Full Paper

www.molinf.com

molecular informatics

Special Issue EUROQSAR

DOI: 10.1002/minf.201400140

Pharmacophore-Based Virtual Screening to Discover New Active Compounds for Human Choline Kinase α1

Lucía Serrán-Aguilera,[a] Roberto Nuti,[b] Luisa C. López-Cara,[a] Miguel Á. Gallo Mezo,[a] Antonio Macchiarulo,*[b] Antonio Entrena,*[a] and Ramón Hurtado-Guerrero*[c]

Abstract: Choline kinase (CK) catalyses the transfer of the ATP γ-phosphate to choline to generate phosphocholine and ADP in the presence of magnesium leading to the synthesis of phosphatidylcholine. Of the three isoforms of CK described in humans, only the α isoforms (*Hs*CKα) are strongly associated with cancer and have been validated as drug targets to treat this disease. Over the years, a large number of Hemicholinium-3 (HC-3)-based *Hs*CKα biscationic inhibitors have been developed though the relevant common features important for the biological function have not been defined. Here, selecting a large number of previous HC-3-based inhibitors, we discover through computational studies a pharmacophore model formed by five moieties that are included in the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment. Using a pharmacophore-guided virtual screening, we then identified 6 molecules that showed binding affinities in the low μM range to *Hs*CKα1. Finally, protein crystallization studies suggested that one of these molecules is bound to the choline and ATP-binding sites. In conclusion, we have developed a pharmacophore model that not only allowed us to dissect the structural important features of the previous HC-3 derivatives, but also enabled the identification of novel chemical tools with good ligand efficiencies to investigate the biological functions of *Hs*CKα1.

Keywords: Choline kinase · Pharmacophore model · Virtual screening · Tryptophan fluorescence · Hits · Binding · Crystallization

1 Introduction

Choline kinase (CK) is a homodimeric enzyme that catalyses the transfer of the γ-phosphate group from ATP to choline rendering ADP and phosphocholine (PCho), and it is the first committed step in the Kennedy pathway leading to the synthesis of phosphatidylcholine (PtdCho).[1] PtdCho is a crucial phospholipid in membrane neogenesis for successful growth and proliferation in human cells, and in other eukaryotic and bacterial pathogens.[2a–c] In humans, there are three isoforms of CK being the α-isoforms (*Hs*CKα1 and *Hs*CKα2) validated drug targets in cancer disease.[3]

Initially the first inhibitor for *Hs*CKα1 was Hemicholinium-3 (HC-3, **1**),[4] though this compound showed multiple life-threatening side effects.[5] To ameliorate the toxicity and keep key *Hs*CKα-binding features of HC-3, a large number of generally large symmetrical or asymmetrical biscationic compounds based on HC-3 structure were synthesized along the last years (Figure 1).[6a–e] One of these molecules, TCD-717 (**2**), has even passed Phase I under clinical trials for advanced solid tumors.[7] Very recently the first crystal structures of several complexes of *Hs*CKα1 bound to some of these inhibitors have enlightened their binding mode majorly through interactions to the choline-binding site.[8a–c] Inhibitors have been also developed using not only a rational approach but also virtual and high-throughput screening of small-molecule libraries.[9]

Despite the extensive information on CK inhibitors, there is not a study dissecting the key chemical features shared among the large number of inhibitors that are important for enzyme recognition. Herein, we propose a pharmaco-

[a] *L. Serrán-Aguilera,[+] L. C. López-Cara,[+] M. Á. G. Mezo,[+] A. Entrena[+]*
Department of Pharmaceutical and Organic Chemistry, Faculty of Pharmacy, University of Granada
Campus Cartuja, Granada 18071, Spain
phone: +34 958 243848
**e-mail: aentrena@ugr.es*

[b] *R. Nuti,[+] A. Macchiarulo[+]*
Dipartimento di Scienze Farmaceutiche, Università degli Studi di Perugia
Via del Liceo, 1, 06123 Perugia, Italy
**e-mail: antonio@chimfarm.unipg.it*

[c] *R. Hurtado-Guerrero[+]*
Institute of Biocomputation and Physics of Complex Systems (BIFI) and BIFI-IQFR (CSIC) Joint Unit, University of Zaragoza
Campus Río Ebro, Zaragoza 50018, Spain; Edificio I+D;
Fundación ARAID, Edificio Pignatelli 36, Spain
phones: +39 075 5855160; +34 976 762997
**e-mail: rhurtado@bifi.es*

[+] *The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/minf.201400140.

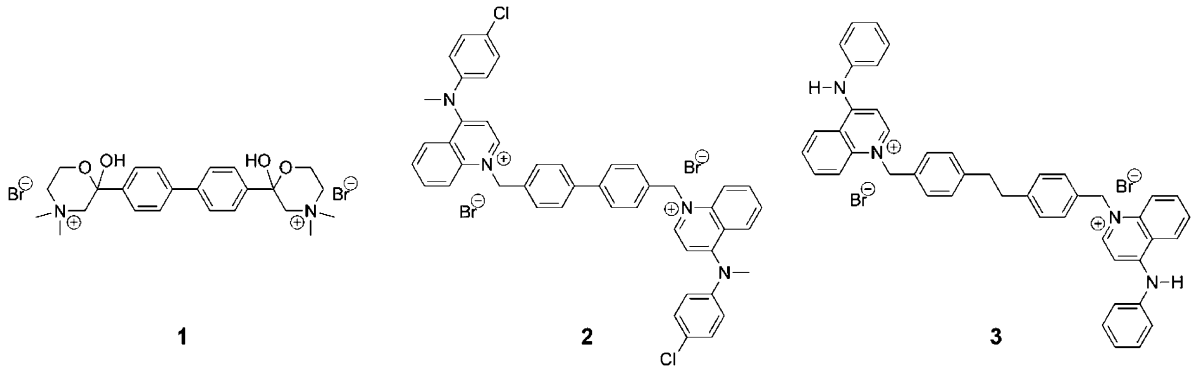


Figure 1. Chemical structure of compounds HC-3 (**1**), TCD-717 (**2**) and CHEMBL-364022 (**3**).

phore model formed by 5 structural features, and validate it by the discovery of new simpler molecules against *Hs*CKα1 using a virtual screening approach. Not only we discover efficient molecules that bind to *Hs*CKα1 in the low μM range but we also describe the binding mode for one of these molecules. Therefore, this thorough study serves to understand the common moieties of hundreds of compounds essential for CK recognition and in turn to design and discover new efficient and simpler molecules that might have potential applications as lead molecules to develop new anticancer agents.

