J Mol Model (2011) 17:3075–3084
DOI 10.1007/s00894-011-0962-2

ORIGINAL PAPER

Integrative computational protocol for the discovery of inhibitors of the *Helicobacter pylori* nickel response regulator (NikR)

Aldo Segura-Cabrera · Xianwu Guo · Arturo Rojo-Domínguez · Mario A. Rodríguez-Pérez

Received: 2 August 2010 / Accepted: 5 January 2011 / Published online: 1 March 2011
© Springer-Verlag 2011

Abstract In order to identify novel inhibitors of the *Helicobacter pylori* nickel response regulator (HpNikR) an integrative protocol was performed for half a million compounds retrieved from the ZINC database. We firstly implement a structure-based virtual screening to build a library of potential inhibitors against the HpNikR using a docking analysis (AutoDock Vina). The library was then used to perform a hierarchical clustering of docking poses, based on protein-contact footprints calculation from the multiple conformations given by the AutoDock Vina software, and the drug-protein interaction analyses to identify and remove potential promiscuous compounds likely interacting with human proteins, hence causing drug side effects. 250 drug-like compounds were finally proposed as non-promicuous potential inhibitors for HpNikR. These compounds target the DNA-binding sites of HpNikR so that HpNikR-compound binding could be able to mimic key interactions in the DNA-protein recognition process. HpNikR inhibitors with promising potential against *H. pylori* could also act against other human bacterial pathogens due to the conservation of targeting motif of NikR involved in DNA-protein interaction.

Keywords Contact activity relationship · Drug discovery · Drug-protein interactions · Structure-based drug design · Virtual screening

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-0962-2) contains supplementary material, which is available to authorized users.

A. Segura-Cabrera (✉)
Laboratorio de Bioinformática, Centro de Biotecnología Genómica, Instituto Politécnico Nacional,
Boulevard del Maestro esquina Elías Piña,
Colonia Narciso Mendoza,
88710 Ciudad Reynosa, Tamaulipas, Mexico
e-mail: asegurac@ipn.mx

A. Segura-Cabrera
e-mail: aldosegura@gmail.com

X. Guo · M. A. Rodríguez-Pérez
Laboratorio de Biomedicina Molecular,
Centro de Biotecnología Genómica,
Instituto Politécnico Nacional,
Boulevard del Maestro esquina Elías Piña,
Colonia Narciso Mendoza,
88710 Ciudad Reynosa, Tamaulipas, Mexico

A. Rojo-Domínguez
Departamento de Ciencias Naturales,
Universidad Autónoma Metropolitana,
Unidad Cuajimalpa, MexicoDF

Introduction

Helicobacter pylori (*H. pylori*) is a Gram-negative, microaerophilic, helical-shaped bacterium that colonizes the human stomach of at least half of the world's population [1, 2]. *H. pylori* has been associated to several diseases such as chronic gastritis, peptic ulcer disease, and predisposition to the development of gastric cancer and other gastric malignancies [1, 2]. A triple combination chemotherapy consisting of a proton pump inhibitor and two antibiotics (either clarithromycin, metronidazole or amoxicillin) has become the election treatment for eradication of *H. pylori* [3, 4]. Although the triple combination chemotherapy succeeds in 70-90% of the cases [5–9], it may fail because of the presence of *H. pylori* antibiotic-resistant strains [5–11]. The resistant *H. pylori* strains to clarithromycin, metronidazole, and amoxicillin vary among regions and countries, but resistant *H. pylori* strains to metronidazole are relatively common worldwide [12, 13]. As resistant

H. pylori strains to antibiotics may eventually compromise the effectiveness of current therapy to fight the disease produced by these resistant strains in humans, the discovery of potential drugs against *H. pylori* may help guide the development of complementary schemes of drug chemotherapy [1, 10, 14, 15].

H. pylori bacteria colonize the mucus layer overlaying the gastric epithelium of humans where the pH is thought to vary between 4 and 6 with temporary acid shocks of pH~2 [16, 17]. Thus, the growth of *H. pylori* in such hostile environment of the stomach depends on particular physiological mechanisms essential for the bacterial survival. Specifically, tolerance of *H. pylori* to acid shocks requires production of ammonia by urease-mediated degradation of urea. *H. pylori* produce large amounts of urease, which is a nickel dependent enzyme [18–20]. Nickel is essential for the activity of some enzymes, and its concentrations must be under strict control because high concentration could be toxic for the microorganisms [21–25]. Therefore, the essentialness of nickel homeostatic systems in *H. pylori* provides many potential targets for the development of novel antimicrobial agents. The sensitive mechanism of *H. pylori* nickel homeostasis is controlled by nickel-responsive regulatory protein, NikR (HpNikR), which is a transcription factor that belongs to the ribbon-helix-helix (RHH) family of DNA binding proteins [26]. The NikR protein is a homotetramer; each subunit consisting of an N-terminal DNA-binding domain, homologous to the Arc/CopG/MetJ/Mnt family of RHH regulators, and a C-terminal domain that is required for binding of nickel and for tetramerization [27–29]. HpNikR is able to regulate multiple genes, including genes that encode to nickel dependent proteins, such as NixA, HPN and UreA, as well as genes that encode proteins that do not use nickel, such as those involved in iron uptake and storage, motility, stress response, and genes that encode outer membrane proteins [22, 30]. Moreover, biochemical, structural, and functional genomics studies of HpNikR have suggested that HpNikR is a master regulator of the expression of a cluster of related genes for *H. pylori* acid adaptation [24, 31–33]. Therefore, drugs targeting HpNikR could break the nickel homeostasis to prevent *H. pylori* infections [15]. In addition, no NikR homolog exists in human that makes NikR an ideal target for antibacterial drugs.

It has recently been suggested that targeting protein-DNA interactions could serve as a new paradigm of drug discovery because a single regulatory protein may disrupt, simultaneously, several physiological processes [34]. The typical antimicrobial drug design targets active sites of enzymes that are essential for metabolic process, which elicits strong selective pressure for resistance development [35]. Bowser et al. [36] identified potent small-molecule inhibitors for the AraC family of bacterial transcription factors. Such compounds targeting regulators may delay the appearance of resistant bacterial strains. The present work reports potential inhibitors of HpNikR using an integrative computational protocol including virtual screening, clustering analyses of docked poses, and protein-drug interaction networks. A set of 250 drug-like compounds were selected as potential inhibitors for HpNikR, which could provide an opportunity to produce novel drugs targeting the protein regulator that interacts with DNA in pathogenic bacteria.

