

Beskrywing, modellering en dok-studies van *Plasmodium falciparum* kinase PfCDPK4

Outeurs:

Thomas Makungo,^a
Tsepo Tsekoa,^b
Anjo Theron,^b
Dalu Mancama,^b
Teunis van Ree^a

Affiliaties:

^aDepartement Chemie,
Universiteit van Venda,
Privaatsak X5050,
Thohoyandou, 0950,
Suid-Afrika

^bDepartement Biowetenskap,
Wetenskaplike en
Nywerheidsnavorsingsraad,
Posbus 395, Pretoria, 0001
Suid-Afrika

Korresponderende outeur:

Teunis van Ree
E-pos:
Teuns.VanRee@univen.ac.za;
teunsvanree@gmail.com

Datums:

Ontvang: 18/06/18
Aanvaar: 19/02/19
Gepubliseer: 11/04/19

Hoe om hierdie artikel aan te haal:

Thomas Makungo, Tsepo Tsekoa, Anjo Theron, Dalu Mancama, Teunis van Ree, Beskrywing, modellering en dok-studies van *Plasmodium falciparum* kinase PfCDPK4, *Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie* 38(1) (2019). <https://doi.org/10.36303/SATNT.2019.38.1.677>

An English copy of this paper is available online at <http://www.satnt.ac.za/index.php/satnt/article/view/677>

Kopiereg:

© 2019. Authors.
Licensee: Die Suid-Afrikaanse Akademie vir Wetenskap en Kuns.
Hierdie werk is onder die Creative Commons Attribution License gelisensieer.

Met die toenemende voorkoms van weerstandige *Plasmodium* stamme het die beheer van malaria-voorkoms en -mortaliteit weer op die voorgrond getree. Nuwe teikens en antimalariamiddels wat effektief is teen weerstandige malaria-parasiete word dus dringend benodig. Kalsium-afhanklike proteïenkinases (calcium dependent protein kinases – CDPKs) is betrokke by die beheer van 'n aantal biologiese prosesse in die malaria-parasiet, *Plasmodium falciparum*, met CDPK4 die belangrikste ensiem in hierdie klas. In hierdie studie is die struktuur van PfCDPK4 gebruik as templaar vir die soeke na nuwe malariamiddels. Die PfCDPK4 modelstruktuur is deur middel van homologiëmodellering gegenereer en die stereochemiese kwaliteit gevalideer. Die molekuleêre modelleringbenadering deur middel van *in silico* sifting teen die teiken-molekuul PfCDPK4 het 'n beskeie biblioteek van 20 000 chemiese verbindings ingesluit, asook 'n aantal aktiewe natuurprodukte en klinies-goedgekeurde kinase-inhibeerders. *In silico* sifting van die Biofocus biblioteek teen PfCDPK4 het 26 verbindings opgelewer; *in vitro* sifting het bevestig dat drie van hierdie verbindings matig aktief is teen *Plasmodium falciparum* NF54, met persentasie inhibisie tussen 42% en 47%.

Characterization, modelling and docking studies of *Plasmodium falciparum* kinase PfCDPK4:

The increasing incidence of *Plasmodium* strains that are resistant to current frontline anti-malarial drugs has become one of the greatest challenges of controlling malaria incidence and mortality. There is, therefore, an urgent need to develop novel targets and anti-malarial drugs that are effective against drug-resistant malarial parasites. At the same time the need to develop anti-malarial drugs that furthermore prevent disease transmission, has become an increasingly important consideration. Calcium dependent protein kinases (CDPKs) regulate a variety of biological processes in the malaria parasite *Plasmodium falciparum*, CDPK4 being of prime importance in *Plasmodium* biology. In this study the structure of PfCDPK4 was used as a template in the discovery of malaria drug leads. The model structure of PfCDPK4 was generated by homology modelling, and model validation confirmed that the model of PfCDPK4 is of stereochemical quality. The molecular modelling approach of *in silico* screening against the target molecule PfCDPK4 utilized a large library of chemical compounds, some natural chemical compounds, and clinically approved kinase inhibitors. *In silico* screening of the Biofocus library against PfCDPK4 resulted in twenty-six compounds being identified; *in vitro* screening confirmed that three of these compounds exhibit moderate anti-malarial activity against the NF54 strain of *Plasmodium falciparum*, with the percentage inhibition ranging between 42% and 47%.

Inleiding

Die toenemende voorkoms van *Plasmodium* variante wat weerstandig is teen die huidige eerste-linie antimalariamiddels sorg dat malaria steeds 'n globale openbare gesondheidsuitdaging bly. Daar is steeds behoefte aan nuwe antimalaria-middels wat aktief sal wees teen weerstandige parasiete. Die siekte word oorgedra deur geïnfecteerde vroulike *Anopheles* muskiete (Biamonte et al., 2013; Dorin-Semlat et al., 2013; Lounnas et al., 2013). Vyf malaria spesies veroorsaak die siekte in mense, maar *Plasmodium falciparum* veroorsaak die dodelikste infeksies (Anastassiadis et al., 2011; Klein, 2013; Vidadala et al., 2014; WHO, 2017), veral in Wes-, Oos- en Sentraal-Afrika (Murray et al., 2012). Meer as 216 miljoen mense word jaarliks aan die risiko van malaria-infeksie blootgestel, daarvan 90% in Afrika suid van die Sahara, met veral kinders en swanger vroue wat kwesbaar is (WHO, 2016).

Die bekendste antimalaria-middels chlorokien (CQ), sulfadoksien-pirimetamien (SP) en artemisinien is hoogs doeltreffend teen die eritrosiet-stadium van *P. falciparum* infeksies (Klein, 2013). *P. falciparum* weerstandigheid teen CQ, meflokin en SP het egter teen die 1960s

in die Kambodja-Thailand grensgebied begin ontwikkel en vandaar geleidelik vanaf die Mekong delta na Afrika versprei. Studies van 2001 tot 2009 het weerstandigheid in Kambodja en Thailand bevestig (WHO, 2017), en het aanleiding gegee tot die globale plan vir die beperking van artemisinien-weerstand (Global Plan for Artemisinin Resistance Containment – GPARC) (WHO, 2011). Die huidige benadering van artemisinien kombinasietherapie (ACT) slaag ook steeds minder daarin om malaria-oordrag van mens na muskiet te verhinder (Ojo et al., 2012).

Proteïenkinases kataliseer die oordrag van fosfaat vanaf ATP na spesifieke aminosuur-residue op proteïen-teikens of substrate, en beïnvloed die aktiwiteit, stabiliteit en interaksie van die teikenproteïen(e) met ligande (Lim et al., 2012). Hierdie kritieke rol wat proteïenkinases in sellulêre kommunikasie speel, en die bewese teenwoordigheid van ATP-aktiewe posisies wat deur chemiese verbindings bereik en dus geïnhibeer kan word, maak die proteïenkinases 'n belowende klas as terapeutiese teikens (Anastassiadis et al., 2011). Die CDPKs word direk deur kalsium gereguleer en beïnvloed 'n verskeidenheid biologiese prosesse in die malariaparasiet (Kumar et al., 2017). CDPKs is teenwoordig in *P. falciparum*, plante, en sommige eensellige organismes (Ranjan et al., 2009). Hierdie ensieme en verwante homoloë kom nie in die mens voor nie en is dus ideale teikens vir geneesmiddels. In *Plasmodium* funksioneer verskillende CDPKs tydens verskillende stadia in die parasiet se lewensiklus (Chapman et al., 2013).

CDPK4 word in die gametosietstadium van die malaria parasiet uitgedruk (Bansal et al., 2013). Dit is 'n belangrike seinmolekuul vir die transisie na mikrogameet- en makrogameetvorms wat saamsmelt om 'n sigoot te vorm (Billker et al., 2004; Ranjan et al., 2009). Die *Plasmodium* CDPK4 vorm dus 'n belangrike skakel in die malaria-lewensiklus en transmissie (Ojo et al., 2012, 2014; Tewari et al., 2010). Genoom-inligting dui daarop dat *Plasmodium berghei* CDPK4 (*PbCDPK4*) homoloog is met *Plasmodium falciparum* CDPK4 (*PfCDPK4*) (91% gemeenskaplike aminosure en 97% ooreenkoms) en *Toxoplasma gondii* CDPK1 (*TgCDPK1*) (met 73% gemeenskaplike aminosure en 88% ooreenkoms) (Billker et al., 2004). *PbCDPK4* is essensieel vir die parasiet se seksuele ontwikkelingsfase in die muskiet (Dorin-Semlat et al., 2013; Lounnas et al., 2013; WHO, 2017), en waarskynlik speel *PfCDPK4* dieselfde rol. Die ontdekking van verbindings wat CDPK4 inhibeer sal dus 'n belangrike bydrae maak tot die bekamping van malaria. 'n Geneties-chemiese benadering het reeds *PfCDPK4* as 'n moontlike teiken vir nuwe geneesmiddels gevalideer (Ojo et al., 2014).

PfCDPK4 word in die *Plasmodium* gameet/gametosietstadium uitgedruk en onderdrukking van die CDPK4 geen veroorsaak ernstige defekte in die parasiet se seksuele reproduksie (Ranjan et al., 2009). Die ensiem en sy homoloë is ook afwesig in die mens, sodat *PfCDPK4* 'n belowende teiken is vir gametosietstadium parasiete en dus transmissie

van die siekte. Die struktuur van *PfCDPK4*, veral die kinase-domein, kan dus benut word om nuwe middels teen malaria te vind deur chemiese verbindings of inhibiteurs van die teiken-ensiem te ontwerp. Om hierdie idee te evalueer, het ons in hierdie studie 'n beskeie biblioteek van chemiese verbindings *in silico* gesif vir moontlike aktiwiteit teen die teikenensiem, *PfCDPK4*. 'n Homologie modelstruktuur van *PfCDPK4* is gebruik om moontlik aktiewe anti-malariaverbindings te vind, en om die *PfCDPK4* model as 'n teiken vir ontwikkeling van nuwe geneesmiddels te valideer. Inligting verkry deur *in silico* virtuele sifting is vervolgens gebruik om 'n versameling verbindings te identifiseer vir *in vitro* sifting teen 'n *P. falciparum* kultuur en het 'n aantal belowende verbindings opgelewer vir moontlike verdere geneesmiddelontwikkeling.

