Theor Chem Account (2008) 120:523–531
DOI 10.1007/s00214-008-0431-x

REGULAR ARTICLE

Computational methods applied to the discovery of stem cell factor ligands

Stefano Alcaro · Lorenzo Gontrani · Ottaviano Incani · Francesco Ortuso

Received: 14 November 2007 / Accepted: 18 February 2008 / Published online: 13 March 2008

Abstract A computational study of the stem cell factor (SCF) and potential ligands was carried out starting with a crystallographic model deposited in the protein data bank. The inhibition of the SCF dimerization equilibrium was considered as the rationale for the lead identification of specific ligands. A preliminary molecular dynamics characterization of the SCF dimer allowed to verify the most flexible loop involved in the dimeric area. Then a virtual screening, coupled with energy minimization in GB/SA water, scored the compounds implemented in the NCI diversity molecular database. Ten top ranked ligands were analyzed considering both the SCF loop perturbation in the dimerization area and the network of intermolecular hydrogen bonds. Among these ten compounds two natural agents were identified. The computational work revealed useful new insights for rational design of novel SCF dimerization inhibitors.

Keywords SCF · c-kit · Virtual screening · Molecular dynamics · NCI-diversity database · Protein–protein interface

Contribution to the Nino Russo Special Issue.

S. Alcaro (✉) · F. Ortuso
Laboratorio di Chimica Farmaceutica Computazionale,
Dipartimento di Scienze Farmacobiologiche,
Università degli Studi "Magna Græcia" di Catanzaro,
Edificio delle Bioscienze, Viale Europa, 88100 Catanzaro, Italy
e-mail: alcaro@unicz.it

L. Gontrani
Consorzio per le Applicazioni di Supercalcolo Per Università
e Ricerca (CASPUR), Via dei Tizii, 6, 00185 Rome, Italy

O. Incani
OpenShell s.a.s. Via Papa Giovanni XXIII, 60,
67039 Sulmona (AQ), Italy

1 Introduction

The stem cell factor (SCF) is an endogenous growth factor involved in the haematopoiesis cell proliferation and differentiation. It plays a crucial role in the development of melanoma and several intestinal tumors [1]. Similar to other growth factors, the SCF dimerization is a necessary prerequisite to exert the activation of its natural tyrosine kinase receptor c-kit [2]. The SCF c-kit interaction leads the first step of a biochemical cascade responsible of several effects, including the cell hyperproliferation.

It has been postulated that the growth factor can exist in two forms: the monomeric soluble SCF and dimeric cell surface bound SCF [3]. The equilibrium between the two forms can be considered as an attractive target for the design and/or the identification of selective inhibitors of c-kit activation. The availability of a crystallographic model of SCF [4] into the protein data bank (PDB) [5] allows to apply the receptor-based paradigm for the drug discovery of leads able to interact in the dimerization area, thus inducing the growth factor toward the monomeric inactive form.

The goal of our scientific project is the identification of potentially active inhibitors against the SCF dimerization. Unlike classical drug discovery, where small molecules capable of binding to the active site of an enzyme are used for the design of new ligands, the development of inhibitors of protein–protein associations is a very difficult task [6]. The key points we considered were: (1) the molecular surface area involved in the recognition of a ligand at the protein–protein interface is quite large (750 $Å^2$ or more); (2) the interface area is usually not as well defined as the binding pocket of an enzyme, since it is often shallow and featureless; and (3) the type and the extent of rearrangements of protein structure when protein–protein interaction is involved may

be different from that resulting from the induced fit due to small molecule recognition.

These points influenced our procedure as follows: (1) the need to consider not just drug-like molecules as larger ligands could also be effective in the discovery process; (2) the requirement for high chemical diversity in the screened compounds; and (3) protein flexibility has to be taken into account fully, when exploring the possible recognition sites.

In this communication we describe a virtual screening study carried out using as target the dimerization SCF surface and as ligands the compounds constituting the National Cancer Institute (NCI) molecular databases [7].

2 Computational details

2.1 Molecular dynamics

Human SCF conformational mobility was investigated by studying the dynamic behavior of the aminoacid residues located at the dimerization interface. The residues of each subunit falling within 8 Å from the center of the dimer were considered. In total, 44 residues (22 aminoacids for each omodimer subunit) were selected as the flexible portion. The simulation protocol can be summarized as follows. The omodimer crystal structure taken from the PDB repository (code: 1SCF) was first neutralized with the necessary Na^+ counterions and then solvated with a water spherical shell of radius 30 Å centered on the center of mass of the 44 residues. Acidic and basic side-chains were considered in ionized state. A total of 1,934 molecules, constituting this water cap, were added to the system. In doing so, an efficient, yet not too extended, hydration model for the interface was obtained. The global protein folding was preserved with a positional restraint which imposes a penalty (5 kcal/mol) on Cartesian coordinate change for backbone atoms far from the interface (i.e., outside the flexible portion), while the virtual evaporation of the solvent was prevented using a central attracting potential. The system was relaxed using standard conjugate gradient minimization methods (10,000 cycles) and then equilibrated at 300 K. Afterwards, a 1.5 ns-long production phase followed. Both equilibration and dynamics were carried out in NPT ensemble, using Berendsen's coupling scheme for temperature and pressure control (a value of 1.0 ps was used for both heat and pressure bath time constants). All simulations were performed with AMBER code (version 7) [8] using the computational facilities available at CASPUR.