Materials and methods

Docking

The AutoDock Vina software [37] (Vina) was used for flexible docking simulations into DNA binding site of HpNikR. The coordinates of crystal structure of HpNikR were downloaded from the Protein Data Bank [38], PDB code 2CA9 [39], and were setup as the receptor for docking protocol. Before docking, small ligands and water molecules were removed manually from the receptor. Polar hydrogens, Gasteiger charges and Vina configuration file were assigned using the AutoDock Tools interface [40, 41].

The compound for positive control (as defined below) was used to determine the size of search spaces on the DNA-binding site where the docking simulations were performed. We adapted the method of Troot et al. [37] to meet this purpose. First, the positive control was placed to the DNA-binding site of HpNikR. Second, several rounds of dockings were carried out to increase the size of search space. Finally, we obtained the optimal search spaces per each dimension (x, y and z) when the longest space kept the ligand inside the DNA-binding site. The size of search spaces in each dimension was of 14 Å and its center of 36.722, 41.543, and 8.499 for x, y and z, respectively. Thus, the search space was large enough for the ligand to rotate, as suggested elsewhere [42].

Using the application of Lipinski's "rule of fives" [43] (molecular weight<500, partition coefficient logP<5, number of H-bond donor<5, and number of H-bond acceptor<10) and the three-dimensional structures in mol2 format, half a million compounds were extracted from ZINC 8 database [44]. Therefore, these compounds could be having molecular properties important for a drug's pharmacokinetics in the human body.

Compound library

Bowser et al. [36], found a set of *p*-amino-substituted analogs of 1-hydroxybenzimidazole as potent inhibitors of bacterial transcription factors. In that paper, the compound named as compound 40 showed the highest in vitro activity

[36]. Hence, the three-dimensional structure of compound 40 (Table 1) was introduced in the ZINC dataset as a positive control. The binding energy of the positive control (-3.45 kcal mol^{-1}) plus the standard deviation in the predicted binding free energy from Vina force field (-2.85 kcal mol^{-1}) [37] was set as the cut-off value to select potential inhibitors. Compounds with high binding-energy values above the cut-off were considered inactive compounds (inactive compounds set).

ZINC dataset per compound was prepared as follows: prepare_ligand4.py python script of AutoDockTools software [40], used to merge nonpolar hydrogens, add Gasteiger charges, and set up rotatable bonds for each ligand via AutoTors. This produces the corresponding pdbqt file format necessary for Vina. All AutoDockTools python scripts were embedded into our "in-house" Perl scripts. The Vina default optimization parameters were used. Thus, the top-binding energy compounds were selected by using the positive control binding-energy as the cut-off value.

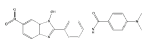
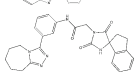
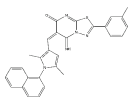
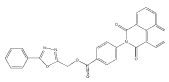
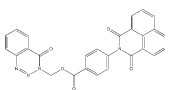
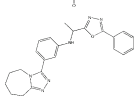
Clustering analysis of docked poses

The clustering methods applied to the analysis of protein-ligand interactions have been showed to improve the identification of correct conformations [45] and the discrimination of active compounds from inactive ones in docking experiments [46]. Then, in order for the improvement of the results of docking protocol above mentioned a clustering analysis was performed. The clustering analysis was carried out by AuPosSOM [46] program (Automatic analysis of Poses using SOM). This program performs the clustering in three steps: (a) a Kohonen self-organizing map (SOM) training phase using drug-protein contact descriptors followed by (b) an unsupervised cluster analysis and (c) a Newick file generation for results visualization as a tree [46]. The docking poses are then analyzed and classified automatically. The performance of clustering analysis was evaluated to discriminate active compounds from the inactive compounds by introducing, as control group, 10, 20, 30 and 40 compounds with high binding energy from inactive compounds set.

The AuPosSOM default parameters were used. These parameters included: the Euclidean SOM which was trained in two phases with the following parameters: map size 5×4, exponential decrease of learning rate and radius; phase 1: starting learning rate 0.2, starting radius 6, 1000 iterations; phase 2: starting learning rate 0.02, starting radius 3, 10,000 iterations. The Newick file produced by AuPosSOM was visualized as a tree with Seaview program [47]. The contact analysis resulted from AuPosSOM was also graphically examined using the software Chimera [48] to determine the key interactions, and complementarity with residues of the DNA-binding site. Thus, molecules are clustered according to their similar binding mode to the protein target. Potentially active and inactive compounds were then clustered into different groups.

In order to assess likely inhibition mechanism of potential inhibitors it was necessary to compare their interactions in the DNA-binding site with those in the HpNikR-DNA complex. However, to date there is no structural basis for the HpNikR-DNA complex. To explore this, prediction studies of putative residues involved in HpNikR-DNA recognition were performed using the TFmodeller [49] and DISPLAR [50] software. TFmodeller compiles information about protein-DNA interfaces deposited in the Protein Data Bank and uses it to model similar interfaces and produce a list of residues interacting with DNA related by homology modeling. DISPLAR employs a neural network method which predicts residues contacting with DNA molecules, if a particular structure of a protein

Table 1 The rank, ZINC code, AutoDock Vina binding score, chemical structure and cluster of top-five potential inhibitors of HpNikR, including the control compound, identified using the integrative protocol

Rank	ZINC code[44]	Vina binding energy (kcal/mol)	Chemical structure	Cluster
0	Positive control	−6.3 (cut-off value)		7
1	ZINC09252524	−7.7		4
2	ZINC09282661	−7.3		7
3	ZINC08805862	−7.2		7
4	ZINC08806007	−7.2		6
5	ZINC12788043	−7.2		3

known to bind DNA is given. The input to the neural network includes position-specific sequence profiles, solvent accessibilities per each residue and its spatial neighbors. The neural network is trained on known structures of protein-DNA complexes and could provide accuracy of prediction over 80%.

Drug-protein interaction networks

STITCH [51, 52] (Search Tool for Interactions of Chemicals) is a tool and database which includes a set of chemicals derived from PubChem [53]. The relationship between chemicals and proteins in STITCH are integrated from pathway and experimental databases, as well as those reported in the scientific literature. Using Openbabel software [54], a SMILES file of lowest-binding energy per compound was produced and later introduced into STITCH. The parameters set in STITCH were as follows: organism= *Homo sapiens*, active prediction methods as experiments, neighborhood, gene fusion, co-occurrence, co-expression, databases, and text-mining; confidence score of 0.4 (which ranged from medium to high confidence interactions); network depth of 1 (including only direct neighbors); and custom limit for interactors shown of 500. The queries in STITCH yield a list with match and Tanimoto score which was associated with their set of chemicals. A Tanimoto score higher than 0.6 was considered to be structurally similar to that of the match in STITCH.