2 Methodology

2.1 Pharmacophore Model Design

2.1.1 First Docking Study

The initial docking study of 83 inhibitors chemically derived from HC-3 (Figure S1) into *Hs*CKα1 was developed after protein and ligands preparation using different utilities included into Maestro 9.3.[10] The crystal structure of *Hs*CKα1 complexed with HC-3 was downloaded from the Protein Data Bank (PDB ID: 3G15). After waters and metal ions deletion, chain A was submitted to Protein Preparation Wizard[12] in order to add hydrogens, assign correct bond orders, convert selenomethionines into methionines and solve sterical conflicts. Hydrogens were minimized at physiological pH, whereas heavy atoms minimization was carried out with Impact[12b] Refinement module using OPLS-2005 force field and terminated when root-mean-square deviation (RMSD) reached a maximum cutoff of 0.3 Å. ADP and HC-3 states at physiological pH were also generated with Protein Preparation Wizard. A formal charge of -2 for ADP and of $+2$ for HC-3 was selected. The set of 83 HC-3 derivatives was downloaded from the Binding Database[13] in SD format and pre-processed with LigPrep 2.5 in order to add hydrogens, remove counterions, generate ionization states at physiological pH with Epik,[12a] generate low-energy ring conformations and optimize their geometries by OPLS-2005. In this set of compounds there were active and inactive molecules with a value of IC_{50} ranged from 0.4 μM to higher than 100 μM. SP flexible docking study was performed with Glide[14] 5.8 based on a receptor grid of 10 Å length centered on HC-3.

2.1.2 HC-3 Pharmacophoric Hypothesis (Hypothesis-1).

An XP precision refinement of HC-3 pose into the PDB entry 3G15 was developed with Glide 5.8 before generating a pharmacophoric hypothesis for HC-3 with E-pharmacophore.[15] This hypothesis (hypothesis-1) included 6 features that were used as input by Phase[16] 3.4 together with 5 Å of excluded volume and a feature-matching tolerance criteria defined by at least 3 out of 6 pharmacophoric features located at a maximum distance of 2.0 Å.

Specifically, the excluded volume is the space occupied by residues of the binding pocket lying in a sphere of 5 Å from HC-3 that cannot be occupied by any moiety of a potential hit compound due to the formation of steric bumps. The feature-matching tolerance is the minimum number and relative distances of pharmacophoric features that a compound should fit to be considered a potential hit molecule.

The training set was composed of 83 preprocessed compounds and the excluded volume was centred on the HC-3 of the PDB entry 3G15 (Figure S2 Supporting Information). A number of 100 conformers was generated per rotatable bond with a maximum of 1000 conformers per structure using ConfGen.[17] Training set compounds were firstly ranked according to the number of matched pharmacophoric features (descending order from 6 to 3), and then to the Fitness score. Specifically, the Fitness score is a parameter measuring the goodness of the fit of the compound on the pharmacophoric features, with higher values meaning better fitting of the pharmacophore model.

2.1.3 1-Benzyl-4-(N-methylaniline)pyridinium Pharmacophore Hypothesis (Hypothesis 2)

This alternative pharmacophore searching was based on the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment. This

fragment was optimized and energy minimized with LigPrep 2.5. Then, it was rigidly docked into the PDB entry 3G15, using a grid of 10 Å centered on HC-3 and the Standard Precision protocol (SP). The binding pose where its benzylic ring was overlaid with one of the phenyls of HC-3 (Figure S3 Supporting Information) was submitted to an XP refinement before the search of the pharmacophoric features with E-pharmacophore. Manual definition of the pharmacophoric features was made including the positive charge, the amino nitrogen or the hydrophobic moiety (bound to this amino nitrogen). The pharmacophoric features defined by E-pharmacophore were used as input in Phase 3.4 calculations together with an excluded volume of 5 Å from the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment, as defined previously. A pool of twelve pharmacophoric hypotheses was thus generated. The feature-matching tolerance was defined by 3 out of the 6 pharmacophoric features located at a maximum distance of 2.0 Å. In this case, 11 active molecules ($IC_{50} < 1.5$ μM) and 11 inactive compounds ($IC_{50} > 100$ μM) of the initial 83 molecules were selected to define the test set (Table S1 Supporting Information) and their conformer generation was made as previously described. In particular, this reduction of the dataset from 83 molecules to 22 compounds was done in order to obtain a clear and balanced separation between active and inactive compounds. The best hypothesis (hypothesis-2) among the twelve ones was selected according to its ability to select not only the minimum number of inactive compounds but also to rank at first positions the active ligands and to show a correlated trend between the IC_{50} values and the Fitness scores of ligands (Table S2 Supporting Information).

2.2 Pharmacophore Based Virtual Screening

All the compounds included into Enamine,[18] Chembridge[19] and Life Chemicals[20] libraries were preprocessed with LigPrep 2.5 and ConfGen in order to add hydrogens, remove counterions, generate ionization states at physiological pH, generate low-energy ring conformations, optimize their geometries and generate conformers. In particular, a maximum number of 100 conformations were generated for each molecule in an energy window of 10 kcal/mol from the relative global minimum. Phase 3.4 was used for the searching and the number of minimum matches that the finally selected compounds should share with the hypothesis-2 was set to 3 pharmacophoric features out of the 5 (Figure S4 Supporting Information). Each screening was launched separately indicating 5000 as maximum number of hits. As a result of the screening of the three libraries, a total of 8621 molecules was selected. As criteria for filtering only those hits that could show a better affinity for *Hs*CKα1, ligands showing a Fitness score better than the most active compound in the test set according to hypothesis-2 (Fitness > 0.366) were taken into account. After this selection, 72 compounds were docked with Glide SP into the choline-binding site of *Hs*CKα1 using a grid of 10 Å centered on the HC-3. After this, Prime[12c] MM-GBSA rescoring calculations were performed using the default parameter set to compute the binding energies of the 72 hits. Among the 20 hits showing Prime MM-GBSA binding scores lower than −36.79 kcal/mol, 11 compounds were selected according to their commercial availability and prices to be tested on the *Hs*CKα1.

2.3 Cloning and Purification of hCK

Details about cloning and purification of human CKα1 and CKβ have been previously reported.[8b]

2.4 Tryptophan Fluorescence

All compounds were prepared in 100% DMSO. Before running the binding experiments, fluorescence property of the compounds was evaluated measuring the fluorescence signal at varying concentrations in a specific buffer (25 mM Tris, 150 mM NaCl, pH 7.5) containing an equivalent % of DMSO as the one used for the binding assay. The excitation wavelength was 280 nm while the emission was between 300–400 nm (typical emission range for protein tryptophan). No fluorescence signal was detected for the compounds that could interfere with the binding assay.