Materiaal en metodes

Molekuulmodellering van *PfCDPK4*

Struktuurvoorspelling van *PfCDPK4* deur molekuulmodellering

Die aminosuurvolgorde van die kinase- of ATP-bindingsdomein van *PfCDPK4* is verkry van die *UniProt Knowledgebase* (UniProtKB) met Uniprot ID Q8IBS5 (UniProt Consortium, 2012). 'n Drie-dimensionele struktuur van *PfCDPK4* is voorheen vasgestel deur middel van homologie-modellering gebaseer op die kristalstruktuur van die *TgCDPK1:BKI-1* kompleks (Ojo et al., 2012). In die huidige studie is die 3D struktuur bepaal deur die aminosuurvolgorde van *PfCDPK4* self as 'n soekterm te gebruik om 'n templatstruktuur in die *Protein Data Bank* (PDB) te vind deur middel van 'n BLAST soektog (Altschul et al., 1997) met *Discovery Studio 3.5* (Accelrys, 2012). Die BLAST soekprotokol is gebruik om 'n soektog binne die PDB_nr95 sekvensdatabasis in *Discovery Studio* te doen, met BLOSUM62 as substitusie/vergelijking matriks. 'n Templaat baie na aan die *PfCDPK4* aminosuurvolgorde was weereens die kinase of ATP-bindingdomein *TgCDPK1* (PDB Id: 3ma6) van die eensellige parasiet *Toxoplasma gondii*. Die drie-dimensionele *TgCDPK1* struktuur is in die *RCSB Protein Data Bank* gevind (Deshpande et al., 2005).

Die *Align 123* toepassing van *Discovery Studio 3.5* is gebruik om die aminosuurvolgorde van *PfCDPK4* met die *TgCDPK1* volgorde te vergelyk. Die gebruik van 'n gewysigde CLUSTAL W (Thompson et al., 1994) algoritme, MODELER 9v8 (Šali et al., 1995), en die *Build Homology Model* protokol het 20 modelle opgelewer. Hierdie 20 modelle is met behulp van die *Verify Protein* (MODELER) protokol en die *Verify Protein (Profiles-3D)* protokol geëvalueer om die beste model te bepaal. Die DOPE (*Discrete Optimized Protein Energy*) telling is vir elke proteïenstruktuur bereken om die verkillende moontlike konformasies te vergelyk. 'n Laer DOPE telling beteken 'n beter kwaliteit model is gevind (Accelrys, 2012; Shen & Šali, 2006). Die model proteïenstruktuur met die laagste DOPE telling en hoogste *Verify* telling is gebruik vir verdere verbetering en validering.

PfCDPK4 modelstruktuurverfyning en validering vir stereochemiese kwaliteit

Die gekose modelstruktuur (*PfCDPK4.F06*) is verbeter deur 'n soektog vir lae-energie konformasies van lus-strukture met die *Loop Refinement* toepassing, gebaseer op *CHARMM* molekulêre meganika (Brooks et al., 1983; Karplus, 2003). Proteïenlusse is buigsame gebiede op proteïenoppervlaktes (Tang et al., 2014) en is dele van die polipeptied sonder sekondêre struktuur (Regad et al., 2010). Vir elke lus wat gekies is, is verskeie lae-energie konformasies onder 'n energie-drumpel van 20 kcal/mol gegeneer en die beste model is gevind deur middel van die *DOPE* en *Verify* tellings van die verskillende konformasies te vergelyk. Die verbeterde modelle is verder ondersoek deur middel van die *Verify Protein* (MODELER) en *Verify Protein* (Profiles-3D) toepassings. Die volgende aminosure is só aangepas: Glu132-Ile140, Gly196-Asn204, Leu55-Lys62, Met54-Lys62, Asn117-Lys125, Lys33-Glu43, Leu55-Lys62, Gly215-Leu221, His151-Lys157, Val119-Pro126, Met240-Ser248, Phe152-Lys157, Asp78-Gly82, Lys33-Asp41, Ile144-Gly147, Asp19-His24, Lys225-Ser228, Gly9-Phe11 en Asp69-Asn71. Die finale model is gevalideer deur gebruik van *PROCHECK* (Arnold et al., 2006, 2009; Laskowski et al., 1993; <http://swift.cmbi.ru.nl/servers/html/index.html>) verkry vanaf die *SWISS-MODEL Workspace* webtuiste (<http://swissmodel.expasy.org/workspace/>). Die *Ramachandran Plot* (Ramachandran et al., 1963) is gebruik om die stereochemiese kwaliteit van die finale *PfCDPK4* model te analiseer (Hovmoller et al., 2002).

Teiken en biblioteekvoorbereiding

PfCDPK4 model-optimering en bindingposisie

Die *PfCDPK4* modelstruktuur is voorberei deur middel van die *Prepare protein* protokol in *Discovery Studio 3.5* (Accelrys), wat oorblywende probleme in die proteïenstruktuur oplos met stappe soos protonering, standardisering van atoomname, en eliminerings van alternatiewe konformasies.

Die bindingposisie is afgelei van die reseptor (*PfCDPK4*) se holruimtes deur middel van 'n uitvee-algoritme (Kang et al., 2017; Taha et al., 2014; Venkatachalam et al., 2003).

Vorbereiding en optimering van die Biofocus en NSV biblioteke en bekende kinase-inhibitore

Drie versamelings van verbindings is in hierdie ondersoek gebruik, nl. die *Biofocus* biblioteek van 20 000 verbindings, 'n kinase-inhibitor versameling bestaande uit 'n klein subversameling van 13 klinies-bewese (FDA-goedgekeurde) verbindings (Zhao et al., 2013), en 'n versameling van 28 Natuurlike en Sintetiese Verbindings (NSV) met bekende of vermoede antimalaria-aktiwiteit. Verbindings van die *Biofocus* en NSV-versamelings is gefiltreer met behulp van Lipinski reëls (Lipinski et al., 1997) in *Discovery Studio*. Hierdie verbindings asook die kinase-inhibitore is geoptimeer deur middel van die *Prepare Ligands* protokol. Tautomere en isomere is gegeneer en duplikate verwyder.

Virtuele en *in vitro* sifting van klein molekule teen die moontlike teiken PfCDPK4

Virtuele sifting van die Biofocus biblioteek teen die PfCDPK4 3D modelstruktuur

Die 83 707 voorbereide ligande van die *Biofocus* biblioteek teen die voorbereide *PfCDPK4* struktuur is gesif deur middel van 'n *CHARMM*-gebaseerde kragveld in die *LibDock* toepassing van *Discovery Studio 3.1* (Yu et al., 2013).

Dok van bekende kinase-inhibitore en NSVs aan die PfCDPK4 3D modelstruktuur

Die 36 verbindings wat verkry is deur optimering van die 13 bewese kinase-inhibitore, sowel as die 28 NSVs, is teen die voorbereide *PfCDPK4* gesif deur middel van die *CHARMM*-gebaseerde *CDOCKER* toepassing in *Discovery Studio 3.5* (Wu et al., 2003). 'n Maksimum van tien lukraak konformere is vir elke ligand gegeneer en die molekulêre dinamika bereken deur volgehoue verwarming- en verkoelingstadias, tot 'n minimum bereik is. Die *poses* wat so verkry is, is geassesseer met behulp van die funksies *PLP1*, *PLP2*, *PMF*, *PMF04*, *Jain*, *LigScore1*, en *LigScore2*, gevolg deur samestelling van 'n 'konsensus' telling; verbindings met die hoogste tellings is uitgesoek vir opvolgwerk.

***In vitro* sifting van voorkeur verbindings teen Plasmodium falciparum NF54 stam in die gametosiet stadium**

Die antiplasmodiale toetse teen laat-stadium (IV-V) gametosiete is deur die Suid-Afrikaanse Wetenskaplike en Nywerheidsnavorsingsraad (WNNR, Pretoria, Suid Afrika) se Molekulêre en Biomediese Tegnologieë groep uitgevoer, soos beskryf in (Reader et al., 2015).

Resultate en bespreking

Molekuulmodellering van PfCDPK4

Struktuurvoorspelling deur homologimodellering

Die volledige *PfCDPK4* aminosuurvolgorde gevind in die *UniProt Knowledgebase* (UniProt Consortium, 2012) bestaan uit vyf domeine met 'n totale lengte van 528 aminosuurresidue. Die proteïenkinase deel bestaan uit posisies 71 tot 329, met 'n domein bestaande uit 259 aminosure.

'n *BLAST* soektog met hierdie volgorde in die *PDB* (Protein Data Bank, <http://www.rcsb.org/pdb/>) het 'n lys van 240 homoloë van die *PfCDPK4* kinasedomein, met bekende proteïenstrukture, opgelewer. In hierdie lys is vyf template (Tabel 1) gevind met voldoende volgorde-identiteit en ooreenkoms met die *PfCDPK4* volgorde om modelkwaliteit te verseker (François et al., 2000; Bordoli et al., 2008; Vyas et al., 2012). 'n Model met meer as 50% templaatt-eiken ooreenkoms dui aan dat die model akkuraat genoeg vir verdere ontwikkeling sal wees (Cavasotto en Phatak, 2009; Vyas et al., 2012). Kalmodulin-domein proteïenkinase 1 (*PDB* Id: 3ma6) van *Toxoplasma gondii* (*TgCDPK1*) is daarom as templaatt gekies weens die hoë volgorde-identiteit (79%) met *PfCDPK4*. Die templaatt 3ma6 het ook 'n goeie resolusie van 2.5 Å en die X-straal kristalstruktuur is vanaf die *Protein Data Bank* (<http://www.rcsb.org/pdb/>) afgelaai.

TABEL 1: Vyf templaattstrukture homologe aan die *PfCDPK4* kinasedomein, verkry deur middel van 'n BLAST soektoeg van die Proteien Databank (PDB).