Results and discussion

The scoring functions used by the docking software are still a major limiting factor in virtual screening process to find out active compounds [55]. For this reason, we implemented an integrative protocol that includes a structure-based high-throughput virtual screening, drug-protein interaction network and contact activity relationship (CAR) analyses for the discovery of inhibitors of HpNikR (Fig. 1). The docking protocol identified a total of 276 small molecules as potential inhibitors which showed lower binding-energy values, and have the lowest value of the binding-energy of -7.7 kcal mol^{-1} (Table 1). The compounds with lowest binding-energy usually share particular structural features such as the presence of three or four heteroaromatic systems of five or six members. Seven of 276 potential inhibitors showed the rhodanine scaffold, thus they could be considered as rhodanine derivatives. It is well known that rhodanine derivatives show a wide range of biological activities (i.e. antimicrobial, antifungal, antiviral, etc). In addition, no significant correlation (r=-0.13, n=276, P> 0.05) was found between the binding energy of the potential inhibitors and their molecular weight, indicating that predicted affinity is due to the specificity and not to the size of the molecule.

Bouvier et al. [46] have demonstrated that contact clustering (CAR analysis) is able to discriminate active compounds based only on protein-contact footprints calculation from the poses given by the docking software. Thus, the CAR analysis of 276 initial potential inhibitors and control group was performed. The analysis of CAR results showed that initial potential inhibitors and control group were distributed in nine clusters. The clustering tree and protein-contact footprints of one representative hit per cluster are shown in Fig. 2. Each one of the clusters includes compounds that interact in a particular way with residues of HpNikR. For example, the cluster 1 contains 42 compounds that interact predominantly with R37 of chain A and R12, F13, S14 of chain A and B. The cluster 2

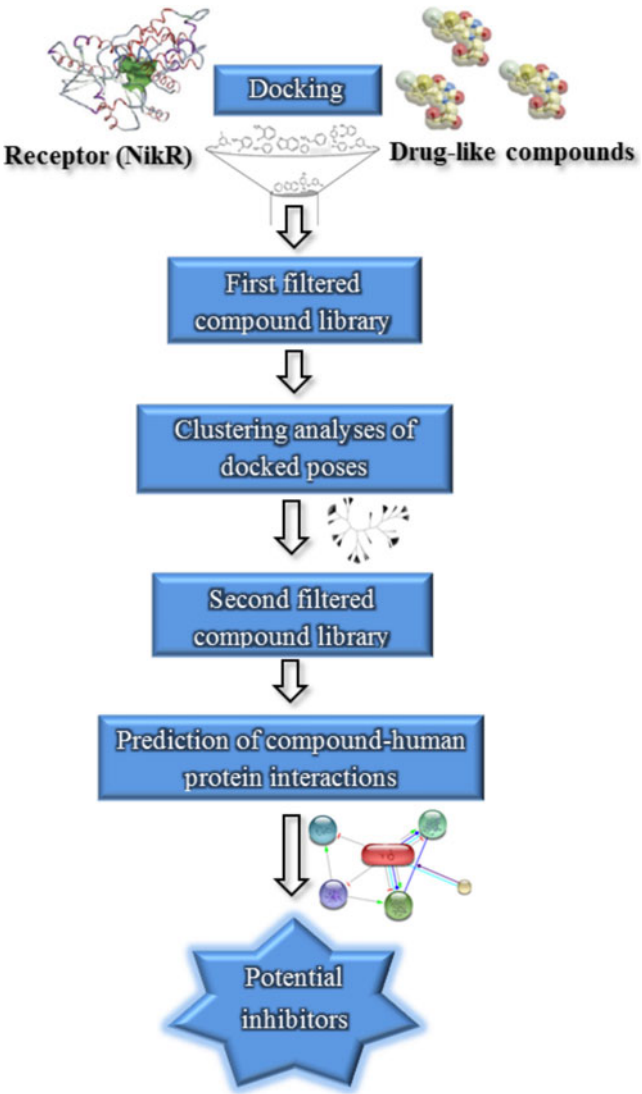


Fig. 1 Workflow diagram of the integrative computational protocol for the discovery of HpNikR inhibitors