The binding affinity constant (K_ds) of the compounds against *Hs*CKα1 and *Hs*CKβ, were measured by monitoring the quenching of tryptophan fluorescence. All experiments were carried out in a Cary Eclipse spectrofluorometer (Varian) at 25 °C with the enzymes at 1 μM, and concentrations of compounds varying from 0.1 to 10 μM for *Hs*CKα1 and 0.5 to 200 μM for *Hs*CKβ in 25 mM Tris, 150 mM NaCl, pH 7.5. Fluorescence emission spectra were recorded in the 300–400 nm range with an excitation wavelength of 280 nm, with slit width of 5 nm. Controls were determined by incubating the enzymes with equivalent amounts of DMSO. As indicated previously, data analysis was performed in Prism 6 (GraphPad software)[21] considering a model with a single binding site (Equation 1), where F0 is the intrinsic fluorescence of the enzyme in the absence of quencher (Q), F1 is the observed fluorescence at a given quencher concentration, f_a is the fractional degree of fluorescence, and K_d is the dissociation constant.

$$1 - \frac{F1}{F0} = \frac{fa * [Q]}{Kd + [Q]} \quad (1)$$

2.5 Protein Crystallography

*Hs*CKα1 at 11.60 mg/ml was preincubated at room temperature with 7.5 mM hit **10** in buffer 25 mM Tris/HCl, 150 mM NaCl pH 7.5 (DMSO is at 5% final concentration in the mix). The sitting-drop vapor diffusion method[22] was used to produce crystals by mixing 0.5 μL of the protein solution and an equal volume of mother liquor (crystals appeared in

18% polyethylene glycol (PEG) 5000, 0.2 M calcium acetate, 0.1 M 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 8.5 and 1.0 M guanidine hydrochloride). Orthorhombic crystals (space group $P2_12_12_1$) grew within 6–7 days. The crystals used in this study were cryoprotected in mother liquor solutions containing 20 % ethylenglycol and frozen in a nitrogen gas stream cooled to 100 K.

Diffraction data of the binary complexes were collected at beamline I04–1 (Diamond, Oxford). The data was processed and scaled using the XDS package[23] and CCP4 software;[24] relevant statistics are given in Table S3 (Supporting Information).

2.6 Structure Determination and Refinement

The structure of the binary complex was solved by molecular replacement using PDB entry 3G15 as a template. Initial phases were further improved by cycles of manual model building in Coot[25]and refinement with REFMAC5.[26] The final model was validated with PROCHECK,[27a] model statistics are given in Table S3. Coordinates and structure factors have been deposited in the Worldwide Protein Data Bank (wwPDB; see Table S3 Supporting Information for the pdb code).

3 Results and Discussion

3.1 Computational Studies Suggest a Pharmacophore Model that is Present in the Fragment 1-Benzyl-4-(*N*-methylaniline)pyridinium

A focused library of 83 HC-3 derivatives (Figure S1) registered in the Binding Database was assembled to carry out our computational studies. This group of molecules varied from poor to highly active compounds according to their IC_{50} values obtained by studies on cancer cell lines and partially purified CK. Most of the molecules as well as HC-3 (**1**) contain a biphenyl-moiety acting like a spacer between two cationic heads that in turn serve as a criterion to classify these compounds as symmetrical or asymmetrical (Figure 2a and Figure S1). The majority of them contain pyridine and quinoline cationic heads with the exception of two rarely cationic fragments of 3-quinuclidinol and 3-dimethylamino1-propanol (Figure S1). Specifically, the pyridine derivatives have the aromatic ring covalently bound to a tertiary nitrogen, leading to a larger delocation of the electrons (Figure S1). Initially these molecules were docked into the choline-binding site of *Hs*CKα1 (PDB entry 3G15) pursuing a structure-based approach. As a result, no difference was found in the binding modes of the top ranked