2.2 Virtual screening

Docking simulations were performed using the version 1.8 of the Glide software [9]. Such an approach required a three-step procedure consisting of (a) receptor model pretreatment, (b) interaction energy maps calculation and (c) evaluation of possible ligands recognition. In the first step, the subunit A, from the 1SCF PDB model, was considered as receptor model. After adding missing hydrogen atoms, the modified crystallographic structure was submitted to 100 interactions of conjugate gradient energy minimization in vacuo (i.e., dielectric constant equal to 1) using the OPLS-AA force field [10] as implemented into the version 1.8 of the Impact program suite [11]. In order to prevent unrealistic distortions of the receptor model, the coordinates of all non-hydrogen atoms located at distances longer than 12 Å from the dimerization surface geometric center (DC) were frozen. The ionization state of acidic and basic side-chains was considered as in the molecular dynamic simulations. The energy and its gradient were adopted as convergence criteria with threshold values of 10^{-7} and 10^{-2} kcal/mol, respectively. In the second step, the optimized conformation was submitted to a default parameters Glide interaction energy maps calculation taking into account all residues within a cubic box of 46,656 $Å^3$ centered onto the DC. Finally, computed maps were used to evaluate the recognition of all compounds available into the NCI diversity database. The conformational properties of the docked molecules were taken into account by means of the Glide flexible docking algorithm. Docking generated poses were sorted by default Glide scoring function sampling only the best configuration for each ligand. With the aim to consider the induced fit phenomena, sampled complexes were submitted to energy minimization. All residues within 12 Å from the DC were left free of moving, while the others were constrained with a constant force of about 24 kcal/mol Å. The optimization procedure was performed by means of 1,000 iteractions of the Polack Ribiere conjugate algorithm coupled to OPLS-AA force field as implemented into the version 7.2 of MacroModel [11,12]. Water environment effects were taken into account using the implicit solvation model GB/SA [13]. According to the MOLINE method [14] the interaction energy of all optimized structures was computed. Partitioning properties at pH = 7 of the best ranked compounds were estimated by the theoretical log *D* calculated with the version 4.1.5 of the MarvinSketch software [15–18]. The graphical inspection the most stable SCF-ligand configurations (Figs. 5, 6, 7, 8, 9) was carried out by the PyMol software ver. 0.99 [19].

3 Results and discussion

Our study started from the assessment that an SCF dimerization negative modulation can be a useful and selective approach to prevent early-acting haematopoietic cancer activation. Actually no c-kit activity selective inhibitors are available, but only generic blockers of the tyrosine kinase modulated pathways (i.e., staurosporin) were developed [20].

The availability of the tetrameric crystallographic model 1SCF in the PDB was the essential structural prerequisite to start our computational work. No other detailed structural information, such as the c-kit binding domain complexed with the SCF, was available in the PDB repository as an alternative starting point.

Since molecules active against the SCF target are unknown, in particular disfavoring the dimerization equilibrium, we have focused our attention on the commercially available anticancer agents categorized and collected in molecular databases by the National Cancer Institute. Most of them exert their biological activity with unknown mechanism of action. So the goal of our study was to identify, among them, lead compounds able to consistently interact in the SCF dimerization area. In order to consider the largest molecular diversity we have selected the database known as "NCI-diversity" that features about 2,000 anticancer agents based on fairly different chemical scaffolds.

The first step of our work was the analysis and the pretreatment of the 1SCF model downloaded from the PDB. Such a model is composed of two dimers with symmetric "head-to-head" association. The molecular recognition of natural growth factor to the c-kit receptor is postulated to occur in a region of the SCF far from its dimerization area [3,4]. So the identification of dimer surface binding agents should lead to selective drugs, in principle better than nonspecific protein-tyrosine kinase inhibitors, such as STI-571 (Gleevec, imatinib), interacting to Abl, Bcr-Abl, Kit, and the PDGF receptor (α and β) [21].

The 1SCF model shows missing residues, such as the first ten in the primary sequence. However, the dimer interface is complete and is characterized by two loop regions, respectively, constituted by residues 17–26 (loop *a*) and 61–72 (loop *b*). Preliminarily we have analyzed the four monomer subunits A–D included into the 1SCF model by superimposition of the two loops. In Fig. 1 we have reported the results of such superimposition analysis.

The root mean square deviation (RMSD) computed separately onto the α carbon atoms of both loops *a* and *b* among all subunits showed valued, respectively, around 0.18–0.23 and 0.71–1.63 Å, confirming that A–D at the interface region were conformationally very conserved. In the crystal structure the relatively most flexible residues resulted those located on the loop *b* around the position 67–69 (Fig. 1).

In order to verify the mobility of all residues involved in the dimeric interface region in both loops *a* and *b* we have carried out a constrained molecular dynamics (MD) simulation of the dimer subunits A and B. The solvation was considered explicitly adding water molecules as reported in Sect. 2 (Fig. 2).

The analysis of the 22 residues in the dimeric interface region during the MD simulation allows to appreciate which

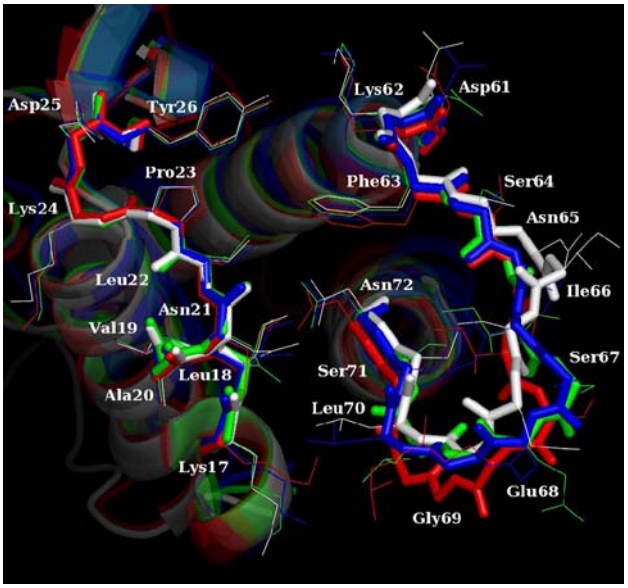


Fig. 1 Superimposition of subunits A (*white*), B (*red*), C (*green*) and D (*blue*) carried out onto the backbone atoms (polytube models) of loops *a* (residues 17–26) and *b* (residues 61–72). Side-chain atoms of loops *a* and *b* are depicted in wireframes, the remainder of the protein subunits are reported in ribbon models