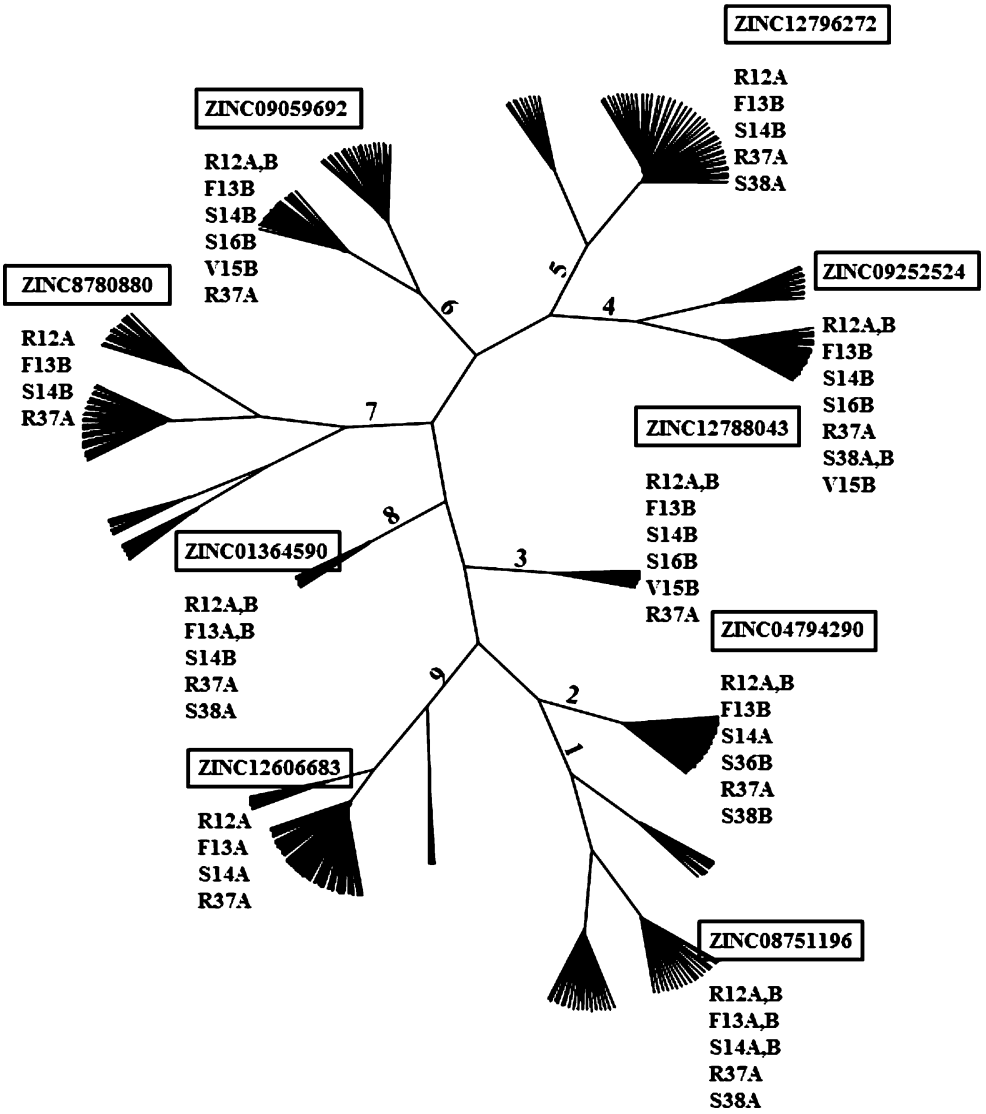


Fig. 2 Tree representation of contact footprints clustering for HpNikR structure-based virtual screening analysis. Numbers on the branches represent each of the nine clusters. Boxes indicate the representative compound in each cluster. The contact footprints per cluster are represented as follows: R12A,B means that Arg (R) in the position 12 of chain A and B interacts with the compounds of respective cluster, and so on

contains 29 compounds that interact with R12 of A and B chains, F13, S36, S38 predominantly of chain B, and the interactions with S14 are dominated by chain A. The cluster 3 contains 10 compounds that interact with R12 of A and B chains, R37 of chain A, and the interactions with F13, S14 and S16 are dominated by chain B. Interestingly, the cluster 9 corresponds to all compounds of the control group and 10 potential inhibitors (Fig. 2). In any case, the size of control group (see the section of Methods) did not affect the results of clustering analysis (data not shown). Hereinafter, we referred to the control group of 40 compounds (Fig. 2). The inspection of cluster 9 showed interactions with R12, F13, and S14 residues, but these contacts were dominated by chain A. It seems that the orientation towards chain A showed by compounds of cluster 9 could minimize the intermolecular forces in its interaction with the DNA-binding site of NikR and, therefore it affect its biding affinity. Moreover, the detailed inspection of interactions showed that, with exception of cluster 9, the most compounds of remaining eight clusters were able to mimic key interactions between amino acids and bases from protein-DNA complexes (it is also discussed further). Hence, the 10 potential inhibitors that clustered in cluster 9 were removed from the list of potential inhibitors. This result highlights the ability of CAR analysis to cluster the compounds according to their binding mode to HpNikR.

In addition, important structural information could be found in the analysis of atomic interactions. Contact clustering allows identification of key residues implicated in ligand binding [46]. These residues can be mapped onto the target structure. However, to better understand the mechanism of inhibition of potential inhibitors it is necessary to know the structural basis of HpNikR-DNA interaction. Further, the mechanisms behind the recognition of DNA molecules by HpNikR are as yet unknown. For this reason, prediction studies of putative residues involved

in HpNikR-DNA recognition were performed. The analysis with TFmodeller [49] predicted contacts in both chains for R12, S14, and S16; DISPLAR [50] detected contacts for I10, I11, R12, S14, S16, Q18, Y34, S35, S36, R37, S38, E39, R42, and D43. The analysis on crystal structure of the homologous complex (NikR-DNA) in *E. coli* [56] showed that the two programs presented congruence on the prediction of important residues in protein-DNA interaction; hence, the differences in the results might be due to the algorithms used by programs to perform predictions as indicated in the section Materials and Methods. The analyses showed that HpNikR DNA-contacting residues are a mixture of charged and polar residues along with a small number of hydrophobic residues which project outside the surface. The analysis of the contacts between the potential inhibitors and HpNikR showed that these compounds interact frequently with DNA-interacting residues such as I11, R12, S14, F13, S16, R37, and S38. Most interactions between these residues and the potential inhibitors were driven mainly by hydrophobic forces (Fig. 3). R12 and S14 could also play a role as strong hydrogen bond donors, as demonstrated elsewhere [56]. Schreiter et al. [56]., reported a complex structure between *E. coli* NikR and its operator DNA sequence suggesting that R3 and T5 (R12 and S14 in HpNikR) interacts with the nucleotide bases in the operator major groove. Schreiter et al. [56], also showed that residues equivalent to R37 and S38 in HpNikR frequently interact with the phosphate

