Bioorganic & Medicinal Chemistry Letters 23 (2013) 6286–6291

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Design checkpoint kinase 2 inhibitors by pharmacophore modeling and virtual screening techniques

Yen-Ling Wang [a,†], Chun-Yuan Lin [a,*,†], Kuei-Chung Shih [b], Jui-Wen Huang [c], Chuan-Yi Tang [d]

[a] Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan 33302, Taiwan
[b] Department of Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan
[c] Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu 30011, Taiwan
[d] Department of Computer Science and Information Engineering, Providence University, Taichung 43301, Taiwan

ARTICLE INFO

Article history:
Received 9 April 2013
Revised 26 August 2013
Accepted 25 September 2013
Available online 3 October 2013

Keywords:
Chk2
Drug design
Pharmacophore
Virtual screening

ABSTRACT

Damage to DNA is caused by ionizing radiation, genotoxic chemicals or collapsed replication forks. When DNA is damaged or cells fail to respond, a mutation that is associated with breast or ovarian cancer may occur. Mammalian cells control and stabilize the genome using a cell cycle checkpoint to prevent damage to DNA or to repair damaged DNA. Checkpoint kinase 2 (Chk2) is one of the important kinases, which strongly affects DNA-damage and plays an important role in the response to the breakage of DNA double-strands and related lesions. Therefore, this study concerns Chk2. Its purpose is to find potential inhibitors using the pharmacophore hypotheses (PhModels) and virtual screening techniques. PhModels can identify inhibitors with high biological activities and virtual screening techniques are used to screen the database of the National Cancer Institute (NCI) to retrieve compounds that exhibit all of the pharmacophoric features of potential inhibitors with high interaction energy. Ten PhModels were generated using the HypoGen best algorithm. The established PhModel, Hypo01, was evaluated by performing a cost function analysis of its correlation coefficient (r), root mean square deviation (RMSD), cost difference, and configuration cost, with the values 0.955, 1.28, 192.51, and 16.07, respectively. The result of Fischer's cross-validation test for the Hypo01 model yielded a 95% confidence level, and the correlation coefficient of the testing set (r_{test}) had a best value of 0.81. The potential inhibitors were then chosen from the NCI database by Hypo01 model screening and molecular docking using the CDOCKER docking program. Finally, the selected compounds exhibited the identified pharmacophoric features and had a high interaction energy between the ligand and the receptor. Eighty-three potential inhibitors for Chk2 are retrieved for further study.

Damage to DNA is caused by ionizing radiation, genotoxic chemicals or collapsed replication forks. When DNA is damaged or cells fail to respond to, the mutation associated with breast or ovarian cancer may occur. To prevent damage to DNA and to repair damaged DNA, mammalian cells control and stabilize the genome using a cell cycle checkpoint. The checkpoint pathway involves many kinases, including ataxia telangiectasia mutated protein (ATM),[1,2] ataxia telangiectasia and Rad3-related protein (ATR),[1,2] checkpoint kinase 1 (Chk1),[3,4] and Checkpoint kinase 2 (Chk2).[5–8]

ATM and ATR are upstream kinases that pass messages to downstream kinases and phosphorylate many proteins that initiate the activation of a DNA-damage checkpoint. ATM is a primary participant in the pathway of the activation of p53 (protein 53)[9] by Chk2, and ATR may affect the phosphorylation of Chk1. Both Chk1 and Chk2 are critical in the damaging of DNA; however, their cellular activities differ. Chk1 participates in S and G2 phases of the cell cycle via an ATR pathway. In contrast, Chk2 is activated in all phases throughout an ATM-dependent pathway and has an important role in response to DNA double-strand breakage and related lesions. Chk1 is an unstable protein and lacks the forkhead-associated domain (FHA), which is involved in many processes that protect against cancer and are present in Chk2. Therefore, this study concentrates on Chk2.