PDB id	Volgorde-identiteit (%)	Volgorde-lengte	Domein-lengte	Bit Score	E-Waarde	Resolusie (Å)	Organisme
3ma6 (B)	79	266	259	427.56	1.02e-120	2.50	<i>T. gondii</i>
3nyv (A)	78	457	259	426.79	1.85e-120	1.88	<i>T. gondii</i>
2wei (A)	68	278	257	362.07	6.61e-101	1.65	<i>C. parvum Iowa II</i>
3lij (A)	63	465	257	323.17	3.03e-89	1.90	<i>C. parvum Iowa II</i>
3dxn (A)	61	258	250	315.08	9.35e-87	2.17	<i>T. gondii</i>

Vergelyking van die *PfCDPK4* en *TgCDPK1* aminosuurvolgordes het 78% volgorde-identiteit en 90% volgorde-ooreenkoms getoon (Figuur 1).

**FIGUUR 1:** Vergelyking van die *PfCDPK4* en *TgCDPK1* aminosuurvolgordes toon 78% volgorde-identiteit en 90% volgorde-ooreenkoms. Donkerblou kleur dui op gekonserveerde aminosuurresidue; Lig blou – sterk ooreenkoms; Flou blou – swak ooreenkoms; Wit – geen ooreenkoms.

Vervolgens is die 3D struktuurmodel van *PfCDPK4* deur middel van die MODELER toepassing in *Discovery Studio V 3.5* gegenereer. Twintig *PfCDPK4* modelle is gegenereer en die beste hiervan gekies op basis van die laagste DOPE (*Discrete Optimized Protein Energy*) telling en hoogste *Verify* telling (Tabel 2). Terwyl die DOPE telling van die *PfCDPK4.F06* model die laagste was (-30 087) en die *Verify* telling van die *PfCDPK4.F09* model die hoogste (114), is besluit om die modelle te verfyn, aangesien die *Verify* telling hoër as die verwagte (117.717) behoort te wees vir 'n goeie kwaliteit model.

***PfCDPK4* modelstruktuur verfyning en validering vir stereochemiese kwaliteit**

Die Profile-3D lynplot (Figuur 2, blou lyn) vir die *PfCDPK4* aminosuurresidue wys dat Leu130 'n negatiewe *Verify* telling het, wat onreëlmatige vou van die proteïen veroorsaak (Huang et al., 2010a). Die *Verify* tellings vir Leu131, Asn135 and Val245 was ook naby 0. Gevolglik is verdere lusverfyning op hierdie residue uitgevoer. Na hierdie optimering van die *PfCDPK4.F06* model het die DOPE en die *Verify* tellings heelwat verbeter (Tabel 3), terwyl die aminosure met lae

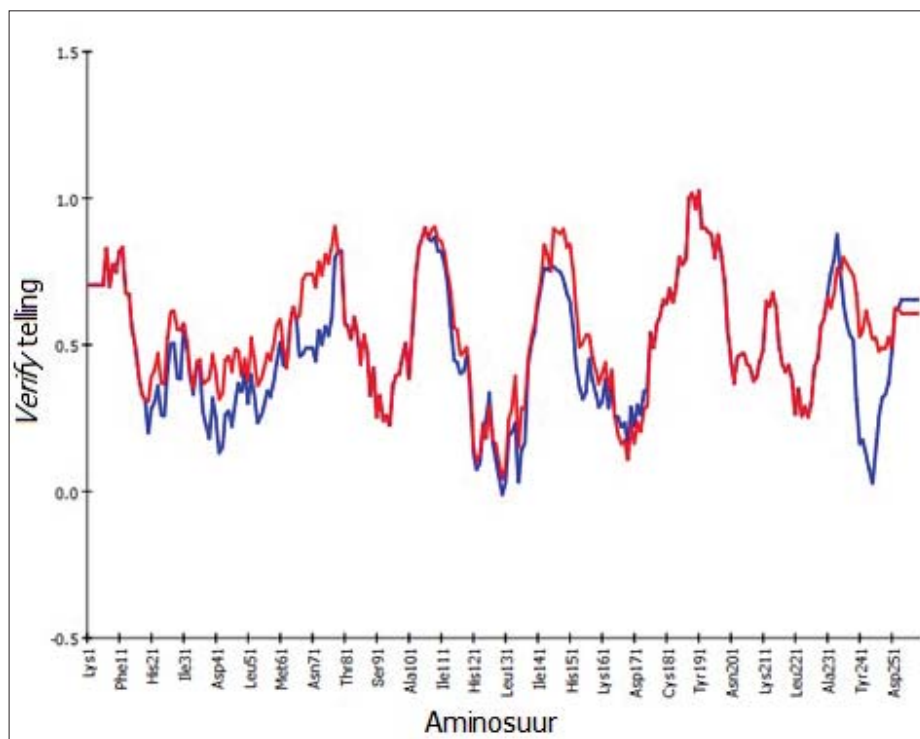
TABEL 2: Die tellings (DOPE en *Verify* tellings) van 3D struktuurmodelle van *PfCDPK4* gegenereer deur MODELER.

Naam	DOPE telling	<i>Verify</i> verwagte lae telling	<i>Verify</i> verwagte hoë telling	<i>Verify</i> telling
<i>PfCDPK4.F01</i>	-29 966	52.973	117.717	112
<i>PfCDPK4.F02</i>	-29 834	52.973	117.717	108
<i>PfCDPK4.F03</i>	-29 830	52.973	117.717	106
<i>PfCDPK4.F04</i>	-29 576	52.973	117.717	110
<i>PfCDPK4.F05</i>	-29 639	52.973	117.717	108
<i>PfCDPK4.F06</i>	-30 087	52.973	117.717	113
<i>PfCDPK4.F07</i>	-29 718	52.973	117.717	106
<i>PfCDPK4.F08</i>	-29 819	52.973	117.717	112
<i>PfCDPK4.F09</i>	-29 944	52.973	117.717	114
<i>PfCDPK4.F10</i>	-29 937	52.973	117.717	110
<i>PfCDPK4.F11</i>	-29 792	52.973	117.717	108
<i>PfCDPK4.F12</i>	-29 924	52.973	117.717	108
<i>PfCDPK4.F13</i>	-29 416	52.973	117.717	103
<i>PfCDPK4.F14</i>	-29 651	52.973	117.717	107
<i>PfCDPK4.F15</i>	-29 933	52.973	117.717	112
<i>PfCDPK4.F16</i>	-29 719	52.973	117.717	107
<i>PfCDPK4.F17</i>	-29 612	52.973	117.717	103
<i>PfCDPK4.F18</i>	-29 767	52.973	117.717	110
<i>PfCDPK4.F19</i>	-29 678	52.973	117.717	106
<i>PfCDPK4.F20</i>	-29 786	52.973	117.717	112

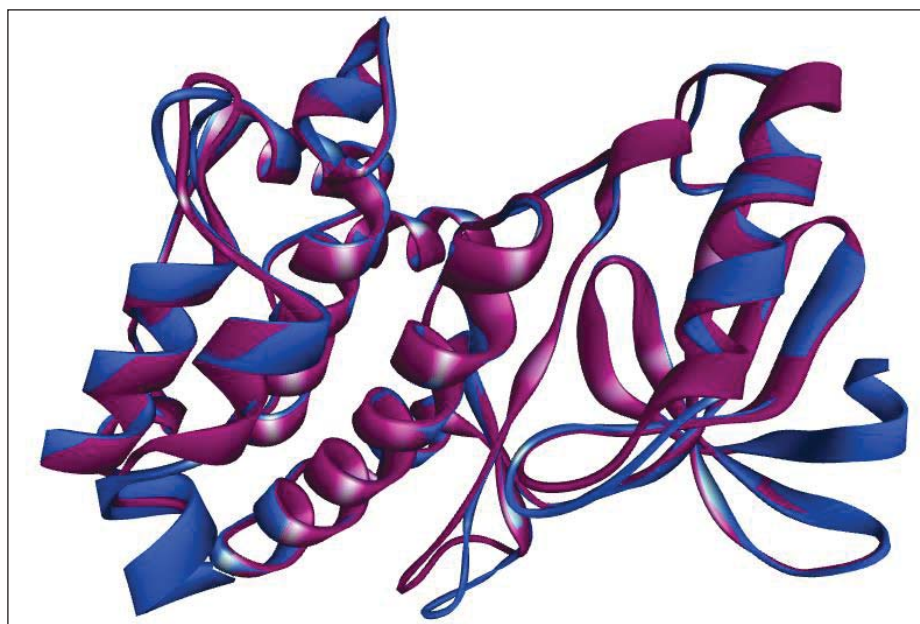
Verify tellings nou almal in die positiewe gebied voorkom (Figuur 2, rooi lyn), wat op die verbeterde kwaliteit van die finale model dui.

Passing van die finale *PfCDPK4* modelstruktuur op die

templaatstruktuur (*TgCDPK1*, PDB id: 3ma6, ketting B), toon die ooreenkomste tussen die twee strukture, met 'n gemiddelde afwyking (*root mean square deviation*, RMSD) van 0.92 Å (Figuur 3).



FIGUUR 2: Die aminosuur *Verify telling* vir die finale model (*PfCDPK4*, rooi) is beter as die van die oorspronklike model (*PfCDPK4.F06*, blou), soos deur die *Verify Protein (Profiles-3D)* lyndiagramme getoon.



FIGUUR 3: Superponering van die *PfCDPK4* homologiemodel (pers) en die *TgCDPK1* (blou) strukture toon die ooreenkomste.

TABEL 3: Vergelyking van die verbeterde met die oorspronklike *PfCDPK4* model. Die Ramachandran Z-telling bereken deur middel van die WHAT-IF toepassing (Hooft et al., 1997) meet die kwaliteit van die Ramachandran diagram (Figuur 4). Die tellings bevestig die kwaliteit van die finale model.