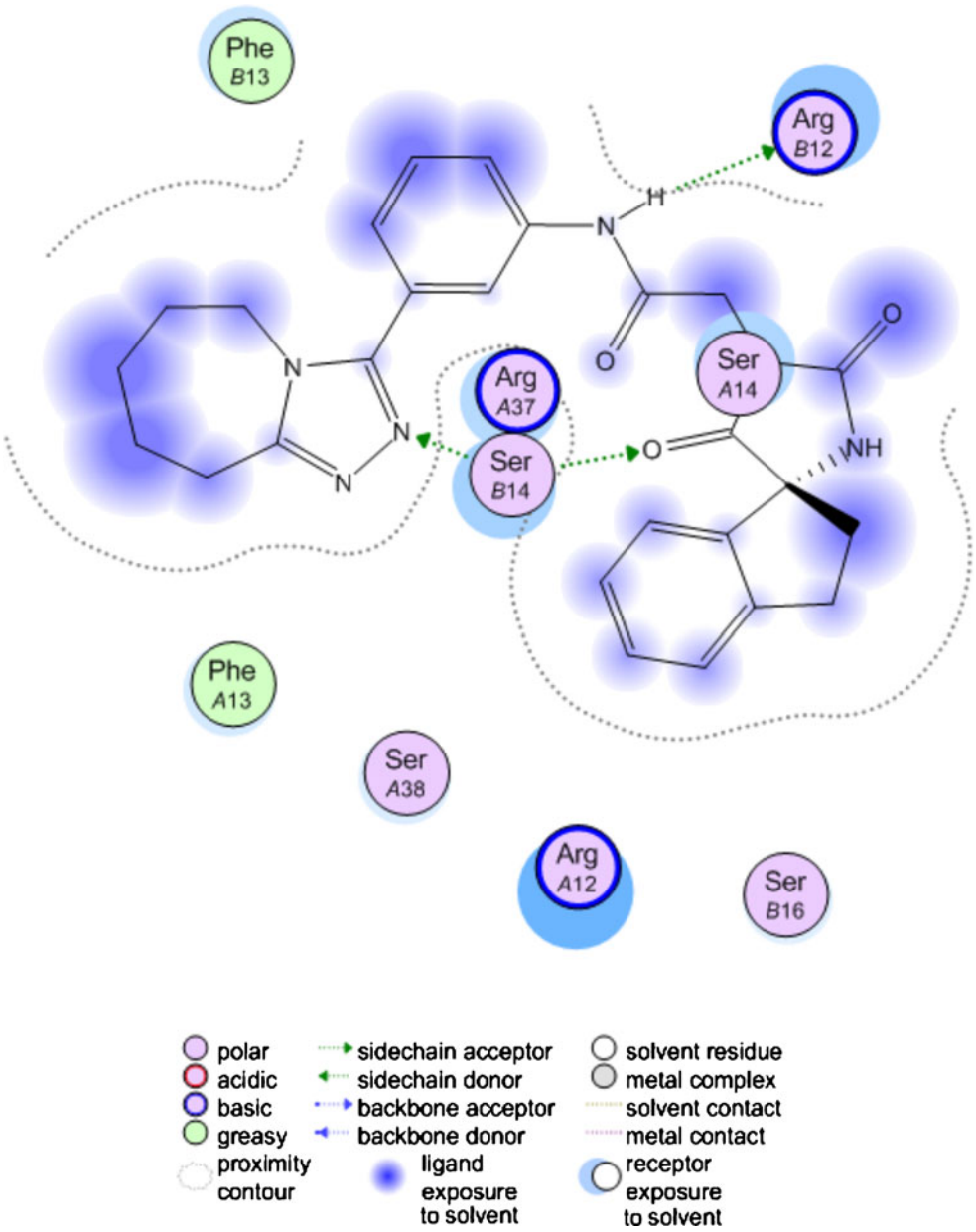


Fig. 3 The interaction between HpNikR residues and one top-scoring compound (ZINC09252524). The image was produced with the program Molecular Operating Environment (MOE) at www.chemcomp.com [59]

backbone of DNA. This kind of interactions is essential as they serve to anchor the domain to DNA and correctly orient the β-sheets within the major groove as shown elsewhere [56]. In addition, we have found that R12 and F13 residues frequently interact by pi-cation and pi-pi interaction, respectively, with at least one of the hetero-aromatic systems of potential inhibitors. Thus, R12 and F13 could be considered as a cationic and stacking center which stabilizes interaction on the face of the hetearomatic rings. It has been demonstrated [55–58] that these interactions are essential for the neurotransmitters-receptor and DNA-protein recognition process.

The physicochemical features of the potential inhibitors could explain why they preferentially interact with those residues, suggesting that they are able to mimic key interactions between nucleotide bases and DNA binding domain of HpNikR. A multiple structured-based alignment of NikR sequences from the HpNikR and other eleven pathogenic bacteria for humans (Fig. 4) revealed striking conservation at the residues that interact with the 266 potential inhibitors (e.g., the hydrophobic residues in position 13, and charged residues at positions 12, 14, 37, and 38); interestingly, some of these residues are involved in recognition and stabilizing of the DNA sequence [56]. Nevertheless, the HpNikR sequence has some unique features such as the replacement of an aliphatic side chain in position 13 by Phe. Overall, the identity between all NikR sequences ranged from 78% to 94%. The high conservation of the residues showed that the NikR of some