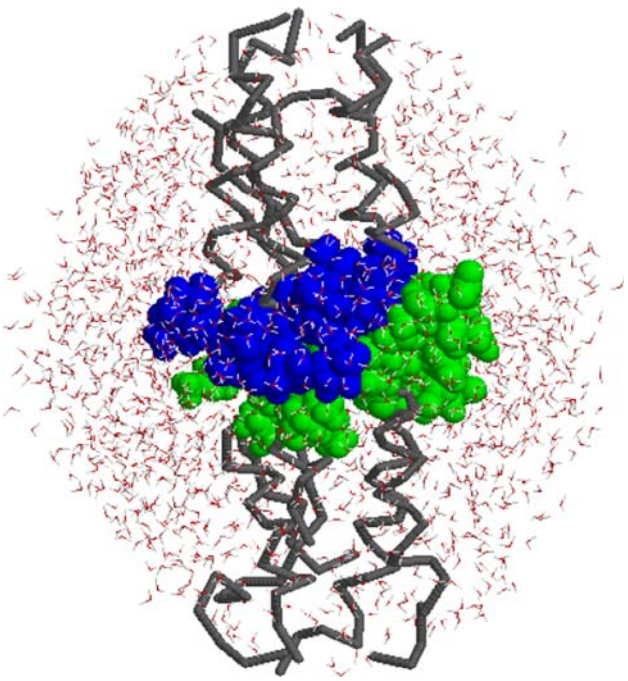


Fig. 2 Model of the SCF dimer used in the MD simulations. The rigidly treated residues are reported as carbon alpha polytubes. The residues free to move are reported in CPK models. *Blue spheres*. residues at dimer interface of chain A (*upper gray backbone*); *green spheres* residues interface of chain B (*gray backbone at the bottom*). Water molecules added explicitly into the model were reported as wireframe

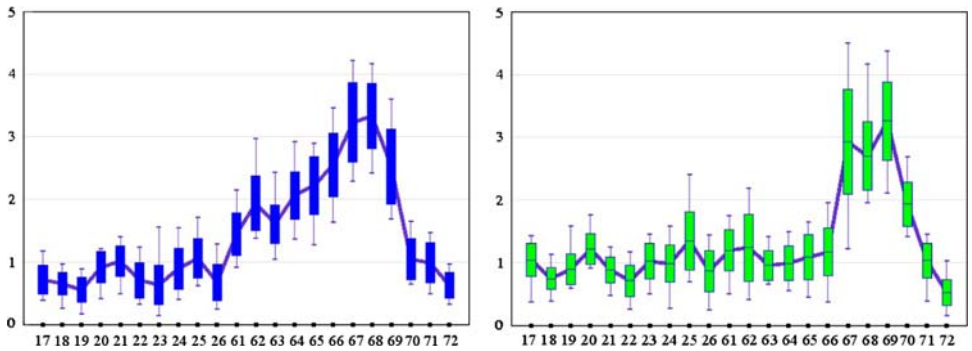


Fig. 3 RMS deviation in the residues 17–26 (loop *a*) and 61–72 (loop *b*) during the MD simulation. *Blue* chain A residues; *Green* chain B residues

Table 1 Summary of the top ten ranked complexes

Compound	NSC number	Log *D* (pH 7)	Total IntEne	vdW IntEne	Elec IntEne	GB/SA IntEne	SCF RMS	HBs
1	119886	10.23	−59.90	−55.32	−36.61	32.03	0.954	4
2	48458	7.71	−57.46	−63.23	−3.22	8.99	0.864	0
3	521777	2.41	−54.59	−40.20	−43.96	29.57	0.563	6
4	211094	1.58	−48.43	−32.05	−33.61	17.23	0.578	5
5	92601	−1.29	−47.53	−32.20	−41.35	26.02	0.513	6
6	135882	6.78	−47.52	−50.84	−0.10	3.42	0.855	0
7	111326	11.40	−47.34	−51.28	−1.97	5.91	0.482	0
8	293161	−4.13	−46.88	−43.26	−41.87	38.25	0.536	4
9	109197	9.05	−46.70	−49.31	−4.90	7.51	1.143	1
10	89818	2.64	−46.70	−28.18	−46.41	27.89	0.459	6

The NSC number is the code identifying uniquely the compound, logD is the calculated lipophilicity, Total IntEne is the sum of the van der Waals (vdW IntEne), electrostatic (Elec IntEne) and the GB/SA solvation (GB/SA IntEne) interaction energies expressed in kcal/mol. The SCF RMS is the root mean deviation, in Å, computed onto the energy minimized complex atomic coordinates with respect to the subunit A crystallographic model. HBs is the number of intermolecular hydrogen bonds between the ligand and the SCF

residues are more likely to undergo consistent folding perturbation.

In agreement with the observation after the crystallographic subunits A–D superimposition, the MD run indicated the loop *b* is particularly subject to perturbation, more than *a*. In both dimer subunits three residues (Ser67, Glu68 and Gly69), located on the solvent most exposed part of the loop *b*, were particularly flexible showing RMS deviations around 3 Å with respect to the crystal model. Such an observation suggested that the loop *b* flexibility could allow small compounds to interfere onto the dimerization process (Fig. 3).