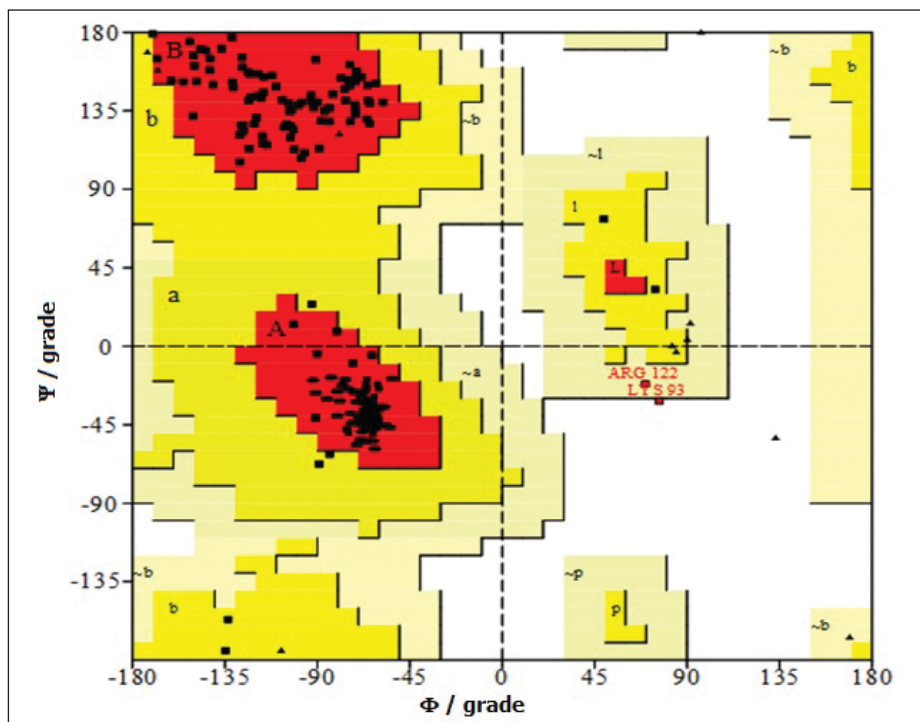
Naam	DOPE telling	Verify telling	Ramachandran Z-telling
Oorspronklike Model (<i>PfCDPK4.F06</i>)	-30 087	113	0.800
Finale Model (<i>PfCDPK4</i>)	-30 813	126	0.072

Vir finale validering van die model is PROCHECK gebruik om die stereochemiese kwaliteit te bepaal met behulp van Ramachandran-diagramme wat die ruggraat dihedrale hoeke van al die aminosure in die struktuur wys en analiseer. Die Ramachandran-diagram wat met behulp van PROCHECK (Arnold et al., 2006), beskikbaar in die SWISS-MODEL *Workspace* (<http://swissmodel.expasy.org/workspace/>), saamgestel is (Figuur 4), toon aan dat 96% van die ϕ en Ψ hoeke van die aminosure in die finale PfCDPK4 model binne die gunstigste of lae-energie gebiede lê (Figuur 4: A, B, L), wat die kwaliteit van die model bevestig.

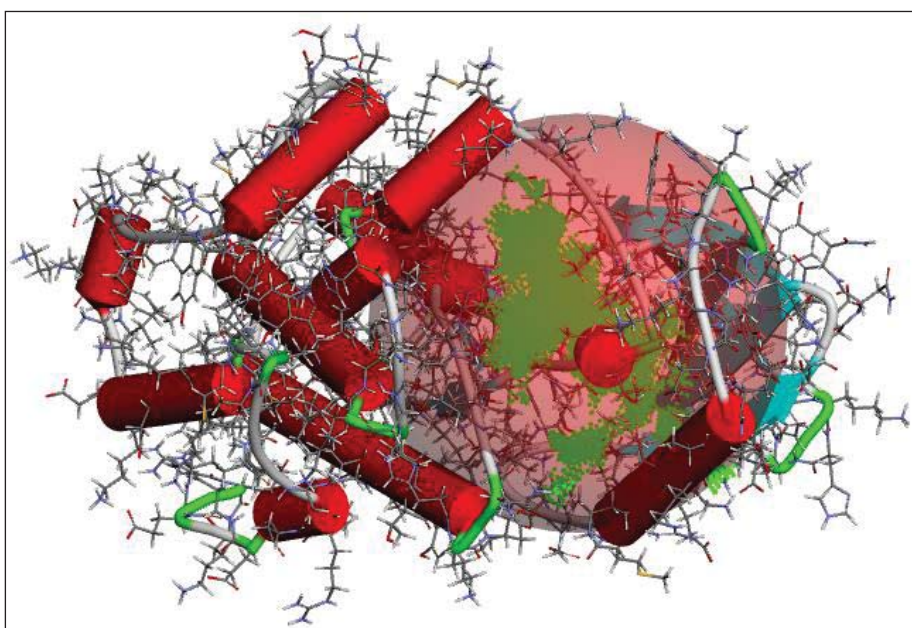
Teiken en biblioteekvoorbereiding

PfCDPK4 modeloptimering en bindingsposisie

Die ATP-bindingsdomein van die modelstruktuur is geïdentifiseer deur middel van 'uitveër'-algoritmes. Sewe verskillende moontlikhede is gevind, waarvan een uitgestaan het (Figuur 5) op grond van bestaande kennis van ander proteïenkinase ATP-bindingsposisies, soos bv. ERK2 (*Extracellular Signal-Regulated Kinase 2*) (Lounnas et al., 2013). Die ATP-bindingsposisies in verskillende kinases stem grootliks ooreen, wat selektiwiteitsprobleme kan veroorsaak (Huang et al., 2010a).



FIGUUR 4: Ramachandran-diagram van die PfCDPK4 model bereken deur middel van PROCHECK (Arnold et al., 2006). Rooi – gunstigste gebied; geel – toegelate gebied; liggeel – minder gunstig; wit – ongunstig.



FIGUUR 5: Die finale PfCDPK4 modelstruktuur met sy ATP bindingsetel aangedui as groen kulle en die deursigtige omringende sfeer in ligrooi.

Die volledige katalitiese domein van die kinase bestaan uit 259 residue (Aher en Roy, 2016) en het twee lobbe (*N*-terminale lob en *C*-terminale lob) verbind deur 'n segment wat 'n 'skarnierlus' genoem word. Die *N*-terminale lob is klein en bestaan hoofsaaklik uit vyf antiparallele β -plate en 'n gekonserveerde α -heliks, terwyl die *C*-terminale lob 'n α -heliks vorm. Die gaping tussen die twee lobbe vorm die aktiewe gebied en bevat die residue wat direk betrokke is by katalise of ATP-binding (Schwartz en Murray, 2011). Vulpetti en Bosotti het 'n reeks van 38 residue binne die vyf ATP-bindingsdomeine gevind wat direk betrokke is by ATP binding deur die proteïenkinases (Vulpetti en Bosotti, 2004), terwyl Huang *et al.* bepaal het dat 36 residue betrokke is (Huang *et al.*, 2010b). Die vyf ATP-bindingsdomeine is die adenosien bindingsgebied ('kinase-skarniergebied'), hidrofobiese gebied, fosfaatbindingsgebied, ribosebindingsgebied, en inganggebied. Met behulp van hierdie inligting was dit moontlik om die korrekte gebied te identifiseer.

Bereiding en optimering van die Biofocus bibliotek, NSV versameling en kinase-inhibitore

In die siftingproses van die Biofocus bibliotek is 19 351 verbindings wat aan die 'Lipinski Reël van 5' (Lipinski *et al.*, 1997) voldoen, geselekteer. Die NSV versameling het 23 verbindings opgelewer. Voorbereiding van al drie versamelings het in totaal 83 780 geoptimeerde strukture opgelewer, wat tautomere, stereoisomere, en geïoniseerde vorms insluit.

Virtuele en in vitro sifting van klein molekule teen die moontlike teiken *PfCDPK4*

Virtuele sifting teen die *PfCDPK4* 3D modelstruktuur

In silico sifting van die 83 707 Biofocus strukture het 195 434 poses opgelewer, wat met behulp van die funksies PLP1, PLP2, PMF, PMF04, Jain, LigScore1, en LigScore2 (Krammer *et al.*, 2005) geëvalueer is. 'n Konsensustelling is bereken om ligande te vind wat konsekwent hoë tellings in meer as een funksie behaal. Dit het 26 prioriteitverbindings opgelewer (Tabel 4, Figuur 6), wat geskeduleer is vir *in vitro* analise.

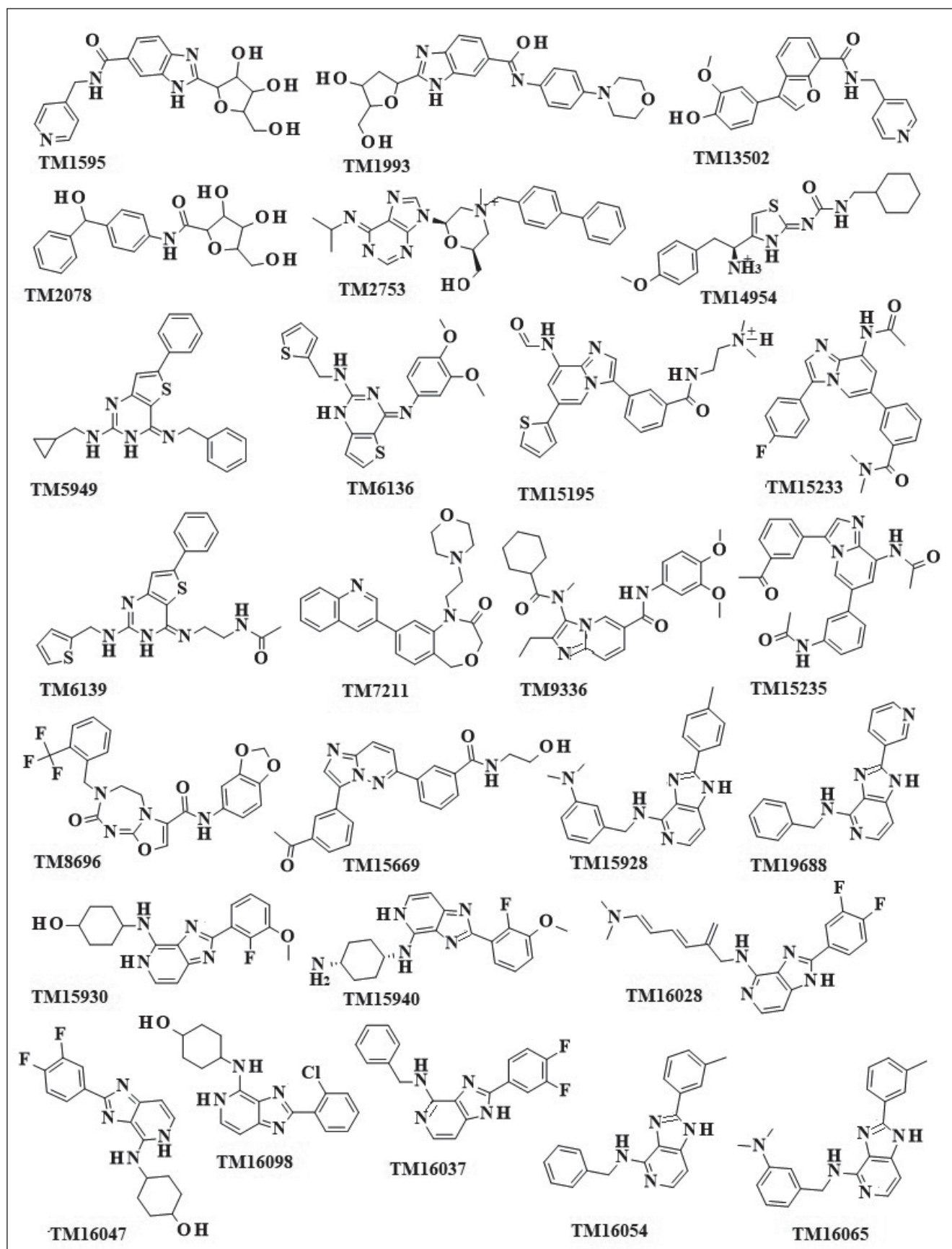
Die hoogste konsensustelling van 6 is vir verbindings TM1595, TM8696 en TM15669 verkry, en is as hoogs-belowend beskou.