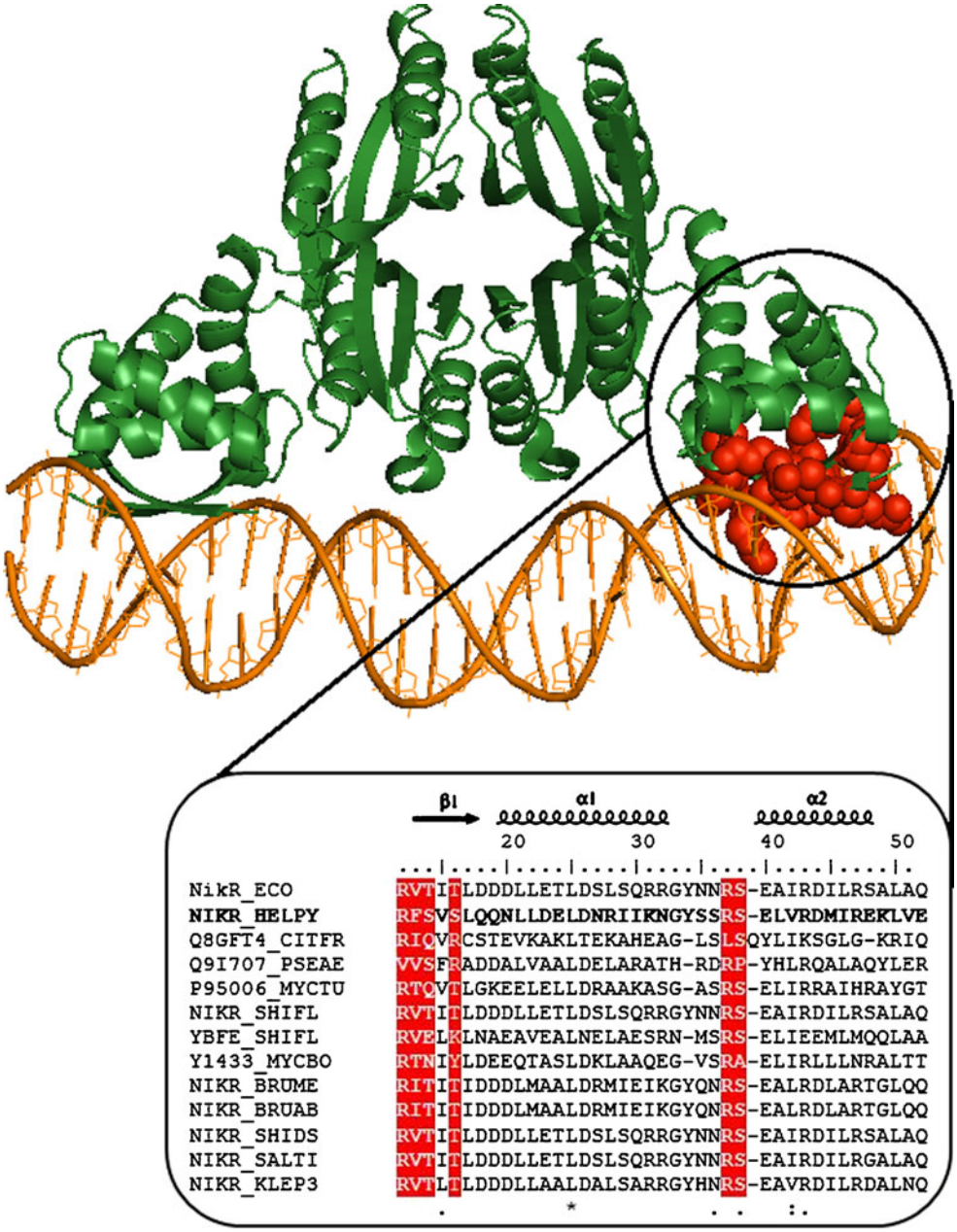


Fig. 4 Structure-based sequence alignment of NikR DNA-binding domain from *H. pylori* and other human pathogenic bacteria. Red boxes indicate residues that interact with the 250 potential inhibitors. HpNikR sequence is highlighted in bold, and the secondary structural elements are shown above the sequence alignment. The NikR-DNA complex corresponds to the structure reported by Schreiter et al. [56], PDB ID: 2HZV and red residues showed in the such structure corresponds to those in the red boxes. The multiple sequence alignment was constructed using Expresso 3D-Coffee [60]

pathogenic bacteria could preserve similar interactions with DNA molecules. Considering that NikR lacks of homologous proteins in humans, the potential inhibitors identified herein may also be useful in the rational development of novel drugs to fight the diseases in humans caused by other pathogenic bacteria of importance in public health such as *Salmonella typhy, Pseudomona aeruginosa* and *Shigella dysenteriae*.

Drug chemotherapy may usually contribute to the appearance of adverse effects after drug administration. Mostly, the complexity of this phenomenon is attributable to the interactions between the drug and unrelated target proteins [57]. Interaction among target proteins is not simple; conversely, the proteins interact within a complex interconnected network. Therefore, the study of the properties of drug-target interactions within the biological networks is of paramount significance for a better understanding of adverse effects associated with specific medications. Known and predicted interactions of the top-binding energy compounds with human proteins were explored using the STITCH server [51]. The STITCH analysis of the 266 top-binding energy compounds shows that 16 of these compounds have the potential to interact with human proteins. For example, the ZINC09025603 and ZINC09056854 compounds were structurally similar to the CRA-013783/L-006235 compound, which interacts with CtsB, CtsK and CtsS proteins. The CtsB, CtsK and CtsS proteins are involved in osteoclastic bone resorption and the removal of the invariant chain from MHC class II molecules, respectively. The ZINC09234386 compound similar to T-0632 compound, which interacts with Glp1R and CckaR proteins. The Glp1R protein activity is mediated by G proteins which activate adenylyl cyclase. The CckaR protein is involved in pancreatic growth and enzyme secretion, smooth muscle contraction of the gall bladder and stomach. Seven compounds were structurally similar to tadalafil and five of them were placed in cluster 2 suggesting that structural similarity obligates them to exhibit the same binding mode to HpNikR. Tadalafil is a cGMP-specific phosphodiesterase type 5 (PDE5) inhibitor, used for treating erectile dysfunction and it has recently been approved for the treatment of pulmonary arterial hypertension. Table 2 summarizes the functional partners in STITCH, and Gene Ontology [58] biological process annotations (GO terms) for human proteins. These 16 compounds were removed from the list of potential inhibitors because they were considered to be promiscuous, ending with an enrichment list of 250 specific compounds for HpNikR with minimal adverse drug side effects (Table 1, and Supplemental Material).

Table 2 The 16 promiscuous compounds removed from the list of potential inhibitors of HpNikR, because their interaction with human proteins

Compound	STITCH[51-52] match	Human functional partners	GO BP annotations[58]
ZINC09454661	GW9662: peroxisome proliferator-activated receptor antagonist	Pga1, RxrA, Tff2, PparA, PparG, RetN, Pga1 and Pprc1.	Fatty acid metabolism, lipid transport, regulation of cell proliferation and intracellular signaling cascades
ZINC09056854 ZINC09025603	CRA-013783/L-006235: cathepsin K inhibitor	CtsS, CtsK and CtsB	Proteolysis
ZINC11398614	6-[(6-fluoro-4-quinolyl)methyl]-5-[(3R)-3-hydroxypyrrolidine-1-carbonyl]-1-isobutyl-3-methyl-thieno[4,5-e]pyrimidine-2,4-dione	Slc16A1	Transmembrane transport
ZINC01060675	AS-252424: PI3K gamma inhibitor	Pik3cG	G-protein coupled receptor protein signaling pathway
ZINC02307440	STK169457	FasN	Fatty acid biosynthesis
ZINC03051369	Flurazepam: a member of the benzodiazepines and a long-acting depressor of the central nervous system with sedative and hypnotic effects	GabrA1, GabrB1, GabrA2, GabrA3, GabrB3, GabrA4, GabrA5, GabrA6, AbcB1 and GSR	Gamma-aminobutyric acid signaling pathway, Transmembrane transport and cell redox homeostasis
ZINC09234386	T-0632: cholecystokinin (CCK)A-receptor antagonist	CckaR and Glp1R	Cellular response to hormone stimulus and cAMP-mediating signaling
ZINC04689725	STOCK2S-39368	FasN	Fatty acid biosynthesis
ZINC02130463 ZINC02091539 ZINC01845382 ZINC02094309 ZINC02122438 ZINC09328404 ZINC02092396	Tadalafil: cGMP-specific phosphodiesterase type 5 inhibitor	Pde5A, AKT1 and Cyp3a4	Signal transduction, apoptosis and drug metabolic process

Conclusions

The implementation of an integrative protocol to identify potential inhibitors of HpNikR has been described. On the basis of ADME/TOX filtering, a library of half a million compounds was retrieved from the ZINC database. The resulting filtered library was docked into the DNA binding site of HpNikR to provide a list of 276 compounds ranked according to AutoDock Vina scoring. The 276 compounds were further screened by the analysis of contact clustering to discriminate active conformations and by the analysis of drug-protein interaction networks to remove the compounds that could lead to side effects on humans after drug administration. The resulting 250 compounds, which are mainly composed of three or four heteroaromatic systems and some rhodanine derivates, were finally identified as potential inhibitors to HpNikR. Visual examination of representative compounds of clusters and their interactions with the amino acids on the DNA binding sites of HpNikR were also determined. The experimental validations of novel compounds presented herein are currently underway.