In the second step of the computational work we have focused our attention onto the virtual screening experiments using the monomer A as target and the NCI-diversity database as source of potential selective SCF ligands. As reported in Sect. 2 we combined a ligand flexible docking search and a constrained energy minimization procedure with the aim to identify the best 10 fitting ligands within those collected in NCI database. The ranking criterion was based on the interaction energy computed onto the energy minimized 1:1 complex between the monomer SCF and the ligand. The analysis of the top 10 complexes was carried out considering the calculated logD at neutral pH [15–18], the perturbation of the SCF surface loops and the network of intermolecular ligand-protein hydrogen bonds established in the global minimum configurations. In Table 1 the analysis of the best 10 ranked compounds is reported.

Table 2 Statistical analysis of the interaction energies (IntEne) in kcal/mol computed in the virtual screening experiment carried out with the entire NCI-diversity database

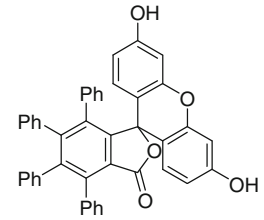
Statistical value	Total IntEne	vdW IntEne	Elec IntEne	GB/SA IntEne
Minimum	−59.90	−63.23	−221.53	−2.37
Maximum	−4.79	7.02	9.76	201.32
Average	−25.58	−23.89	−28.91	27.22
Standard deviation	7.00	8.01	26.64	23.06

In Table 2 a statistical analysis of the virtual screening interaction energies is summarized.

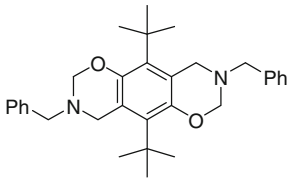
The chemical structures of the top 10 ranked compounds are reported in Fig. 4.

Since they belong to the NCI-diversity set the scaffolds resulted very diverse: derivatives of fluorescein (**1**), tricy-

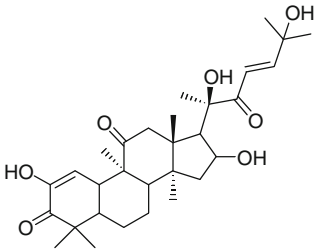
Fig. 4 Chemical formula and NSC identification code of compounds **1**–**10**



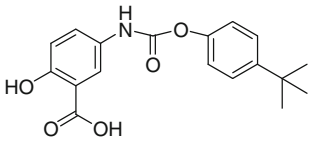
compound **1** (NSC 119886)



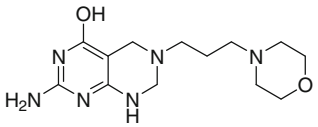
compound **2** (NSC 48458)



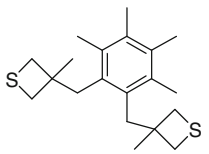
compound **3** (NSC 521777)



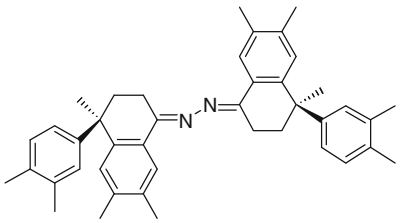
compound **4** (NSC 211094)



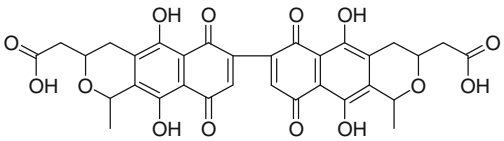
compound **5** (NSC 92601)



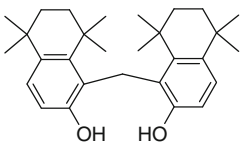
compound **6** (NSC 135882)



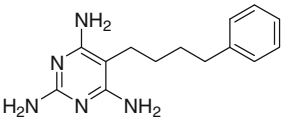
compound **7** (NSC 111326)



compound **8** (NSC 293161)



compound **9** (NSC 109197)



compound **10** (NSC 89818)

clic rings (**2**), cucurbitacin (**3**), salicylic carboxamide (**4**), tetrahydropurine with a morpholine side-chain (**5**), sulfured tetramethylbenzene (**6**), dihydronaphthalen hydrazone (**7**), actinorhodin (**8**), phenolic *bis* tetraline (**9**) and pyrimidine (**10**). It is worth to highlight that five of them (**1**, **2**, **6**, **7**, **9**) are characterized by hydrophobic moieties and log *D*s higher than 5, value generally considered as the maximum limit for a druggable compound according to the Lipinski's rules [22]. Four compounds are symmetric (**2**, **7**, **8**, **9**) and two are natural products (**3**, **8**).