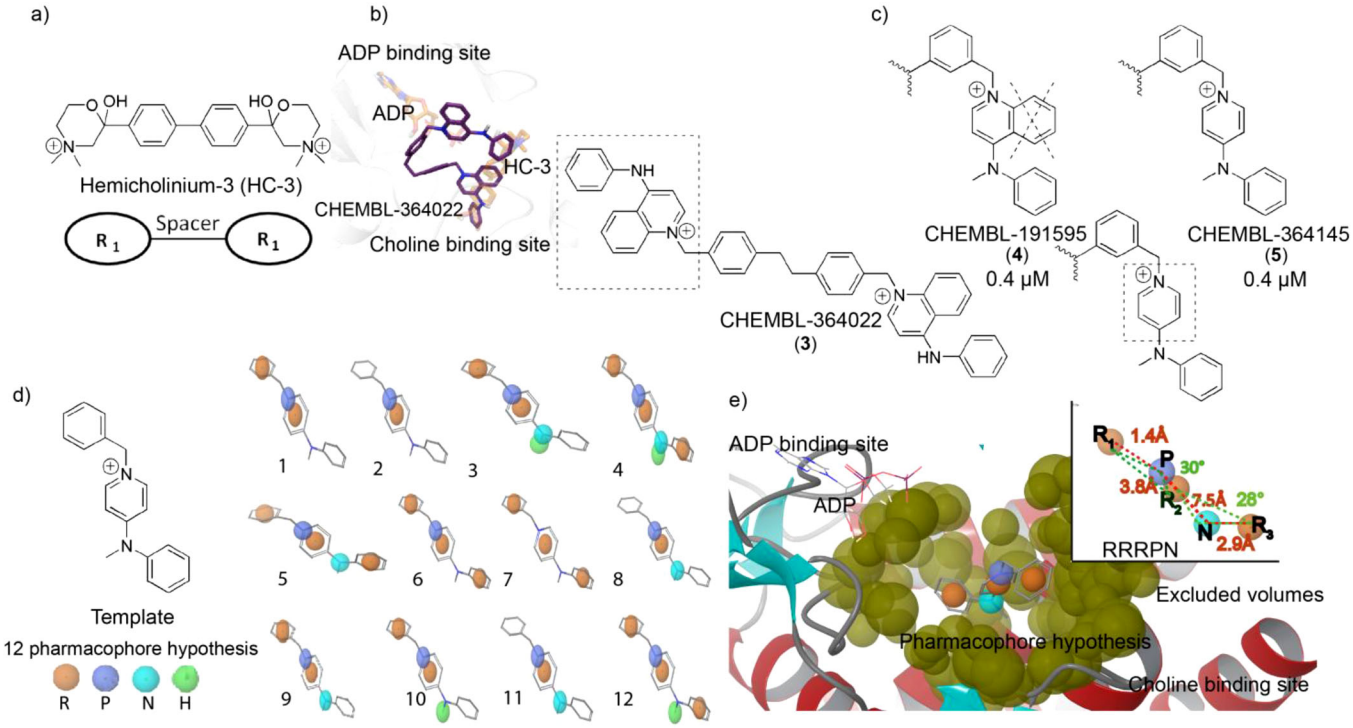


Figure 2. Pharmacophore hypothesis development. a) General chemical structure of the 83 Hemicholinium-3 (HC-3) derivatives that serve as starting point of the study. b) Docking pose of CHEMBL-364022 (**3**). c) Comparison of compounds **4**–**5**: the phenyl ring condensed to the pyridine does not seem to be essential for the enzyme inhibition. d) Feature combinations (R: ring, P: positive charge, N: nitrogen atom and H: hydrophobic) for finding the best hypothesis called hypothesis-2 (number #5). e) Pharmacophore model with the signature PNRRR.

poses of actives and inactives ligands, and thus no attempt to rationalize the different inhibitory activities could be done after this first step.

Consequently, and owing to the non-successful structure-based strategy, we deemed of interest to undertake a ligand-based approach using HC-3 to generate a first attempt of pharmacophoric hypothesis (hypothesis-1) composed of excluded volumes and 6 features (RRDDAP, Figure S2 Supporting Information) as detailed in the method section. The pharmacophoric hypothesis-1 was challenged with the objective of differentiating active from inactive compounds in the focused library of 83 molecules. Again, only 14 compounds of the library were successfully classified, with the pharmacophoric hypothesis-1 being in most cases unable to identify active and inactive compounds on the basis of the Fitness score. Hence, we next focused on one of the most active compound in the library, coded CHEMBL-364022 (**3**) according to the database of bioactive drug-like small molecules, and endowed with an IC_{50} value of 2.3 μM (Figure 2b). CHEMBL-364022, as suggested by docking studies, binds to the choline-binding site with its 1-benzyl-4-(*N*-phenyl)quinolinium moiety, whereas the remaining part of the molecule is exposed to the solvent (Figure 2b). Comparing this fragment to the CHEMBL-191595 (**4**) and other compounds of the library such as CHEMBL-364145 (**5**), we noticed that the phenyl condensed to the quinoline moiety should not be important for the inhibitory activity, since the IC_{50} values of these compounds were in the same range of potency (Figure 2c). For this reason, the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment was selected as interesting scaffold for generating a pool of twelve alternative pharmacophoric hypotheses (Figure 2d) that included different combinations of the positive charge of the pyridinium ring (feature P), a hydrogen bond donor group (feature N), a hydrophobic aliphatic group (feature H), and three aromatic groups (features R). The best hypothesis, namely hypothesis-2 (n. #5, Figure 2d), was the one able to explain the activity of 8 out of 22 compounds (4 actives and 4 inactives compounds, Table S2) and to show a trend between the Fitness score and the IC_{50} values of the compounds (Table S2). As shown in Figure 2e, hypothesis-2 was composed of a positive charge (P), a hydrogen bond donor group (N) and three aromatic groups (RRR).

3.2 A Ligand-Based Virtual Screening on HsCKα1 Provides New Potential Hits that are Validated by Tryptophan Fluorescence Studies

The pharmacophoric model based on hypothesis-2 was used to discover smaller and simpler compounds. To this end, a virtual screening protocol was carried out using three filtering steps (Figure 3a). Firstly, the hypothesis-2 enriched with excluded volumes derived from the binding mode of CHEMBL-364022 to *Hs*CKα1 was employed to screen the Enamine, Chembridge and Life Chemicals virtual compound libraries (8621 compounds in total). Specifically, compounds showing Fitness score higher than 0.366 and compliant to excluded volumes were selected, yielding 72 molecules. In order to further reduce the number of compounds, this set of molecules was next docked into the binding site of *Hs*CKα1.

The criterion for compound selection in the second filtering step was the Prime MM-GBSA docking score. At this regard, it should be mentioned that a recent study has reported the improvement resulting from the incorporation of Prime MM-GBSA scoring function in docking studies.[28] Accordingly, only the first twenty compounds having MM-GBSA scores lower than −36.79 kcal/mol were selected for further evaluations. From the resulting list of 20 compounds, only 11 of them were selected and purchased (Figure S5) in the third filtering step on the basis of their availability and pricing quotes.

The final 11 compounds were evaluated on *Hs*CKα1 using tryptophan fluorescence spectroscopy. Of these molecules, only 6 compounds (hits **6**–**11**) showed binding activity to the enzyme, with K_ds ranging from 0.44 to 7.9 μM (Figure 3b and Table 1).

To assess whether these compounds selectively bind to *Hs*CKα1, we also carried out the same experiments on *Hs*CKβ. A recent manuscript clearly showed that a potent anti-cancer effect inducing maximum apoptosis was only achieved when *Hs*CKα expression was specifically diminished without affecting *Hs*CKβ levels.[29] Therefore it is important to have selective and potent compounds against *Hs*CKα versus the *Hs*CKβ isoform. According to the K_d values, the compounds are selective on the *Hs*CKα1, being hits **7** and **9** the most selective of all, respectively (Table 1). Selectivity has been previously explained by a major flexibility of Trp420 in *Hs*CKα1 compared to a reduced flexibility of its homologous (Trp353) in *Hs*CKβ. This is owed to the location of a leucine at the back of Trp420 in *Hs*CKα1 instead of a more bulky phenylalanine at the back of Trp353 in *Hs*CKβ leading to a less flexible tryptophan.