Acknowledgments This project was supported by the Consejo Nacional de Ciencia y Tecnología-México (Grant Nos. 87935 and 105532) and by the Centro de Biotecnología Genómica of Instituto Politécnico Nacional-México (Secretaría de Investigación y Posgrado, Grant No. 20091288). The authors thank the Computational Systems Biology Group from Centro Nacional de Biotecnología of Consejo Superior de Investigaciones Científicas in Madrid, Spain for providing the access to the "Trueno" cluster, and programming. Aldo Segura was awarded a fellowship by the Agencia Española de Cooperacion Internacional para el Desarrollo (MAEC-AECID-2008), IIB program. Mario A. Rodríguez-Pérez and Xianwu Guo hold a scholarship from Comisión de Operación y Fomento de Actividades Académicas of Instituto Politécnico Nacional-México.

References

1. Loughlin MF (2003) Novel therapeutic targets in Helicobacter pylori. Expert Opin Ther Targets 7:725–735. doi:10.1517/14728222.7.6.725
2. Dunn BE, Cohen H, Blaser MJ (1997) Helicobacter pylori. Clin Microbiol Rev 10:720–741
3. Zschausch HC, Han SR, Meyer HG, Maeurer MJ (2002) No association between Helicobacter pylori genotypes and antibiotic resistance phenotypes within families. Helicobacter 7:364–366
4. Glupczynski Y, Broutet N, Cantagrel A, Andersen LP, Alarcon T, Lopez-Brea M, Megraud F (2002) Comparison of the E test and agar dilution method for antimicrobial suceptibility testing of Helicobacter pylori. Eur J Clin Microbiol Infect Dis 21:549–552. doi:10.1007/s10096-002-0757-6
5. Rimbara E, Noguchi N, Kawai T, Sasatsu M (2008) Novel mutation in 23 S rRNA that confers low-level resistance to clarithromycin in Helicobacter pylori. Antimicrob Agents Chemother 52:3465–3466. doi:10.1128/AAC.00445-08
6. Rimbara E, Noguchi N, Kawai T, Sasatsu M (2008) Mutations in penicillin-binding proteins 1, 2 and 3 are responsible for amoxicillin resistance in Helicobacter pylori. J Antimicrob Chemother 61:995–998. doi:10.1093/jac/dkn051
7. Rimbara E, Noguchi N, Tanabe M, Kawai T, Matsumoto Y, Sasatsu M (2005) Susceptibilities to clarithromycin, amoxycillin and metronidazole of Helicobacter pylori isolates from the antrum and corpus in Tokyo, Japan, 1995-2001. Clin Microbiol Infect 11:307–311. doi:10.1111/j.1469-0691.2005.01099.x
8. Rimbara E, Noguchi N, Yamaguchi T, Narui K, Kawai T, Sasatsu M (2005) Development of a highly sensitive method for detection of clarithromycin-resistant Helicobacter pylori from human feces. Curr Microbiol 51:1–5. doi:10.1007/s00284-004-4488-z
9. Samra Z, Shmuely H, Niv Y, Dinari G, Passaro DJ, Geler A, Gal E, Fishman M, Bachor J, Yahav J (2002) Resistance of Helicobacter pylori isolated in Israel to metronidazole, clarithromycin, tetracycline, amoxicillin and cefixime. J Antimicrob Chemother 49:1023–1026
10. Hu CT, Wu CC, Lin CY, Cheng CC, Su SC, Tseng YH, Lin NT (2007) Resistance rate to antibiotics of Helicobacter pylori isolates in eastern Taiwan. J Gastroenterol Hepatol 22:720–723. doi:10.1111/j.1440-1746.2006.04743.x
11. Boyanova L, Stancheva I, Spassova Z, Katzarov N, Mitov I, Koumanova R (2000) Primary and combined resistance to four antimicrobial agents in Helicobacter pylori in Sofia, Bulgaria. J Med Microbiol 49:415–418
12. Wolle K, Leodolter A, Malfertheiner P, Konig W (2002) Antibiotic susceptibility of Helicobacter pylori in Germany: stable primary resistance from 1995 to 2000. J Med Microbiol 51:705–709
13. Mishra KK, Srivastava S, Garg A, Ayyagari A (2006) Antibiotic susceptibility of Helicobacter pylori clinical isolates: comparative evaluation of disk-diffusion and E-test methods. Curr Microbiol 53:329–334. doi:10.1007/s00284-006-0143-1
14. Dore MP, Osato MS, Kwon DH, Graham DY, el-Zaatari FA (1998) Demonstration of unexpected antibiotic resistance of genotypically identical Helicobacter pylori isolates. Clin Infect Dis 27:84–89
15. Whitmire JM, Gancz H, Merrell DS (2007) Balancing the double-edged sword: metal ion homeostasis and the ulcer bug. Curr Med Chem 14:469–478
16. Bik EM, Eckburg PB, Gill SR, Nelson KE, Purdom EA, Francois F, Perez-Perez G, Blaser MJ, Relman DA (2006) Molecular analysis of the bacterial microbiota in the human stomach. Proc Natl Acad Sci USA 103:732–737. doi:10.1073/pnas.0506655103
17. Van Vliet AH, Kuipers EJ, Stoof J, Poppelaars SW, Kusters JG (2004) Acid-responsive gene induction of ammonia-producing enzymes in Helicobacter pylori is mediated via a metal-responsive repressor cascade. Infect Immun 72:766–773
18. Mobley HL, Island MD, Hausinger RP (1995) Molecular biology of microbial ureases. Microbiol Rev 59:451–480
19. Mulrooney SB, Hausinger RP (2003) Nickel uptake and utilization by microorganisms. FEMS Microbiol Rev 27:239–261
20. Stingl K, De Reuse H (2005) Staying alive overdosed: how does Helicobacter pylori control urease activity? Int J Med Microbiol 295:307–315
21. Krishnaswamy R, Wilson DB (2000) Construction and characterization of an Escherichia coli strain genetically engineered for Ni (II) bioaccumulation. Appl Environ Microbiol 66:5383–5386
22. Contreras M, Thiberge JM, Mandrand-Berthelot MA, Labigne A (2003) Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of Helicobacter pylori. Mol Microbiol 49:947–963
23. Dosanjh NS, Michel SL (2006) Microbial nickel metalloregulation: NikRs for nickel ions. Curr Opin Chem Biol 10:123–130. doi:10.1016/j.cbpa.2006.02.011
24. Van Vliet AH, Ernst FD, Kusters JG (2004) NikR-mediated regulation of Helicobacter pylori acid adaptation. Trends Microbiol 12:489–494. doi:10.1016/j.tim.2004.09.005
25. Wolfram L, Haas E, Bauerfeind P (2006) Nickel represses the synthesis of the nickel permease NixA of Helicobacter pylori. J