Die mees-voorkomende aminosuurreidue wat aan waterstofbindings in hierdie 26 Biofocus-*PfCDPK4* komplekse deelneem, is Asp145, Lys8, Glu84, Tyr80 en Lys125. Die interaksies van sommige van die *PfCDPK4* model se aminosure met atome van drie geselekteerde verbindings word in Figuur 7 gewys.

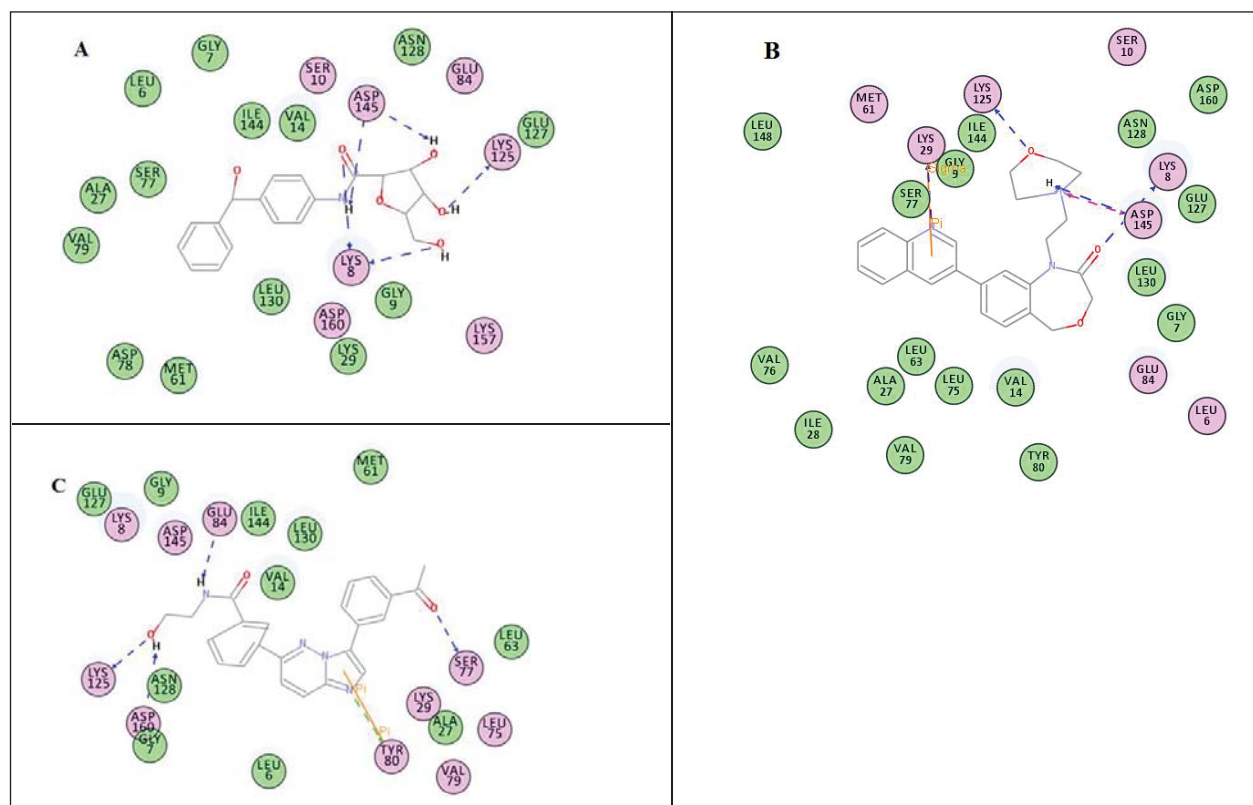
Die dok van die 37 geoptimeerde NSVs teen *PfCDPK4* het 308 poses opgelewer. Gebaseer op die konsensustellings is ses verbindings geselekteer (Tabel 5 en Figuur 8). Neurolefin C (Francois *et al.*, 1996) en calothrixin A en B (Rickards *et al.*, 1999) is van die treffers met die hoogste

TABEL 4: Dok-resultate teen die *PfCDPK4* modelstruktuur vir die prioriteitverbindings uit die Biofocus bibliotek.

Verbinding	LigScore1	LigScore2	PLP1	PLP2	Jain	PMF	PMF04	Konsensus
TM1595	4.77	3.59	102.27	103.59	7.29	117.29	66.76	6
TM1993	3.82	3.00	102.85	103.14	6.28	132.98	62.55	5
TM2078	4.28	3.88	90.44	94.28	6.49	132.18	66.39	5
TM2753	-0.04	-0.93	120.34	108.71	7.39	114.52	60.11	5
TM5949	4.30	5.74	111.52	104.54	7.74	90.67	42.12	5
TM6136	3.90	5.46	100.43	93.18	2.87	106.47	50.88	5
TM6139	4.27	4.93	102.28	90.35	5.17	119.36	64.92	5
TM7211	3.15	3.47	99.38	92.95	8.89	112.91	67.88	5
TM8696	3.73	3.67	111.20	97.13	6.69	115.14	74.85	6
TM9336	-0.50	-3.36	102.97	104.00	6.56	127.65	70.04	5
TM13502	4.61	4.75	98.21	95.67	4.31	108.29	42.51	5
TM14954	4.02	4.52	97.09	94.03	6.51	104.40	35.76	5
TM15195	2.68	0.94	100.73	102.02	7.21	129.57	59.23	5
TM15233	4.83	5.29	97.48	90.04	5.68	116.60	62.29	5
TM15235	4.55	4.80	103.18	95.55	6.32	115.46	46.57	5
TM15669	4.17	4.57	100.52	97.84	3.47	126.25	59.32	6
TM15928	3.97	3.38	98.84	98.10	7.33	82.90	65.13	5
TM15930	3.46	4.31	99.54	98.11	4.12	110.19	75.57	5
TM15940	2.18	2.89	99.48	96.33	6.65	114.89	66.61	5
TM16028	3.58	4.11	102.76	99.06	4.99	104.92	77.23	5
TM16037	4.28	3.28	106.62	105.12	7.49	68.36	64.65	5
TM16047	4.28	2.87	103.48	105.94	6.74	94.53	76.42	5
TM16054	4.49	4.41	101.51	101.00	7.73	86.10	64.18	5
TM16065	3.44	2.57	102.01	102.32	8.01	94.92	70.02	5
TM16098	4.00	4.23	97.01	101.90	7.11	102.03	70.72	5
TM19688	4.14	4.04	97.22	96.32	6.43	87.94	66.52	5