The computed interaction energy was based on the combination of three terms as reported in Table 1. As expected, the most hydrophobic compounds showed a major role of vdW term with respect to the electrostatic one between the

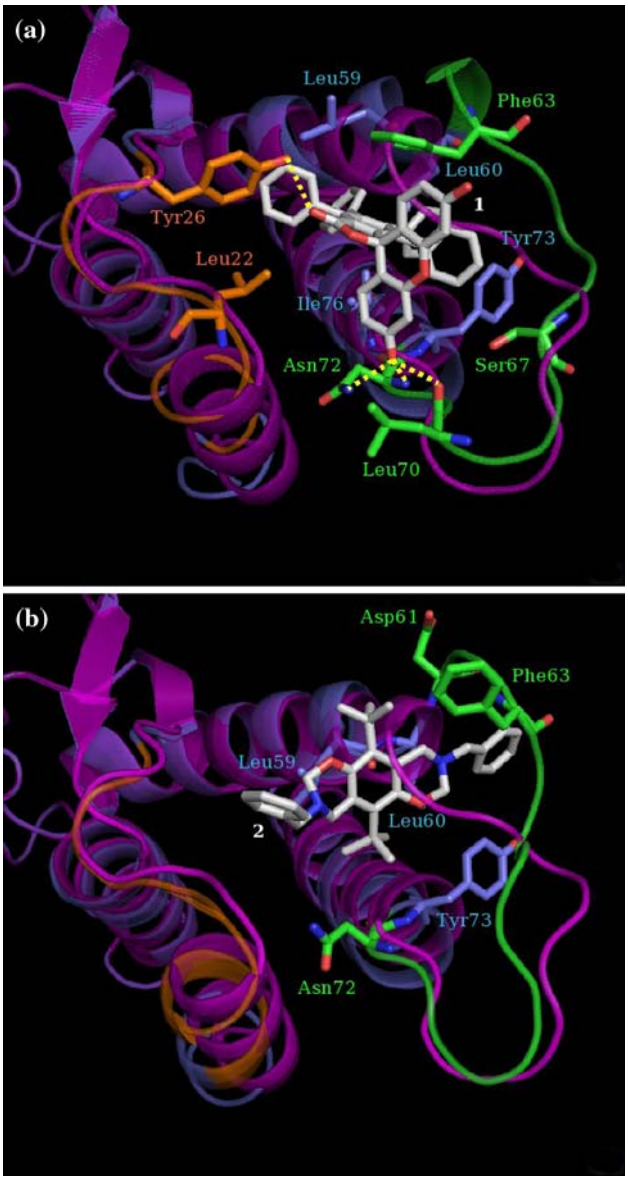


Fig. 5 Best poses of compounds **1** (*a*) and **2** (*b*). Loops *a* and *b* of SCF complexed and energy minimized with the ligands are, respectively, depicted in *orange* and *green ribbons*. Most relevant interacting residues and ligands are labeled and reported in polytube models. The crystallographic model of SCF is reported as *ribbon violet polytube*. Intermolecular hydrogen bonds are reported as *dotted yellow lines*

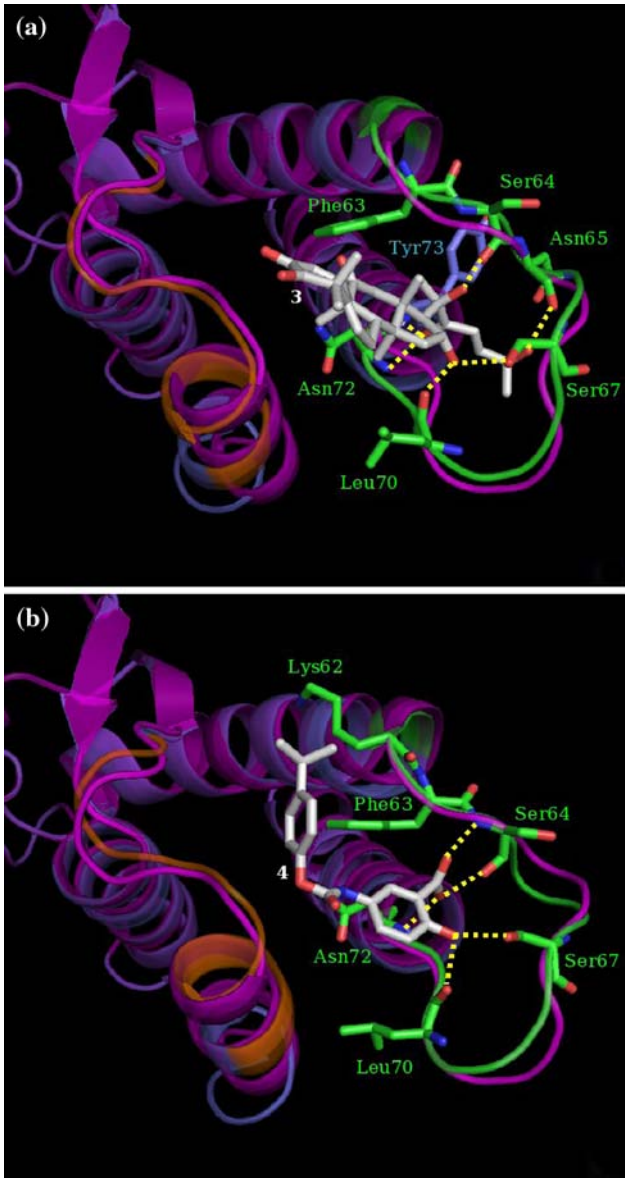


Fig. 6 Best poses of compounds **3** (*a*) and **4** (*b*). Loops *a* and *b* of SCF complexed and energy minimized with the ligands are, respectively, depicted in *orange* and *green ribbons*. Most relevant interacting residues and ligands are labeled and reported in polytube models. The crystallographic model of SCF is reported as *ribbon violet* polytube. Intermolecular hydrogen bonds are reported as *dotted yellow lines*

attractive interaction contributions. In compounds **3**, **4**, **5** and **10** the opposite trend was observed.

The visual analysis of the binding modes of the top 10 ranked compounds is reported in Figs. 5, 6, 7, 8, and 9.