Another important aspect of these compounds comes from the analysis of their ligand efficiency (L.E.) values[30] compared to the ones previously described for the HC-3 derivatives. This feature is defined as the ratio of ΔG and the number of non-hydrogen atoms of the molecules. The majority of the molecules in this study present better L.E. than previous heavier compounds such as HC-3, compound **12** (PDB ID: 3ZM9)[8c] and compound **13** (PDB ID: 4BR3)[8b] (Table 1). In particular hit **9** with the best K_d is the second more efficient molecule with a L.E. of −0.34 whereas hit **10** despite having a moderate K_d, has a L.E. of −0.38, and therefore it is the more efficient molecule (Table 1). A plausible explanation for the lower L.E. values found in HC-3 derivatives is that a significant fragment of these HC-3 derivatives is exposed to the solvent and consequently does not show interactions with the enzyme. This is further supported by our docking studies (Figure 2b) and previous crystal structures in complex with these heavier compounds.[8b]

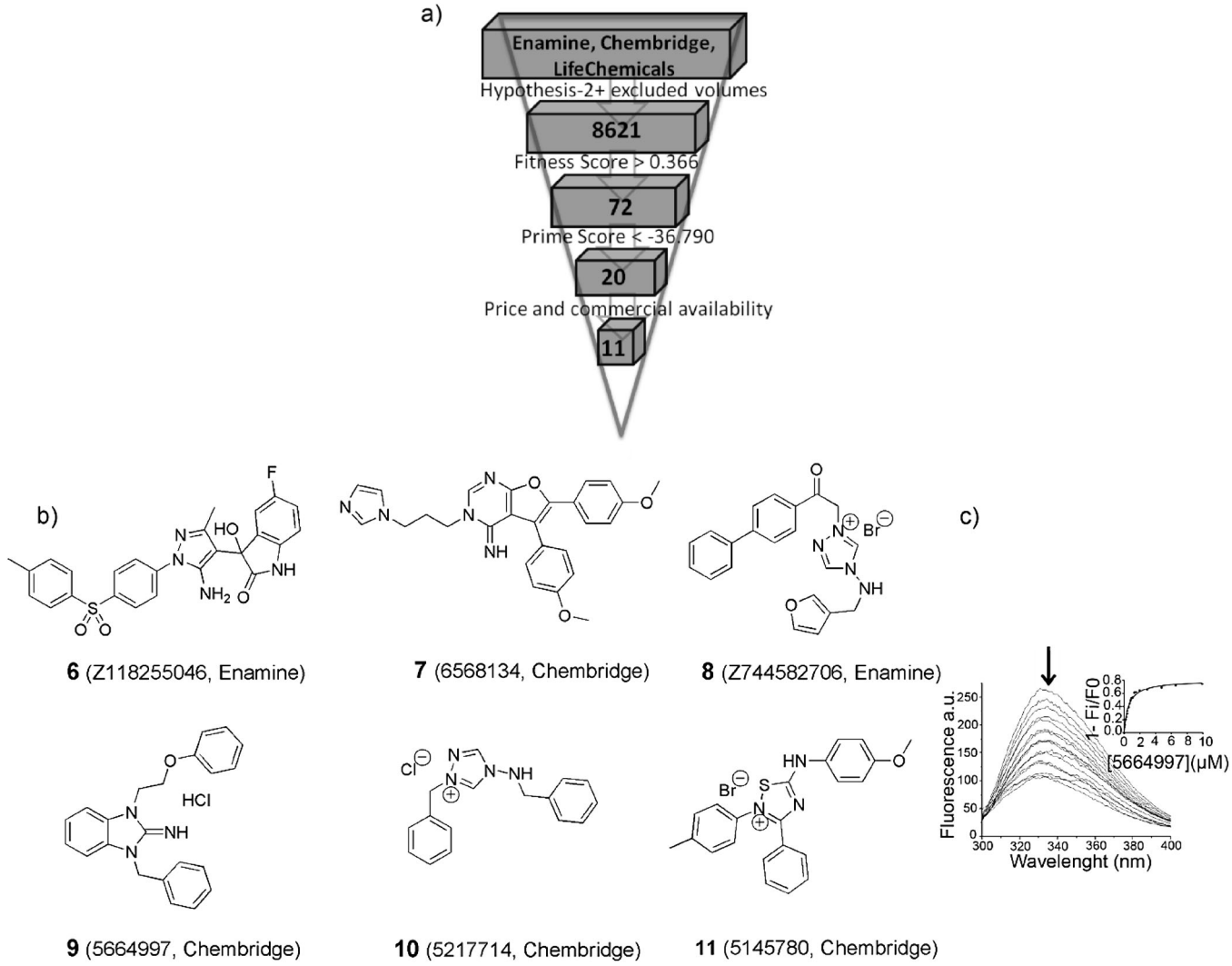


Figure 3. a) Hits filtering steps. b) Hits that showed binding to *Hs*CKα1 in the low μM range, and companies from which they are purchased. c) Quenching of intrinsic *Hs*CKα1 tryptophan fluorescence measured at increasing concentrations of hit **9**. A black arrow indicates the decrease of the maximal fluorescence signal with increasing concentrations of the compound.

Table 1. K_d values of the six hits for *Hs*CKα1 and *Hs*CKβ. All the hits are 3–160 fold selective on the *Hs*CKα1 over *Hs*CKβ. Hit **9** shows the best binding to the *Hs*CKα1, the greatest selectivity for this isoform and is also one of the compounds with the best ligand efficiency (L.E.). If the hits are compared with HC-3, most of them show a better value of ligand efficiency.

Hit (Supplier)	K_d (μM)		*Hs*CKα1	
	*Hs*CKα1	*Hs*CKβ	Selectivity (K_d for CKβ)/(K_d for CKα1)	L.E. (kcal/(mol atom))
9[19]	0.438 ± 0.1	70 ± 9	160	−0.34
8[18]	0.526 ± 0.1	36.5 ± 3.2	69	−0.32
10[19]	2.60 ± 0.7	59.5 ± 3.7	23	−0.38
11[19]	4.14 ± 1.3	41.3 ± 7.8	10	−0.27
7[19]	7.4 ± 1.6	23 ± 0.8	3	−0.21
6[18]	7.9 ± 0.5	68.3 ± 2.3	9	−0.20
HC-3 (1)	0.180 ± 0.05	–	–	−0.26