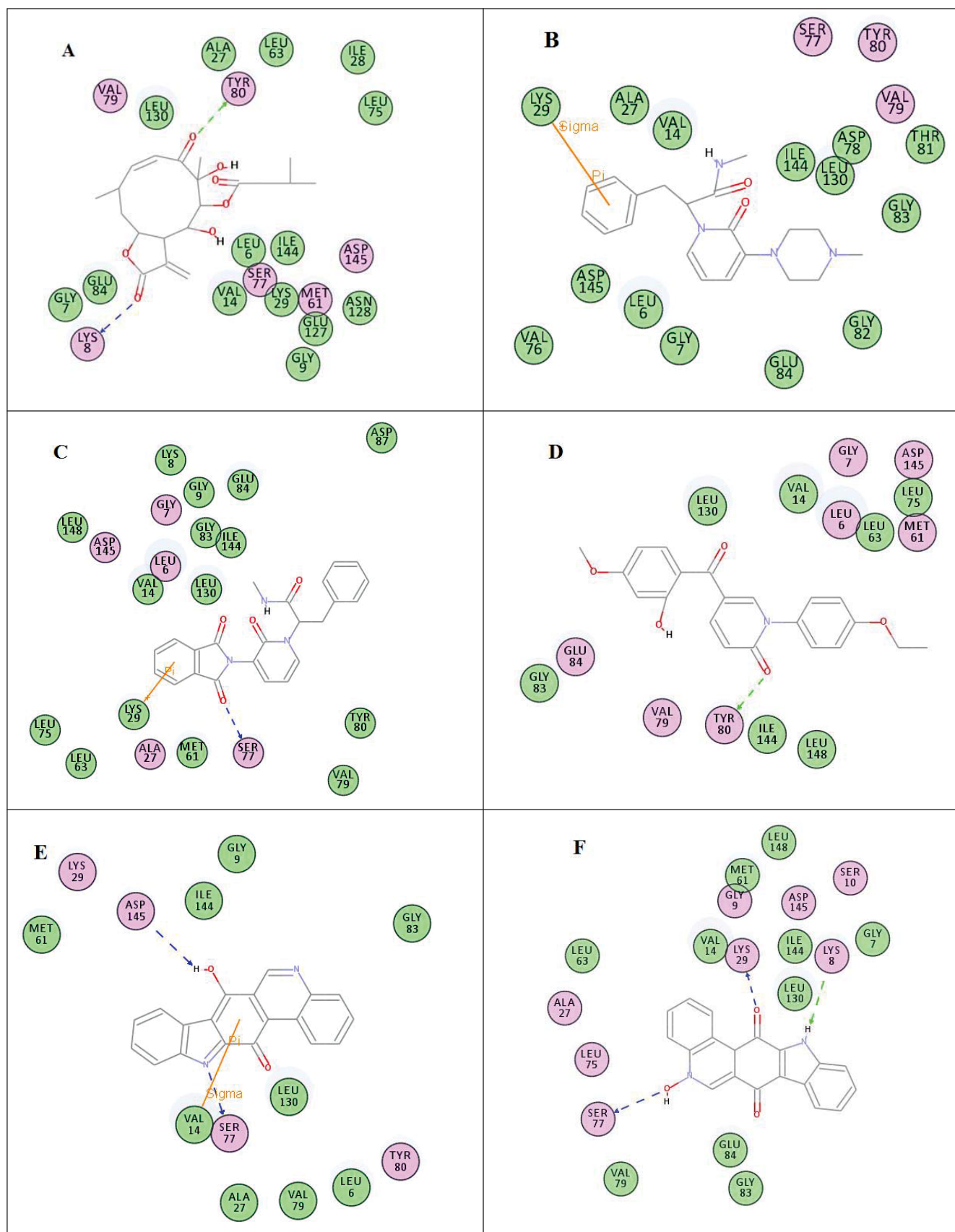
FIGUUR 6: Strukture van die prioriteitbindings verkry deur virtuele sifting van die geoptimeerde Biofocus biblioteek teen die *Pf*CDPK4 modelstruktuur.



FIGUUR 7: Twee-dimensionele voorstellings toon die interaksies van die Biofocus verbindings met die hoogste moontlike affiniteit vir die *PfCDPK4* aktiewe setel. Kompleks A: TM2078-*PfCDPK4*; kompleks B: TM7211-*PfCDPK4*; kompleks C: TM15669-*PfCDPK4*. Die blou en groen pyle dui op waterstofbindingsinteraksies, terwyl die oranje pyle π - σ en π - π interaksies aandui.

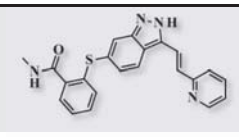
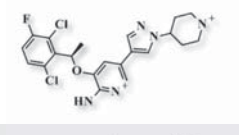
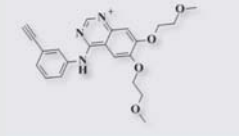
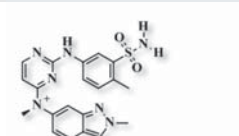
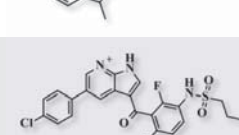
TABEL 5: Geselekteerde Natuurlike en Sintetiese Verbindings (NSVs) met die hoogste voorspelde bindingsaffiniteit vir die *PfCDPK4* model. NeuroC; Neurolefen C; CaloA; Calothrixin A; CaloB; Calothrixin B; Compd3; Compd4; Compd5.

Verbinding	LigScore1	LigScore2	PLP1	PLP2	Jain	PMF	PMF04	Konsensus-telling	Struktuur
NeuroC	3.12	4.72	76.89	72.40	5.49	77.17	14.59	5	
Compd5	2.75	5.56	85.65	77.57	4.60	54.23	14.98	6	
Compd4	3.56	5.76	94.67	83.01	3.85	91.67	26.07	7	
Compd3	2.56	5.22	76.51	69.47	2.13	84.06	29.38	5	
CaloB	2.95	5.44	77.24	71.24	3.02	90.05	38.30	5	
CaloA	4.17	5.72	83.52	77.28	3.68	71.03	35.18	6	



FIGUUR 8: Twee-dimensionele voorstellings toon die interaksies van die Natuurlike en Sintetiese Verbindings (NSVs) met die hoogste moontlike affiniteit vir die *PfCDPK4* aktiewe posisie. Kompleks A: NeuroC; kompleks B: Compd5; kompleks C: Compd4; kompleks D: Compd3; kompleks E: CaloB; kompleks F: CaloA. Die blou en groen pyle dui op waterstofbindingsinteraksies, terwyl die oranje pyle π - σ en π - π interaksies aandui.

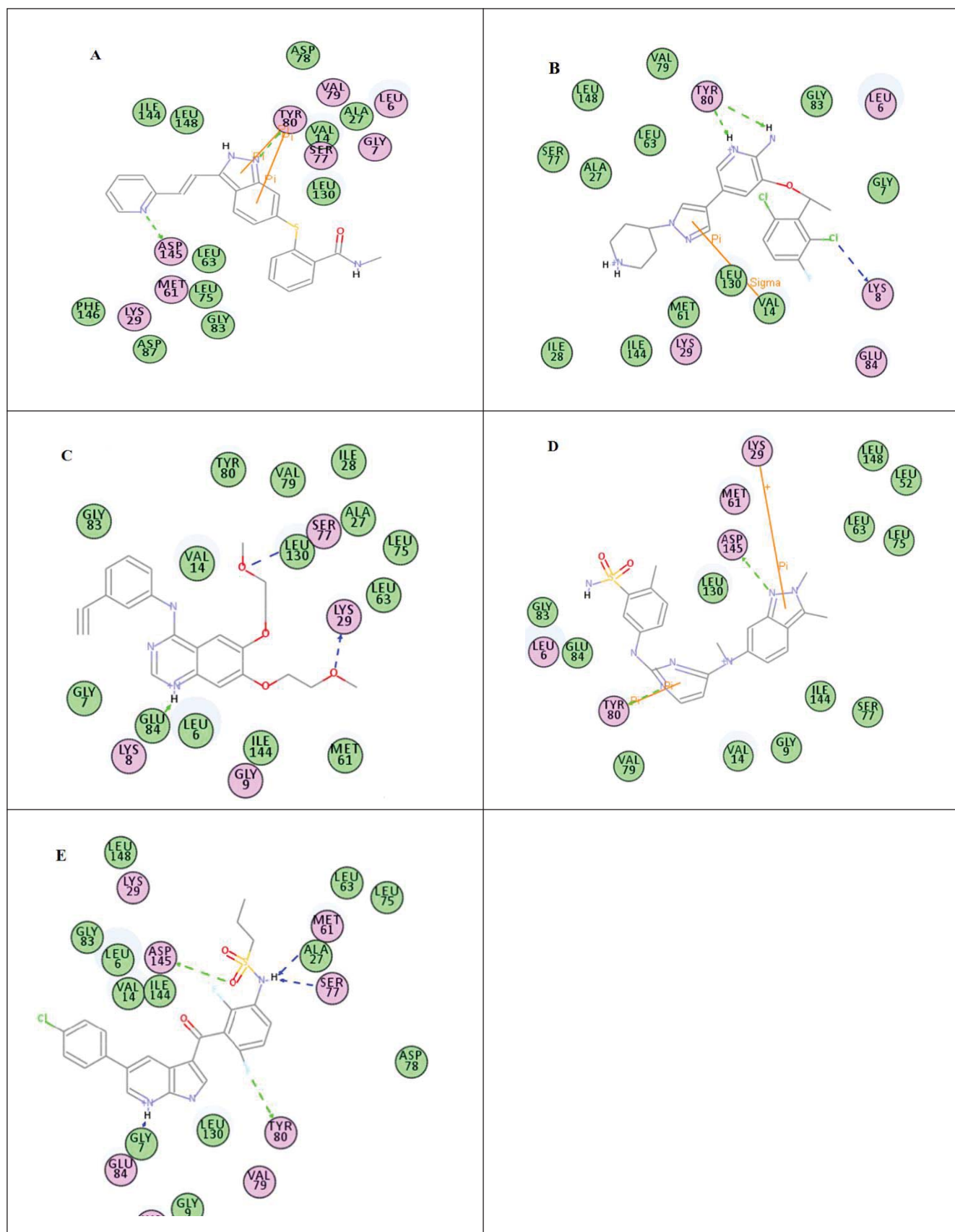
TABEL 6: Vyf belowende verbindings / geselekteerde kinase-inhibitore vanuit die groep van 36 bekende kinase-inhibitore (Zhao et al., 2013) wat met die *Pf*CDPK4 model gedok is.

Verbinding Nr	LigScore1	LigScore2	PLP1	PLP2	Jain	PMF	PMF04	Konsensus-telling	Struktuur
6450551	3.22	5.51	90.79	79.70	3.67	99.60	40.30	7	
11626560	3.43	6.16	96.62	82.45	3.88	82.74	51.45	5	
176870	3.34	5.41	85.66	79.34	2.44	65.39	32.29	5	
10113978	3.44	5.49	94.38	86.11	6.01	101.59	50.07	6	
42611257	4.37	5.74	81.07	79.53	5.22	87.50	43.14	5	

voorspelde bindingsaffiniteit. Dit is natuurprodukte met bewese *in vitro* antimalaria-aktiwiteit, alhoewel hulle sellulêre teikens nog nie vasgestel is nie. Verdere ondersoek (bv. deur kristallisasie) word dus benodig om vas te stel of *Pf*CDPK4 inderdaad die teiken van hierdie verbindings is.

Die 36 bekende kinase-inhibitore (Zhao et al. 2013) wat met

die *Pf*CDPK4 model gedok is, het 175 poses opgelewer. Die konsensusstelling het vyf belowende verbindings (6450551, 11626560, 176870, 10113978 en 42611257) opgelewer (Tabel 6, Figuur 9). Hierdie verbindings was ongelukkig nie beskikbaar vir *in vitro* toetsing nie, maar die *in silico* resultate vir hierdie verbindings is belowend genoeg om hulle verder te ondersoek as moontlike middels wat *Pf*CDPK4 sal teiken.