The compound **1**, ranked as the best one, was characterized the presence of several phenyl rings and a spyrolactone moiety. This feature is responsible for a rigid T-like conformation delimited by the tricyclic xantene and the benzofurane rings. Despite the high level of hydrophobicity of this compound, the presence of two phenolic hydroxyl and the lactone oxygen atoms allowed to create multiple hydrogen bonds with the SCF receptor (Fig. 5a). The main interactions were established mainly with the loop *b* with two hydrogen bonds to Asn72 and one to Leu70. Hydrophobic contacts were detected to Phe63 (hindered stacking with the benzofurane ring) and to Tyr73 and Leu59, proximal residues to the same loop. Conversely with the loop *a* one hydrogen bond was found with the Tyr26 phenolyc OH and hydrophobic contacts to Leu22. Globally the compound **1** induced a consistent perturbation in the SCF dimeric surface,

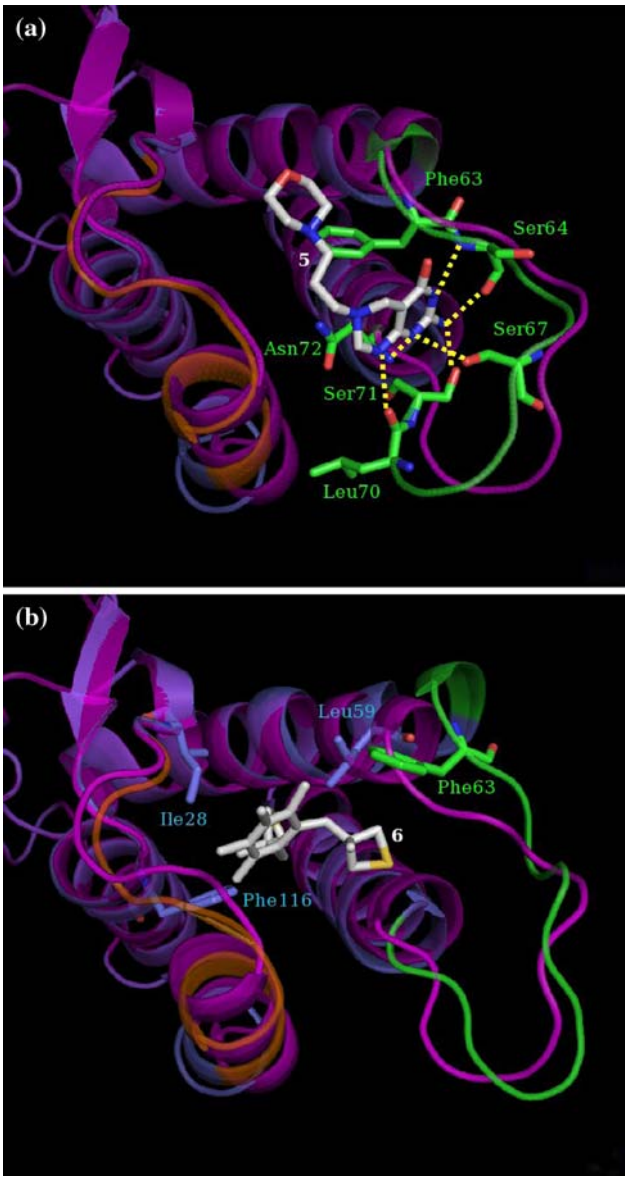


Fig. 7 Best poses of compounds **5** (*a*) and **6** (*b*). Loops *a* and *b* of SCF complexed and energy minimized with the ligands are, respectively, depicted in *orange* and *green ribbons*. Most relevant interacting residues and ligands are labeled and reported in polytube models. The crystallographic model of SCF is reported as *ribbon violet polytube*. Intermolecular hydrogen bonds are reported as *dotted yellow lines*

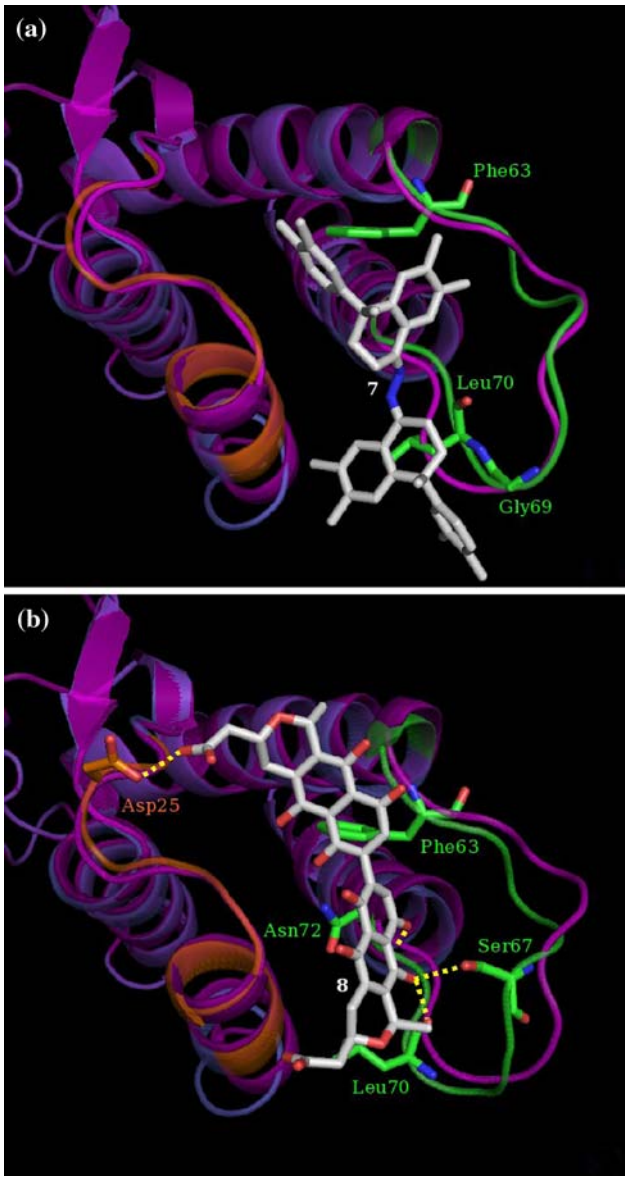


Fig. 8 Best poses of compounds **7** (*a*) and **8** (*b*). Loops *a* and *b* of SCF complexed and energy minimized with the ligands are, respectively, depicted in orange and green ribbons. Most relevant interacting residues and ligands are labeled and reported in polytube models. The crystallographic model of SCF is reported as *ribbon violet polytube*. Intermolecular hydrogen bonds are reported as *dotted yellow lines*

as demonstrated by the RMS deviation equal to 0.954 Å (Table 1).