3.3 Hit 10 Binds to the Choline and ATP-Binding Sites

Although docking studies clearly suggested that the compounds bound to the choline-binding site (Figure S6 Supporting Information), we further carried out crystallization experiments with all the positive hits (Figure 3b) to have more convincing data. Despite of a large number of trials

we only managed to get the crystal structure of *Hs*CKα1 in complex with hit **10** at 2.25 Å (Methods and Table S3). Successive iterative model building and refinement cycles were carried out to produce a final model with good refinement statistics ($R = 0.224$, $R_{free} = 0.268$ for *Hs*CKα1/hit **10**, Table S3 Supporting Information). The structure is a dimer in which each monomer is formed by a small *N*-terminal and a large *C*-terminal. Whereas the ATP binding site is located in a cleft formed by *N*- and *C*-terminal residues, the choline-binding site is found in a deep hydrophobic pocket. Two molecules of hit **10** were visualized into the choline and ATP binding sites of one of the monomers (molecule-1 and molecule-2. Figure 4a). Conversely, only one molecule was found in the choline-binding site of the second monomer (molecule-1, Figure 4a). Molecule-1, which is the one that occupies a deeper location in the choline-binding site, has also better defined electron density than molecule-2. Molecule-1 establishes π–π interactions with Trp423 and Phe435, a cation–π interaction with Trp420 and a hydrogen bond with Glu349 (Figure 4a) whereas molecule-2 establishes an electrostatic interaction with Asp330 and a hydrogen bond with Asn311 (Figure 4a). Overall, this compound contains 3 of the 5 moieties present in our proposed pharmacophore (Figure S4) that in turn are required to interact with the protein as described above (Figure 2e and 4a). Owing to these interactions, hit **10** shows two binding modes that allow on one hand to compete with the magnesium atom (this metal is coordinated by Asp330 and Asn311 and the phosphate groups) and the ADP phosphate groups (Figure 4b), and on the other hand to mimic the positive charge present in HC-3,[31] choline/phosphocholine, and HC-3 derivatives such as compound **12**,[8c] compound **13**[8b] and compound **14**[8a] (Figures 4b–f). In all these compounds the positive charge is key to maintain a cation-π interaction with Trp420. Compound **1,** as well as hit **10,** is also located in the ATP and choline-binding sites and therefore both of them act as dual competitive inhibitors of choline and ATP.

Overall, the results of crystallization studies of *Hs*CKα1 in complex with hit **10,** combined with previous structural data,[8] contribute to support a scenario in which the majority of *Hs*CKα1 inhibitors generally occupy the choline-binding site, using only one positive charged moiety of the molecule to engage aromatic residues in cation-π interactions.

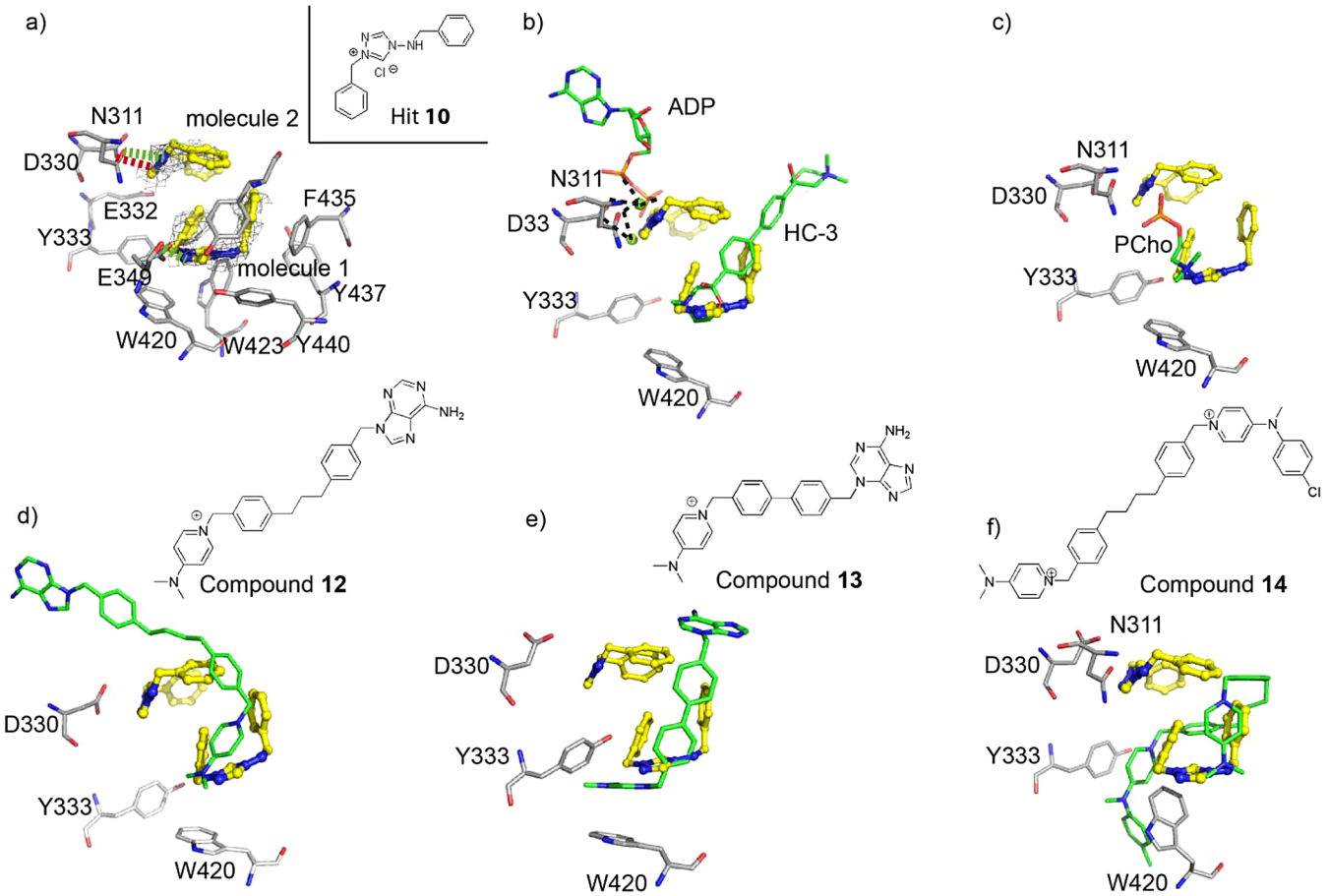


Figure 4. Crystal structures of *Hs*CKα1 complexes. Close-up views of hit **10** (a), HC-3 and ADP (b, PDB entry 3G15), phosphocholine (c, PDB ID entry 2CKQ), compound **12** (d, PDB ID entry 3ZM9), compound **13** (e, PDB ID entry 4BR3) and compound **14** (f, PDB ID entry 4CG8) bound to the *Hs*CKα1 active site, respectively. Hydrogen bonds and electrostatic interactions are represented in green and pink dashes, respectively.