Chk2 is a protein that contains 543 amino acid residues and the structure of Chk2 comprises some functional elements, including the SQ/TQ cluster domain (SCD), FHA, and the serine/threonine kinase domain (KD).[5–8] The SCD is known to be a preferred site at which is located the residue Thr68 for phosphorylation in response to damage to DNA by ATM/ATP kinases. The FHA domain is a phosphopeptide recognition domain that is found in many regulatory proteins and thought to bind to the phosphoThr68 segment of SCD.[5–8,10–14] Therefore, it is a good candidate in the interactions of Chk2 with its upstream regulators or downstream targets in cell-cycle-checkpoint signaling. The KD occupies almost the whole

* Corresponding author. Tel.: +886 32118800; fax: +886 32118700.
E-mail address: cyulin@mail.cgu.edu.tw (C.-Y. Lin).
† These authors contributed equally to this work.

0960-894X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.09.080

carboxy-terminal half of Chk2, and has been identified by their homology with serine/threonine kinases.

Some studies have reported that when DNA is damaged, Chk2 is activated by ATM/ATR via the phosphorylation of residue Thr68. Additionally, Chk2 induces the trans-autophosphorylation of residues Thr383 and Thr387, and then the *cis*-phosphorylation of residue Ser516.[5–8,10–14] Thereafter, Chk2 will phosphorylate many downstream substrates, including BRCA1 (breast cancer 1, early onset),[15,16] Cdc25A (cell division cycle 25 homolog A), Cdc25C, and p53.[7,8,10] Many works have shown that Chk2 phosphorylates Cdc25A, which is considered to be an oncogene on residue Ser123 in the S phase of the cell cycle, and it also phosphorylates Cdc25C on residue Ser216 in the G2 phase, helping to prevent mitotic entry into cells with damaged DNA.[5] Furthermore, BRCA1 and p53 participate in DNA repair in breast or ovarian cancer. BRCA1 is a human caretaker gene that helps to repair damaged DNA or destroys cells that cannot be repaired. p53 is a tumor suppressor protein that is involved in the prevention of cancer in humans, and has an important role in the G1 checkpoint in response to DNA-damaging agents. Sites of phosphorylations are important in the design of drugs to promote cell survival when DNA is damaged.

Recently, several studies have identified a few inhibitors of Chk2,[6–8,10–14] including 2-arylbenzimidazole, VRX0466617, isothiazole carboxamides, PV1019 and NSC 109555. They have also shown the crystal structures of Chk2 complex, such as PDB: 1GXC, 2W7X, and others. These inhibitors are selective, reversible, and ATP-competitive. They effectively inhibit the radiation-induced phosphorylation of Chk2. The quantitative structure-activity relationship model (QSAR model) is a regression or classification model, which is important in rational drug design. It is used to relate the structural properties of compounds to their biological activities. The use of QSAR to predict the quality was improved using the three-dimensional QSAR (3D-QSAR)[19–24] of the targeted inhibitor, allowing the structure to be directly optimized in 3D space. Comparative molecular field analyses (CoMFA)[18,25–30] and comparative molecular similarity indices (CoMSIA)[18,27–32] for Chk2 inhibitors were performed using ligand-based and receptor-guided alignment.[18] These methods use the co-crystal structure from a protein data bank (PDB code: 2CN8)[7] with 25 Chk2 inhibitors, and then they identify new plausible binding modes for use as templates in 3D-QSAR.[18] Other research on Chk2 in QSAR/QSPR[17] has elucidated structures that reduce the side effects of Chk2 inhibitors.