The compound **2** is a symmetric benzoxazine derivative characterized by a binding mode consistently different from the previous ligand (Fig. 5b). Despite the presence of N and O heteroatoms no hydrogen bonds were formed with the SCF receptor. Moreover, the recognition was almost exclusively done with the loop *b* via hydrophobic contacts to Asn72, Phe63 and Asp61. Other proximal residues involved in the same kind of interaction were Tyr73, Leu60 and Leu59. Since the compound **2** docked SCF inserting one benzyl moiety into the loop *b* the effect in the RMS deviation (0.864 Å) was remarkable (Table 1).

The compound **3**, the natural anticancer agent cucurbitacin, has been recently reported as able to generically interfere with the tyrosin kinase signaling pathway [23]. The presence of polar moieties substituting a polycyclic rigid scaffold is responsible for the lower log *D* estimation in the range of the Lipinski's rules. Regarding the other three rules, this compound is relatively fine because the number of hydrogen

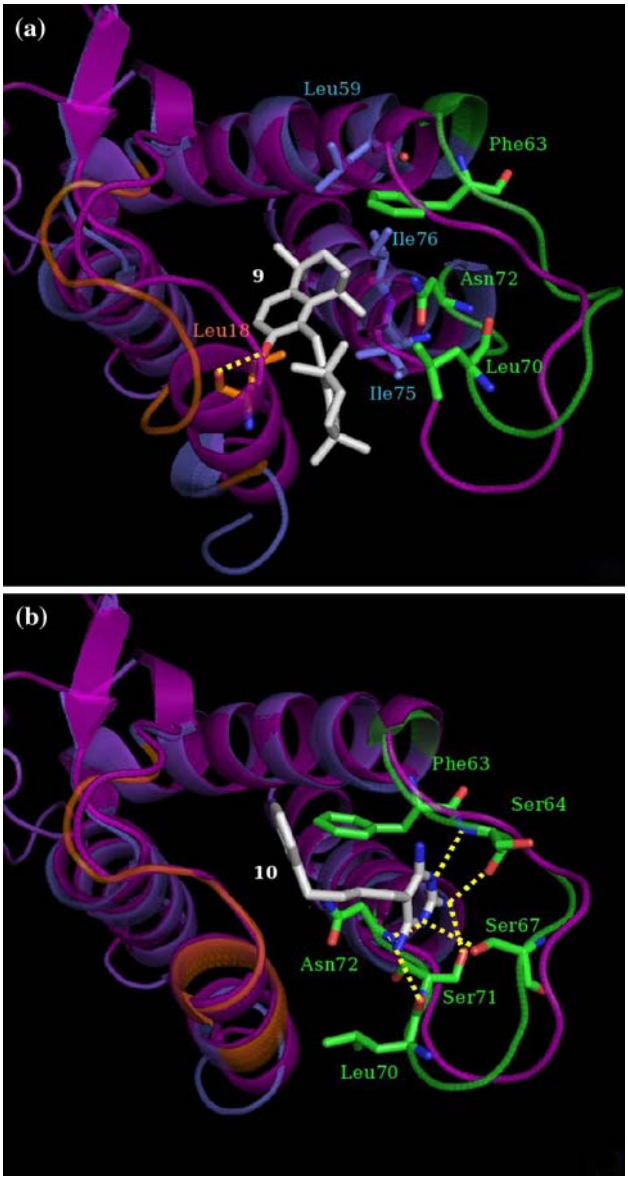


Fig. 9 Best poses of compounds **9** (*a*) and **10** (*b*). Loops *a* and *b* of SCF complexed and energy minimized with the ligands are, respectively, depicted in *orange* and *green ribbons*. Most relevant interacting residues and ligands are labeled and reported in polytube models. The crystallographic model of SCF is reported as *ribbon violet polytube*. Intermolecular hydrogen bonds are reported as *dotted yellow lines*

bond donors and acceptors is lower than the limit of 5 and 10. Only its molecular weight, equal to 514.29, is slightly higher than the limit of 500. Such a molecule mainly recognized the SCF loop *b* (Fig. 6a) through hydrogen bonds to Ser64, Asn65, Ser67, Leu70, Asn72 and Tyr63. Interestingly these contributions always involved the aminoacid backbones. Compound **3** reported hydrophobic contacts to Phe63 and its target recognition produced a relatively modest perturbation of the SCF structure, as indicated by the RMS deviation equal to 0.563 Å and mainly addressable to the loop *b* interaction.

Compound **4** is a salicylic carboxamide derivative showing a similar best pose with respect to **3** (Fig. 6b). Actually the SCF loop *b* was the main docking involved area. Five hydrogen bonds were reported to Ser64 (twice), Ser67, Leu70 and Asn72. Relevant hydrophobic contacts were highlighted with respect to Lys62 and Phe63 side-chains. Induced fit modification of the receptor structure and the RMS deviation (0.578 Å) resulted similar to those reported for compound **3**. Moreover, this compound fully satisfied the Lipinski's rules, i.e., the lipophilicity expressed as log *D* is lower than 5 (Table 1), the molecular weight equal to 329.35 is lower than 500, the 3 hydrogen bond donors and 5 acceptors are, respectively, lower than the limits of 5 and 10.