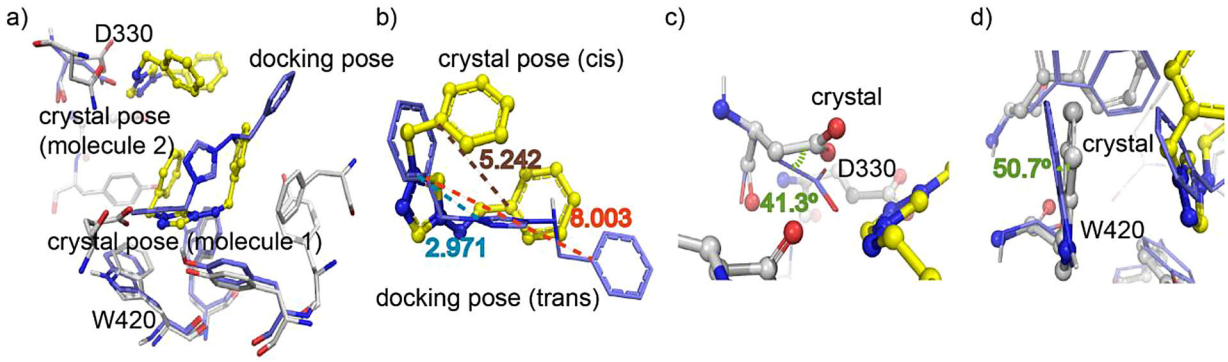


Figure 5. Differences in the pose of hit **10** resulting from docking studies and X-ray crystallography. Residues in the crystal structure are colored in light grey whereas the predicted docking and experimental poses are shown in blue and yellow, respectively.

Although the second positive charge of biscationic compounds is usually exposed to the solvent, for one particular compound inducing the aperture of an adjacent new binding site to the choline binding site, it was found that both positive charges are engaged in cation-π interactions.[8a] Remarkably, a second copy of hit 10 has been identified in the ATP binding site, mimicking a magnesium atom that is coordinated by the ATP γ-phosphate and making electrostatic and hydrogen bond interactions with polar residues.

Despite the accuracy in the prediction of the binding mode of the compounds by docking studies, some deviations were found analyzing a superposition of the docked pose of hit **10** with the adopted one in the crystal structure (Figure 5). These deviations are in terms of the number of molecules of the compound that are present in the binding site, the bioconformation of the compound, the location of the positive charge and the rotation of some important residues involved in the binding. More exactly, the docking studies are not able to predict: a) the entering of two molecules of hit **10** into the choline-binding pocket (Figure 5a); b) its cis-bioconformation, where the real distance between the aromatic moieties is 5.242 Å instead of 8.003 Å (Figure 5b); c) the location of the quaternary nitrogen, which in the crystal structure is 2.971 Å far from its docking pose (Figure 5b) and d) the noticeable rotation (more than 40°) that some residues such as Asp330 (Figures 5c) and Trp420 (Figure 5d) adopt in the crystal structure. All these data provide clues about the current limitations of the docking programs, which must be taken into account in the interpretation of any theoretical result.

4 Conclusions

CK has an essential role in the biosynthesis of cellular membrane phospholipids and in cell proliferation. In humans, the α isoform has been proposed as a validated drug target to treat cancer and consequently along the last years a large number of HC-3 derivatives acting as inhibitors of *Hs*CKα1 have been discovered. Generally these compounds are large molecules whose common structural features, important for their binding mechanism and biological activity, have not been elucidated. To dissect the key features of HC-3 derivatives, we have conducted a computational study to propose a pharmacophore model formed by five moieties that are included in the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment. Then this pharmacophore model has been used to carry out a virtual screening of several commercial libraries leading to the finding of six small monocationic compounds with better ligand efficiencies than the previously reported ones. Although docking studies suggest that these 6 molecules bind to the choline-binding site, our crystal structure of *Hs*CKα1/hit **10** suggests that this particular molecule has a dual mechanism and might compete with ATP/Mg^{+2} and choline. In conclusion, a pharmacophore model is provided to unravel the key moieties of HC-3 derivatives that probably account for their biological functions, and to discover new efficient and simpler *Hs*CKα1 ligands that may represent starting lead compounds on the way to develop novel anticancer agents.

Supporting Information

HC-3 derivatives found in the Binding Database; HC-3 hypothesis (hypothesis-1); Rings superimposition between the crystal pose of HC-3 and the docked pose of the 1-benzyl-4-(*N*-methylaniline)pyridinium fragment; features agreement between the final pharmacophore hypothesis and the hits; docking pose of the hits that showed a K_d in the low μM range; 22 compounds of the test set; 11 selected compounds to be tested; Fitness score of the test set compounds selected by the final hypothesis; X-ray crystal structure data collection and refinement statistics.

Acknowledgement

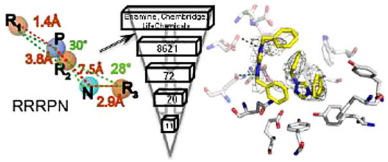
The research leading to these results has received funding from the European Community Seventh Framework Pro-

gramme (FP7/2007–2013) under BioStruct-X (Grant 724 N° 283570). We thank the *Ministerio de Ciencia e Innovación of Spain* (SAF2009-11955, BFU2010-19504, AP2009-0555) and the Fundación *Agencia Aragonesa para la Investigación y el Desarrollo* (*ARAID*, Spain). *The Diamond Light Source* (*DLS*) at Oxford, especially beamline I04-1 (experiment MX 8035-11) is gratefully acknowledged.