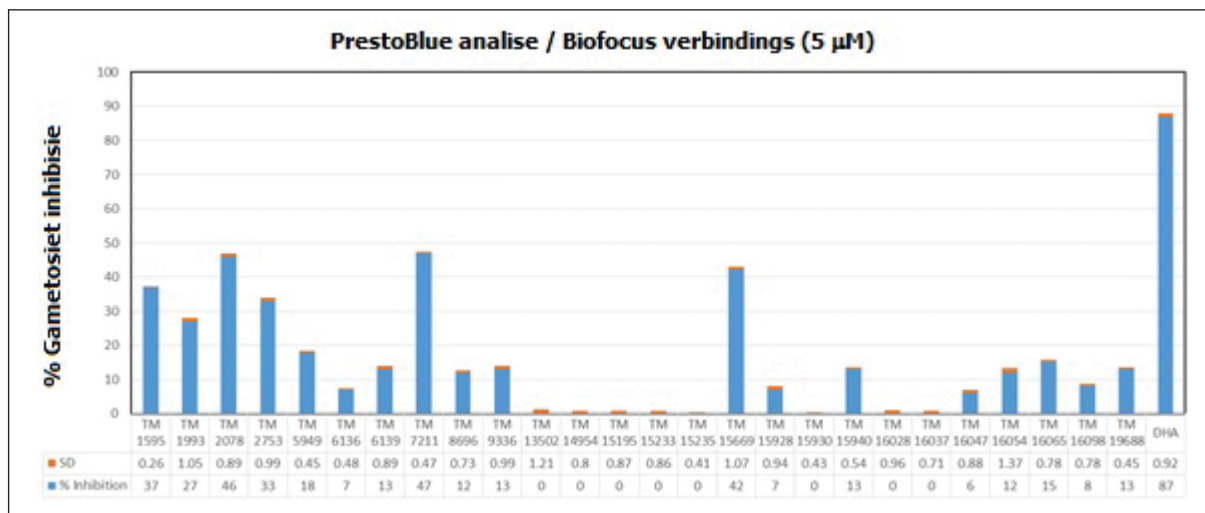


FIGUUR 9: Twee-dimensionele voorstelling van interaksies van bekende kinase-inhibitore gedok met *PfCDPK4*. A: 6450551. B: 11626560. C: 176870. D: 10113978. E: 42611257. Blou en groen pyle dui op waterstoffbinderingsinteraksies, terwyl die oranje pyle π - σ en π - π interaksies aandui.

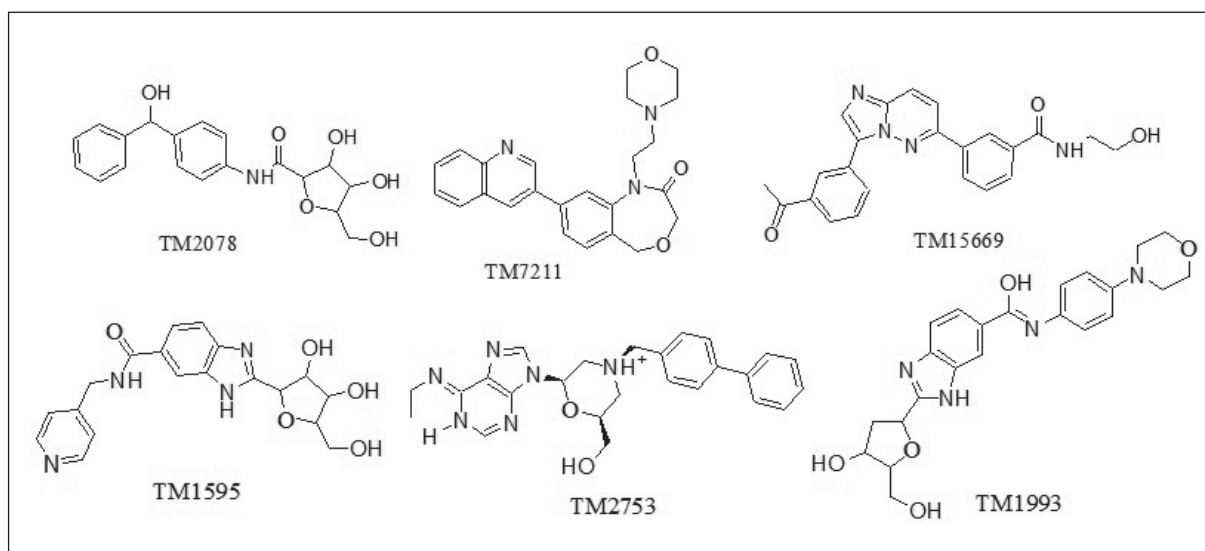
In vitro sifting van Biofocus verbindings teen *Plasmodium falciparum* stam NF54

Die aktiwiteit van CDPK4 word geassosieer met die gametosietstadium van die parasiet, en dus het ons besluit om die potensiaal van die 26 voorkeur verbindings om hierdie proteïene te teiken, te ondersoek. Die verbindings is *in vitro* getoets vir gametosiet-dodende aktiwiteit teen die *Plasmodium falciparum* NF54 stam, teen 'n konsentrasie van 5 μ M (Figuur 10). Verbindings wat minder as 20% inhibeer

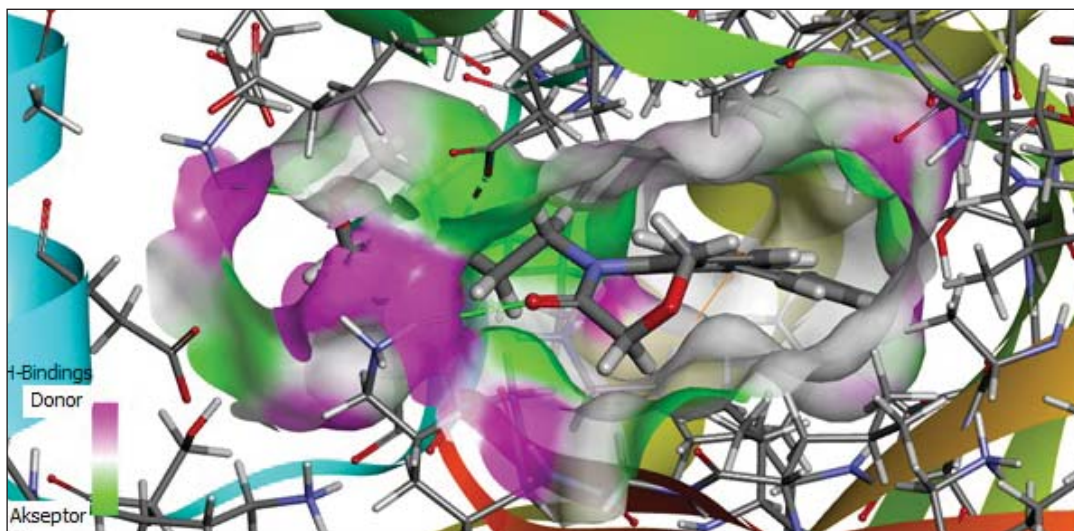
is as onaktief geklassifiseer, die wat minstens 50% inhibeer as marginaal aktief, en die wat meer as 70% inhibeer, as aktief. Van hierdie verbindings is drie as marginaal aktief geklassifiseer (TM2078 46%, TM7211 47%, en TM15669 42%, Figuur 11), in ooreenstemming met hulle dok-resultate. Drie ander verbindings het meer as 20% inhibisie getoon: TM1595 37%, TM2753 33% en TM1993 27% (Figuur 11). Die plasing van TM7211 in die model word in Figuur 12 getoon.



FIGUUR 10: *In vitro* siftingsresultate teen *Plasmodium falciparum* malaria stam NF54 vir 26 geselekteerde Biofocus verbindings.



FIGUUR 11: Ses Biofocus verbindings met lae to marginale aktiwiteit teen *Plasmodium falciparum* malaria stam NF54.



FIGUUR 12: Passing van verbinding TM2711 in die aktiewe setel van die *PfCDPK4* model.

Gevolgtrekkings

Homologie-modellering is 'n betroubare en akkurate metode om die drie-dimensionele strukture van meeste proteïene te bepaal waarvoor eksperimenteel-bepaalde strukture nog nie bekend is nie. Ten tyde van hierdie studie was die struktuur van *PfCDPK4* nie in proteïenstruktuur-databasisse beskikbaar nie, sodat ons homologiemoellering toegepas het met die X-straal kristallografiese struktuur van *TgCDPK1* (PDB kode: 3ma6-b) as 'n templaar om 'n geskikte *PfCDPK4* struktuur te genereer vir geneesmiddelontdekking. Met 79% identiteit was dit duidelik dat *TgCDPK1* 3ma6 die beste templaar is om die drie-dimensionele struktuur van *PfCDPK4* op te baseer. Die verfynde struktuur het ook goeie stereochemiese gehalte getoon en is dus verder ontwikkel om die geskikte bindingsposisie te vind.

Die ligande wat in die dok-studies gebruik is, is eerstens gefiltreer met behulp van die Lipinski reëls, en vervolgens verder geselekteer deur middel van *in silico* sifting. Daar is aangetoon dat etlike *Biofocus* verbindings en sommige klinies-goedgekeurde kinase-inhibitore en natuurlike en sintetiese verbindings 'n affiniteit vir die teiken het. *In vitro* sifting het die moontlike aktiwiteit van minstens drie verbindings teen *Plasmodium falciparum* stam NF54 aangetoon. Hierdie resultate bevestig die gebruik van die *PfCDPK4* modelstruktuur as teiken in die soektog na nuwe verbindings vir die ontwikkeling van antimalaria geneesmiddels, veral deur die bestryding van transmissie van die siekte.

Erkennings

Die Nasionale Navorsingstigting en die Universiteit van Venda word bedank vir finansiële ondersteuning. Die Wetenskaplike en Nywerheidsnavorsingsraad (WNNR-Biowetenskappe) word bedank vir gebruik van hulle hulpbronne, en die Sentrum vir Hoëverrigtingberekening

(Centre for High Performance Computing – CHPC) word bedank vir toegang tot die BIOVIA (voorheen Accelrys) *Discovery Studio* lisensie.

