station de pipetage automatisé (Beckman NX^P) où 60 µL de DMSO est ajouté à chaque puits et mélangé par aspiration/expulsion répétée. Une fois les extraits solubilisés, la plaque est centrifugée à 2056 g, 22°C pendant 15 min, afin de séparer les particules insolubles des particules solubles.

Criblage des extraits d'invertébrés marins

Le NADPH ($\epsilon_{340\text{nm}} = 6\,200\text{ M}^{-1}\text{ cm}^{-1}$) et le DHF ($\epsilon_{282\text{nm}} = 28\,400\text{ M}^{-1}\text{ cm}^{-1}$) sont quantifiés par absorbance comme expliqué précédemment (1). Le DHF est synthétisé à partir de l'acide folique selon la méthode décrite (2). Le β -NADPH provient d'Alexis Biochemicals (San Diego, CA). La DfrB1 est purifiée selon la méthode décrite (1). La station de travail automatisée Laboratory Automation Workstation Biomek NX^P (Beckman Coulter, Montréal, Qc) couplée à une lectrice à plaques DTX 880 Multimode Detector (Beckman Coulter, Montréal, Qc), dépose dans chaque puits : 50 µM de NADPH, 50 µM de DHF, 90 nM de DfrB1 et 10 µL d'extrait solubilisé. Chaque essai est effectué dans le tampon 50 mM de phosphate de potassium à pH 7,0 et le volume final du milieu réactionnel est de 100 µL. L'activité de l'enzyme DfrB1 en présence de chaque extrait est mesurée durant dix minutes par la diminution de l'absorbance correspondant à la consommation du NADPH et DHF ($\Delta\epsilon_{340\text{nm}} = 12\,800\text{ M}^{-1}\text{ cm}^{-1}$) par l'enzyme (1). Pour chaque extrait, l'essai cinétique est réalisé en parallèle avec le contrôle négatif (tous les composants sauf la DfrB1). L'activité enzymatique est calculée en considérant la vitesse initiale de la réaction correspondant à la consommation de $\leq 10\%$ des substrats. Le contrôle positif se réalise en absence des extraits. Puisque le criblage d'une plaque d'extraits prend près

de 3 heures, l'activité du contrôle positif est mesurée en quadruplicata au début et à la fin du criblage. La moyenne de ces huit activités constitue le 100% d'activité. Le pourcentage d'activité enzymatique relative (%) de l'enzyme DfrB1 en présence de chacun des extraits est calculé par la moyenne de l'activité enzymatique de la DfrB1 en présence du composé, divisé par la moyenne du contrôle positif, multipliée par 100. Le pH est suivi tout au long du criblage avec du papier indicateur de pH.

ANNEXE 4 - REFERENCES

1. Bastien D, Ebert MCCJC, Forge D, Toulouse J, Kadnikova N, Perron F, et al. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *Journal of Medicinal Chemistry*. 2012;55:3182-92.
2. Blakley RL. Crystalline dihydropteroylglutamic acid. *Nature*. 1960;188:231-2.