Pharmacophore[20–23,33–36] refers to a set of structural features of a molecule that are responsible for its biological activity. Compounds with diverse structures and common chemical features can be found by ligand pharmacophore mapping, which differs from CoMFA and CoMSIA in having the common substructure constraint. Therefore, a pharmacophore can be used to explain how diverse ligands bind to a receptor site, and to visualize potential chemical interactions between ligands and receptors. Additionally, a pharmacophore can be used easily and quickly to identify candidate inhibitors of a target protein based on a 3D query. Therefore, in this work, 3D-QSAR was used as in a previous work[37] to establish hypotheses about pharmacophores (PhModels) for Chk2 inhibitors. Virtual screening is a computational method that is used in drug discovery. The two categories of screening techniques are structure-based and ligand-based. In this work, in ligand-based virtual screening, the PhModel was used in 3D structure queries with model screening such that each compound in National Cancer Institute (NCI) database is mapped onto the pharmacophoric features. Each compound that has all of the chemical features of the PhModel was selected. Finally, the potential inhibitors were retrieved from selected compounds by using a molecular docking program to predict the conformation and interaction energy between Chk2 and the ligands. Unlike in earlier work,[37] Lipinski's Rule-of-Five filtration was not applied to selected compounds, to prevent the removal of potential inhibitors. We believe that PhModels with information about potential chemical interactions can help medicinal chemists to identify or design new Chk2 inhibitors. The potential inhibitors of Chk2 that were retrieved in this work should be further examined by biologists.

To construct PhModels, Chk2 inhibitors with two-dimensional structures and their biological activity values were collected from the ChEMBL database.[36] From structural variations and chemical variation in the kinase inhibitor activity, 158 known Chk2 inhibitors were selected and retrieved. The biological activity of each of 158 known Chk2 inhibitors was specified as IC_{50} (nanomolar, nM). 260,071 compounds from the NCI database (release version 3, http://cactus.nci.nih.gov/download/nci/) were used for database screening and in molecular docking.

Before PhModels were generated, the 158 Chk2 inhibitors were divided into the training and testing sets. The rules that were utilized to select training set inhibitors satisfy the following requirements, as suggested by the Accelrys Discovery Studio. (1) Clear and concise information must be available on all selected inhibitors, including structural features and range of activity; (2) at least 16 diverse inhibitors must be selected as the training set to ensure statistical significance; (3) the training set should contain the most and the least active inhibitors; (4) the biological activities of the inhibitors span at least four orders of magnitude. Based on the above four rules, the 158 Chk2 inhibitors were divided into training and testing sets. Figure 1 presents the scatter diagram of the training and testing sets of inhibitors. Figure 1 presents the distribution of the inhibitors in the training and testing sets, and the representative points of the testing set are close to those of the training set. The training set, which has 25 inhibitors, is then used to construct PhModels, and the IC_{50} values of these 25 inhibitors ranged from 2.3 to 100,000 nM. Supplementary Figure S1 presents the chemical structures of the training set inhibitors; Supplementary Table S1 presents the experimental and estimated IC_{50}, and the energies of interaction of the training set inhibitors. The testing set the remaining 133 inhibitors whose chemical structures and IC_{50} values are shown in Supplementary Table S2 is used to test the predictive ability of the generated PhModels. The IC_{50} values of the 133 testing set inhibitors ranged from 3.4 to 74,000 nM. Therefore, in this study, the generated PhModels are used to identify diverse candidate inhibitors of Chk2.

After the training set inhibitors were selected, ten PhModels were established. Then, cost function analysis, correlation analysis, and Fischer's cross-validation analysis, were used to estimate the predictive capacities of the ten PhModels. In this study, the

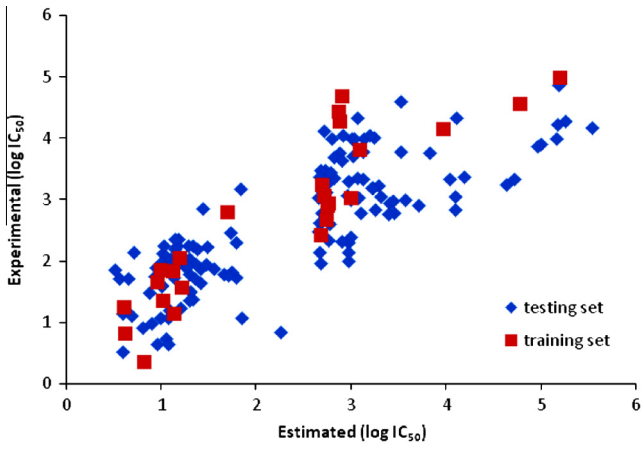

Figure 1. Scatter diagram of training set and testing set of inhibitors.