References

[1] P. Fagone, S. Jackowski, *Biochim. Biophys. Acta* **2013**, *1831*, 523 – 532.
[2] a) G. C. Whiting, S. H. Gillespie, *FEMS Microbiol. Lett.* **1996**, *138*, 141 – 145; b) V. Sampels, A. Hartmann, I. Dietrich, I. Coppens, L. Sheiner, B. Striepen, A. Herrmann, R. Lucius, N. Gupta, *J. Biol. Chem.* **2012**, *287*, 16289 – 16299; c) T. Miyake, S. J. Parsons, *Oncogene* **2012**, *31*, 1431 – 1441.
[3] a) D. Gallego-Ortega, A. Ramirez de Molina, M. A. Ramos, F. Valdes-Mora, M. G. Barderas, J. Sarmentero-Estrada, J. C. Lacal, *PLoS One* **2009**, *4*, e7819; b) A. Ramirez de Molina, A. Rodriguez-Gonzalez, R. Gutierrez, L. Martinez-Pineiro, J. Sanchez, F. Bonilla, R. Rosell, J. Lacal, *Biochem. Biophys. Res. Commun.* **2002**, *296*, 580 – 583; c) A. de la Cueva, A. Ramirez de Molina, N. Alvarez-Ayerza, M. A. Ramos, A. Cebrian, T. G. Del Pulgar, J. C. Lacal, *PLoS One* **2013**, *8*, e64961.
[4] a) K. S. Crilly, M. Tomono, Z. Kiss, *Arch. Biochem. Biophys.* **1998**, *352*, 137 – 143; b) R. Hernandez-Alcoceba, L. Saniger, J. Campos, M. C. Nunez, F. Khaless, M. A. Gallo, A. Espinosa, J. C. Lacal, *Oncogene* **1997**, *15*, 2289 – 2301.
[5] a) G. B. Ansell, S. G. Spanner, *J. Neurochem.* **1974**, *22*, 1153 – 1155; b) J. J. Freeman, J. W. Kosh, J. S. Parrish, *Br. J. Pharmacol.* **1982**, *77*, 239 – 244.
[6] a) V. Gomez-Perez, T. McSorley, W. C. See Too, M. Konrad, J. M. Campos, *ChemMedChem* **2012**, *7*, 663 – 669; b) A. Conejo-Garcia, J. M. Campos, R. M. Sanchez-Martin, M. A. Gallo, A. Espinosa, *J. Med. Chem.* **2003**, *46*, 3754 – 3757; c) S. Schiaffino-Ortega, L. C. Lopez-Cara, P. Rios-Marco, M. P. Carrasco-Jimenez, M. A. Gallo, A. Espinosa, C. Marco, A. Entrena, *Bioorg. Med. Chem.* **2013**, *21*, 7146 – 7154; d) R. Hernandez-Alcoceba, F. Fernandez, J. C. Lacal, *Cancer. Res.* **1999**, *59*, 3112 – 3118; e) R. Sanchez-Martin, J. M. Campos, A. Conejo-Garcia, O. Cruz-Lopez, M. Banez-Coronel, A. Rodriguez-Gonzalez, M. A. Gallo, J. C. Lacal, A. Espinosa, *J. Med. Chem.* **2005**, *48*, 3354 – 3363.
[7] K. Kirk, R. E. Martin, S. Broer, S. M. Howitt, K. J. Saliba, *Curr. Top. Microbiol. Immunol.* **2005**, *295*, 325 – 356.
[8] a) B. Rubio-Ruiz, A. Figuerola-Conchas, J. Ramos-Torrecillas, F. Capitan-Canadas, P. Rios-Marco, M. P. Carrasco, M. A. Gallo, A. Espinosa, C. Marco, C. Ruiz, A. Entrena, R. Hurtado-Guerrero, A. Conejo-Garcia, *J. Med. Chem.* **2014**, *57*, 507 – 515; b) M. Sahun-Roncero, B. Rubio-Ruiz, A. Conejo-Garcia, A. Velazquez-Campoy, A. Entrena, R. Hurtado-Guerrero, *Chembiochem.* **2013**, *14*, 1291 – 1295; c) M. Sahun-Roncero, B. Rubio-Ruiz, G. Saladino, A. Conejo-Garcia, A. Espinosa, A. Velazquez-Campoy, F. L. Gervasio, A. Entrena, R. Hurtado-Guerrero, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 4582 – 4586.
[9] a) B. F. Clem, A. L. Clem, A. Yalcin, U. Goswami, S. Arumugam, S. Telang, J. O. Trent, J. Chesney, *Oncogene* **2011**, *30*, 3370 – 3380; b) C. S. Hudson, R. M. Knegtel, K. Brown, P. A. Charlton, J. R. Pollard, *Biochim. Biophys. Acta* **2013**, *1834*, 1107 – 1116.
[10] *Maestro*, Version 93, Schrödinger, LLC, New York, NY, **2012**.
[11] http://www.rcsb.org/.■not cited in the text■
[12] L. S. Schrödinger Suite, **2012**, Protein Preparation Wizard; a) *Epik*, Version 2.3, New York, NY, **2012**; b) *Impact,* Version 5.8, Schrödinger, LLC, New York, NY, **2012**; c) *Prime,* Version 3.1, Schrödinger, LLC, New York, NY, **2012**.
[13] http://www.bindingdb.org/.
[14] V. *Glide*, Schrödinger, LLC, New York, NY, **2012**, *Glide*, Version 58, Schrödinger, LLC, New York, NY, 2012.
[15] N. K. Salam, R. Nuti, W. Sherman, *J. Chem. Inf. Model.* **2009**, *49*, 2356 – 2368.
[16] P*hase,* Version 34, Schrödinger, LLC, New York, NY, **2012**.
[17] C*onfGen,* Version 23, Schrödinger, LLC, New York, NY, **2012**.
[18] http://www.enamine.net/.
[19] http://www.chembridge.com/.
[20] http://www.lifechemicals.com/.
[21] www.graphpad.com.
[22] H. Adachi, K. Takano, M. Morikawa, S. Kanaya, M. Yoshimura, Y. Mori, T. Sasaki, *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 194 – 196.
[23] W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125 – 132.
[24] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658 – 674.
[25] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126 – 2132.
[26] A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, G. N. Murshudov, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2184 – 2195.
[27] a) R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, *J. Biomol. NMR* **1996**, *8*, 477 – 486; b) http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/index.html.
[28] P. A. Greenidge, C. Kramer, J. C. Mozziconacci, W. Sherman, *J. Chem. Inf. Model.* **2014**, *54*, 2697 – 2717.
[29] J. Gruber, W. C. See Too, M. T. Wong, A. Lavie, T. McSorley, M. Konrad, *Febs. J.* **2012**, *279*, 1915 – 1928.
[30] C. Abad-Zapatero, *Expert Opin. Drug Discov.* **2007**, *2*, 469 – 488.
[31] B. S. Hong, A. Allali-Hassani, W. Tempel, P. J. Finerty Jr. , F. Mackenzie, S. Dimov, M. Vedadi, H. W. Park, *J. Biol. Chem.* **2010**, *285*, 16330 – 16340.

Received: October 17, 2014
Accepted: March 10, 2015
Published online: ■■ ■■, 0000



L. Serrán-Aguilera, R. Nuti, L. C. López-Cara, M. Á. G. Mezo, A. Macchiarulo, A. Entrena,* R. Hurtado-Guerrero**

■■ – ■■

Pharmacophore-Based Virtual Screening to Discover New Active Compounds for Human Choline Kinase α1