Verwysings

- Accelrys Software, 2012. *Discovery Studio Modeling Environment*, in *Release 3.5*. Accelrys Software Inc, San Diego. [<http://www.accelrys.com/products/discovery-studio/>]
- Aher RB, Roy K. 2016. Exploring structural requirements for the inhibition of *Plasmodium falciparum* calcium-dependent protein kinase-4 (*PfCDPK4*) using multiple *in silico* approaches. *RSC Advances* 6, 51957-51982.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389-3402.
- Anastasiadis T, Deacon SW, Devarajan K, Ma H, Peterson JR. 2011. Comprehensive assay of kinase catalytic activity reveals features of kinase. *Nature Biotechnology* 29, 1039-1045.
- Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195-201.
- Arnold K, Kiefer F, Kopp J, Battey JD, Podvina M, Westbrook J, et al. 2009. The protein model portal. *Journal of Structural and Functional Genomics* 10, 1-8.
- Bansal A, Singh S, More KR, Hans D, Nangalia K, Yogavel M, et al. 2013. Characterization of *Plasmodium falciparum* calcium-dependent protein kinase 1 (*PfCDPK1*) and its role in microneme secretion during erythrocyte invasion. *Journal of Biological Chemistry* 288, 1590-602.
- Biamonte MA, Wanner J, Le Roch KG. 2013. Recent advances in malaria drug discovery. *Bioorganic & Medicinal Chemistry Letters* 23, 2829-2843.
- Billker O, Dechamps S, Tewari R, Wenig G, Franke-Fayard B, Brinkmann V. 2004. Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* 117, 503-514.
- Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. 2008. Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols* 4, 1-13.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* 4, 187-217.
- Cavasotto CN, Phatak SS. 2009. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* 14, 676-683.
- Chapman TM, Osborne SA, Bouloc N, Large JM, Wallace C, Birchall K, et al. 2013. Substituted imidazopyridazines are potent and selective inhibitors of *Plasmodium falciparum* calcium-dependent protein kinase 1 (*PfCDPK1*). *Bioorganic & Medicinal Chemistry Letters* 23, 3064-3069.
- Deshpande N, Address KJ, Bluhm WF, Merino-Ott JC, Townsend-Merino W, Zhang Q, et al. 2005. The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Research* 33, D233-D237; <http://www.rcsb.org/pdb/>

- Dorin-Semblat D, Carvalho, TG, Nivez M-P, Halbert J, Poulet P, Semblat J-P, et al. 2013. An atypical cyclin-dependent kinase controls *Plasmodium falciparum* proliferation rate. *Kinome* 1, 4-16.
- François CJJ, Klomp JPG, Knegtel RMA. 2000. Sequence annotation of nuclear receptor ligand-binding domains by automated homology modeling. *Protein Engineering* 13, 391-394.
- Francois G, Passreiter CM, Woerdenbag HJ, Van Looveren M. 1996. Ant plasmodial activities and cytotoxic effects of aqueous extracts and sesquiterpene lactones from *Neurolaena lobata*. *Planta Medica* 62, 126-129.
- Hooti RWW, Sander C, Vriend G. 1997. Objectively judging the quality of a protein structure from a Ramachandran plot. *CABIOS Bioinformatics* 13(4), 425-430.
- Hovmoller S, Zhou T, Ohlson T. 2002. Conformations of amino acids in proteins. *Acta Crystallographica Section D* 58, 768-776.
- Huang H-J, Yu HW, Chen C-Y, Hsu C-H, Chen H-Y, Lee K-J, et al. 2010. Current developments of computer-aided drug design. *Journal of the Taiwan Institute of Chemical Engineers* 41, 623-635.
- Huang D, Zhou T, Lafleur K, Nevado C, Cafilisch A. 2010. Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* 26, 198-204.
- Kang BS, Pugalendhi GK, Kim KJ. 2017. Binding direct on-based two-dimensional flattened contact area computing algorithm for protein-protein interactions. *Molecules* 22(10), 1722. doi: 10.3390/molecules22101722
- Karplus M. 2003. Molecular dynamics of biological macromolecules: a brief history and perspective. *Biopolymers* 68, 350-358. doi: 10.1002/Bip.10266.
- Klein EY. 2013. Ant malarial drug resistance: a review of the biology and strategies to delay emergence and spread. *International Journal of Antimicrobial Agents* 41, 311-317.
- Krammer A, Kirchhoff PD, Jiang X, Venkatachalam CM, Waldman M. 2005. LigScore: a novel scoring function for predicting binding affinities. *Journal of Molecular Graphics and Modelling* 23, 395-407.
- Kumar S, Kumar M, Ekka R, Dvorin JD, Paul AS, Madugundu AK, et al. 2017. PfCDPK1 mediated signaling in erythrocytic stages of *Plasmodium falciparum*. *Nature Communications* 8, Article 63.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM, 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26, 283-291.
- Lim DC, Cooke BM, Doerig C, Saeji JPI. 2012. Toxoplasma and *Plasmodium* protein kinases: roles in invasion and host cell remodelling. *International Journal for Parasitology* 42, 21-32.
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* 23, 3-25.
- Lounnas V, Ritschel T, Kelder J, McGuire R, Bywater R, Foloppe N. 2013. Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Computational and Structural Biotechnology Journal* 5, 1-14.
- Murray CJL, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, et al. 2012. Global malaria mortality between 1980 and 2010: A systematic analysis. *The Lancet* 379, 413-431.
- Ojo KK, Pfander C, Mueller NR, Burstroem C, Larson ET, Bryan CM, et al. 2012. Transmission of malaria to mosquitoes blocked by bumped kinase inhibitors. *The Journal of Clinical Investigation* 122, 2301-2305.
- Ojo KK, Eastman RT, Vidadala R, Zhang Z, Rivas KL, Choi R, et al. 2014. A specific inhibitor of PfCDPK4 blocks malaria transmission: chemical-genetic validation. *Journal of Infectious Diseases* 209, 275-284.
- Ramachandran GN, Ramakrishnan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* 7, 95-99.
- Ranjan R, Ahmed A, Gourinath S, Sharma P. 2009. Dissect on of mechanisms involved in the regulation of *Plasmodium falciparum* calcium-dependent protein kinase 4. *Journal of Biological Chemistry* 284, 15267-15276.
- Reader J, Botha M, Theron A, Lauterbach SB, Rossouw C, Engelbrecht D, et al. 2015. Nowhere to hide: interrogating different metabolic parameters of *Plasmodium falciparum* gametocytes in a transmission blocking drug discovery pipeline towards malaria elimination. *Malaria Journal* 14, 213.
- Regad L, Martin J, Nuel G, Camproux AC. 2010. Mining protein loops using a structural alphabet and statistical exceptionality. *BMC Bioinformatics* 11, 75.
- Rickards RW, Rothschild JM, Willis AC, De Chazal NM, Kirk J, Kirk K, et al. 1999. Calothrixins A and B, novel pentacyclic metabolites from Calothrix cyanobacteria with potent activity against malaria parasites and human cancer cells. *Tetrahedron* 55, 13513-13520.
- Schwartz PA, Murray BW. 2011. Protein kinase biochemistry and drug discovery. *Bioorganic Chemistry* 39, 192-210.
- Šali A, Potterton L, Yuan F, Van Vlijmen H, Karplus M. 1995. Evaluation of comparative protein modeling by MODELER. *Proteins: Structure, Function, and Bioinformatics* 23, 318-326, 1995.
- Shen MY, Šali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Science* 15, 2507-2524.
- Taha MO, Habash M, Khanfar MA. 2014. The use of docking-based comparative intermolecular contacts analysis to identify optimal docking conditions within glucokinase and to discover of new GK activators. *Journal of Computer-Aided Molecular Design* 28, 509-547; doi10.1007/s10822-014-9740-4
- Tang K, Zhang J, Liang J. 2014. Fast protein loop sampling and structure predict on using distance-guided sequential chain-growth Monte Carlo Method. *PLoS Computational Biology* 10, 1-16.
- Tewari R, Straschil U, Bateman A, Bohme U, Cherevach I, Gong P, et al. 2010. The systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators of mosquito transmission. *Cell Host & Microbe* 8, 377-387.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weight ng, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-4680.
- UniProt Consortium, 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Research* 40, D71-5; hc p://www.uniprot.org/uniprot/Q8IB55.
- Venkatachalam CM, Jiang X, Oldfield T, Waldman M. 2003. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics & Modelling* 21, 289-307.
- Vidadala RS, Ojo KK, Johnson SM, Zhang Z, Leonard SE, Mitra A. et al. 2014. Development of potent and selective *Plasmodium falciparum* calcium-dependent protein kinase 4 (PfCDPK4) inhibitors that block the transmission of malaria to mosquitoes. *European Journal of Medicinal Chemistry* 74, 562-573.
- Vulpetti A, Bosotti R. 2004. Sequence and structural analysis of kinase ATP pocket residues. *Farmaco* 59, 759-65.
- Vyas VK, Ukawala RD, Ghate M, Chintha C, 2012. Homology modeling a fast tool for drug discovery: current perspectives. *Indian Journal of Pharmaceutical Sciences* 74, 1-17.
- WHO (World Health Organization), 2011. Global plan for artemisinin resistance containment (GPARC). WHO, Geneva.
- WHO (World Health Organization), 2016. World Malaria Report 2016 Fact Sheet. WHO, Geneva.
- WHO (World Health Organization), 2017. World Malaria Report. WHO, Geneva.
- Wu G, Robertson DH, Brooks CL, Vieth M. 2003. Detailed analysis of grid-based molecular docking: A case study of CDOCKER – A CHARMM-based MD docking algorithm. *Journal of Computational Chemistry* 24, 1549-1562.
- Yu H, Jin H, Sun L, Zhang L, Sun G, Wang Z, et al. 2013. Toll-like receptor 7 agonists: chemical feature based pharmacophore identification and molecular docking studies. *PLoS One* 8, 1-11.
- Zhao HY, Wei H, Wang X. 2013. The reciprocal interaction of small molecule protein kinase inhibitors and ATP-binding cassette transporters in targeted cancer therapy. *Journal of Cancer Research Updates* 2, 68-86.