Bacteriol 188:1245–1250. doi:10.1128/JB.188.4.1245-1250.2006
26. Chivers PT, Sauer RT (1999) NikR is a ribbon-helix-helix DNA-binding protein. Protein Sci 8:2494–2500. doi:10.1110/ps.8.11.2494
27. Schreiter ER, Sintchak MD, Guo Y, Chivers PT, Sauer RT, Drennan CL (2003) Crystal structure of the nickel-responsive transcription factor NikR. Nat Struct Biol 10:794–799. doi:10.1038/nsb985
28. Chivers PT, Sauer RT (2002) NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. Chem Biol 9:1141–1148
29. Wang SC, Dias AV, Bloom SL, Zamble DB (2004) Selectivity of metal binding and metal-induced stability of Escherichia coli NikR. Biochemistry 43:10018–10028. doi:10.1021/bi049405c
30. Delany I, Ieva R, Soragni A, Hilleringmann M, Rappuoli R, Scarlato V (2005) In vitro analysis of protein-operator interactions of the NikR and fur metal-responsive regulators of coregulated genes in Helicobacter pylori. J Bacteriol 187:7703–7715. doi:10.1128/JB.187.22.7703-7715.2005
31. Abraham LO, Li Y, Zamble DB (2006) The metal- and DNA-binding activities of Helicobacter pylori NikR. J Inorg Biochem 100:1005–1014. doi:10.1016/j.jinorgbio.2005.10.014
32. Ernst FD, Homuth G, Stoof J, Mader U, Waidner B, Kuipers EJ, Kist M, Kusters JG, Bereswill S, van Vliet AH (2005) Iron-responsive regulation of the Helicobacter pylori iron-cofactored superoxide dismutase SodB is mediated by Fur. J Bacteriol 187:3687–3692. doi:10.1128/JB.187.11.3687-3692.2005
33. Li Y, Zamble DB (2009) pH-responsive DNA-binding activity of Helicobacter pylori NikR. Biochemistry 48:2486–2496. doi:10.1021/bi801742r
34. Sperandio O, Miteva MA, Segers K, Nicolaes GA, Villoutreix BO (2008) Screening Outside the Catalytic Site: Inhibition of Macromolecular Inter-actions Through Structure-Based Virtual Ligand Screening Experiments. Open Biochem J 2:29–37. doi:10.2174/1874091X00802010029
35. Ivanov AS, Gnedenko OV, Molnar AA, Mezentsev YV, Lisitsa AV, Archakov AI (2007) Protein-protein interactions as new targets for drug design: virtual and experimental approaches. J Bioinform Comput Biol 5:579–592
36. Bowser TE, Bartlett VJ, Grier MC, Verma AK, Warchol T, Levy SB, Alekshun MN (2007) Novel anti-infection agents: small-molecule inhibitors of bacterial transcription factors. Bioorg Med Chem Lett 17:5652–5655. doi:10.1016/j.bmcl.2007.07.072
37. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461. doi:10.1002/jcc.21334
38. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 58:899–907
39. Dian C, Schauer K, Kapp U, McSweeney SM, Labigne A, Terradot L (2006) Structural basis of the nickel response in Helicobacter pylori: crystal structures of HpNikR in Apo and nickel-bound states. J Mol Biol 361:715–730. doi:10.1016/j.jmb.2006.06.058
40. Sanner MF (1999) Python: a programming language for software integration and development. J Mol Graph Model 17:57–61
41. Sanner MF, Duncan BS, Carrillo CJ, Olson AJ (1999) Integrating computation and visualization for biomolecular analysis: an example using python and AVS. Pac Symp Biocomput:401-412
42. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30:2785–2791. doi:10.1002/jcc.21256
43. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26
44. Irwin JJ, Shoichet BK (2005) ZINC–a free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182. doi:10.1021/ci049714+
45. Chang MW, Belew RK, Carroll KS, Olson AJ, Goodsell DS (2008) Empirical entropic contributions in computational docking: evaluation in APS reductase complexes. J Comput Chem 29:1753–1761. doi:10.1002/jcc.20936
46. Bouvier G, Evrard-Todeschi N, Girault JP, Bertho G (2010) Automatic clustering of docking poses in virtual screening process using self-organizing map. Bioinformatics 26:53–60. doi:10.1093/bioinformatics/btp623
47. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221–224. doi:10.1093/molbev/msp259
48. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. doi:10.1002/jcc.20084
49. Contreras-Moreira B, Branger PA, Collado-Vides J (2007) TFmodeller: comparative modelling of protein-DNA complexes. Bioinformatics 23:1694–1696. doi:10.1093/bioinformatics/btm148
50. Tjong H, Zhou HX (2007) DISPLAR: an accurate method for predicting DNA-binding sites on protein surfaces. Nucleic Acids Res 35:1465–1477. doi:10.1093/nar/gkm008
51. Kuhn M, Szklarczyk D, Franceschini A, Campillos M, von Mering C, Jensen LJ, Beyer A, Bork P (2010) STITCH 2: an interaction network database for small molecules and proteins. Nucleic Acids Res 38 (Database issue):D552-556. doi:10.1093/nar/gkp937
52. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P (2008) STITCH: interaction networks of chemicals and proteins. Nucleic Acids Res 36 (Database issue):D684-688. doi:10.1093/nar/gkm795
53. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH (2010) The NCBI BioSystems database. Nucleic Acids Res 38 (Database issue):D492-496. doi:10.1093/nar/gkp858
54. Guha R, Howard MT, Hutchison GR, Murray-Rust P, Rzepa H, Steinbeck C, Wegner J, Willighagen EL (2006) The Blue Obelisk-interoperability in chemical informatics. J Chem Inf Model 46:991–998. doi:10.1021/ci050400b
55. Waszkowycz B (2008) Towards improving compound selection in structure-based virtual screening. Drug Discov Today 13:219–226. doi:10.1016/j.drudis.2007.12.002
56. Schreiter ER, Wang SC, Zamble DB, Drennan CL (2006) NikR-operator complex structure and the mechanism of repressor activation by metal ions. Proc Natl Acad Sci USA 103:13676–13681. doi:10.1073/pnas.0606247103
57. Liebler DC, Guengerich FP (2005) Elucidating mechanisms of drug-induced toxicity. Nat Rev Drug Discov 4:410–420. doi:10.1038/nrd1720
58. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29. doi:10.1038/75556
59. Clark AM, Labute P (2007) 2D depiction of protein-ligand complexes. J Chem Inf Model 47:1933–1944. doi:10.1021/ci7001473
60. Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V, Notredame C (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res 34 (Web Server issue):W604-608. doi:10.1093/nar/gkl092