Compound **5** is a tetrahydropurine with a morpholino side-chain. It interacted mainly with loop *b*, in particular with residues 67–72 (Fig. 7a). The effect on the RMS deviation was relatively modest (0.513 Å) despite the consistent hydrogen bond network established with residues Ser64, Ser67, Leu70, Ser71 and Asn72. In particular, the tetrahydropurine was located between Ser67 and Asn72 creating a bridge which strongly perturbed the loop *b* conformation. The compound **5** best pose was, moreover, stabilized by hydrophobic contacts between the morpholino ring and the Phe63 side-chain. Also this compound satisfied the Lipinski's rules, i.e., the lipophilicity expressed as log *D* is lower than 5 (Table 1), the molecular weight equal to 294.35 is lower than 500, the 3 hydrogen bond donors and 7 acceptors are, respectively, lower than the limits of 5 and 10.

Compound **6**, a sulfured tetramethylbenzene, showed a different binding mode to SCF. Actually, the ligand in best docking pose lies in the middle of the dimerization surface interacting with low solvent exposed aminoacids, such as Ile28, Phe116 and Leu59, and with Phe63 located into the loop *b* (Fig. 7b). Due to its simple structure, **6** could not produce hydrogen bonds and its recognition was stabilized only by remarkable hydrophobic contributions inducing relevant modifications of both loop *a* and *b* structures, as highlighted by the SCF RMSD equal to 0.855 Å.

Compound **7** is the most hydrophobic molecule of the top 10 list (Table 1). It is based on an hindering bis dihydronaphthalen hydrazone chemical scaffold without hydrophilic substituents. Its SCF recognition was driven by vdW contacts to loop *b* residues such as Phe63, Gly69 and Leu70. No hydrogen bonds were observed and the SCF inducing fit effect was the lowest found (RMSD equal to 0.482 Å), almost exclusively concentrated in the loop *b* (Fig. 8a).

Compound **8** is the natural antibiotic actinorhodin [24] showing a bis anthraquinonic chemical scaffold. As opposed to the previous ligand, it was the most hydrophilic structure among our virtual screening selected hits (Table 1). Such a ligand demonstrated the capability of interacting with both SCF loop *a* and *b* but, again, it could induce relevant conformational changes only onto the last one (Fig. 8b). Due to the

presence of several oxygen atoms, its recognition was stabilized by four hydrogen bonds, respectively, to Asp25, Ser67, Leu70 and Asn72. Hydrophobic contacts were reported to Phe63, Leu70 and Asn72 side-chains. Even if the lipohilicity resulted lower than 5 (Table 1), the molecular weight (634.54), the 6 hydrogen bond donors and 14 acceptors are over the accepted limits for the Lipinski's rules.

Compound **9** has a phenolic bis tetraline structure. This ligand showed an SCF binding mode similar to that reported for compound **6**. It was located in the middle of the dimerization area recognizing low solvent exposed hydrophobic residues, such as Ile59, Ile75 and Ile76 (Fig. 9a). Compound **9** produced the largest modification of SCF binding site conformation (SCF RMSD equal to 1.143 Å) and it was able to interact with both loop *a* and *b*, respectively, by hydrogen bond to Leu18 backbone and vdW contacts to Phe63, Leu70 and Asn72 side-chains.

Compound **10** is an aminopyrimidine derivative mainly interacting with the loop *b* (Fig. 9b). Six hydrogen bonds were observed to the backbone of Phe63, Ser64, Leu70 and Asn72 and to the side-chains of Ser67 and Ser71. Moreover, Phe63 and Asn72 were also involved in hydrophobic interactions with the ligand phenylbutyl moiety. The low SCF dimerization area perturbation (RMSD equal to 0.459 Å) can be mainly addressed to the loop *b* conformational changes. Finally this compound also satisfied the Lipinski's rules, i.e., the lipophilicity expressed as log *D* is lower than 5 (Table 1), the molecular weight equal to 257.33 is lower than 500, the 3 hydrogen bond donors and 5 acceptors are, respectively, lower than the limits of 5 and 10.

4 Conclusions

The analysis of the SCF dimerization area, performed by superimposition of the monomer crystal structures and by MD simulation, revealed the loop *b* as the most flexible portion. The virtual screening study, carried out with the NCI molecular diversity set, confirmed that all 10 top scored compounds interacted in the dimerization area mainly with the loop *b*. As expected the nature of the compounds identified as potentially best SCF ligands is diverse in terms of chemical scaffolds, flexibility and lipophilicity. This issue is probably the most critical aspect to be considered for potentially druggable compounds of this series. Four of them (**3**, **4**, **5** and **10**) passed the Lipinski's rules and it is interesting to note that the first compound of this subset is the cucurbitacin **3**, recently reported as an active anticancer agent involved in the tyrosine kinase pathway.

The top scored compounds are now under biological investigations with the aim to confirm their ability to inhibit the proliferation of cell lines significantly expressing the c-kit receptor and to check the SCF dimerization interference as the main mechanism of action. The information of the computational work will be useful for the rational design of novel SCF dimerization inhibitors.

References

1. Galli SJ, Zsebo KM, Geissler EN (1994) Adv Immunol 55:1
2. Hunter T (2000) Cell 100:113
3. Zhang Z, Zhang R, Joachimiak A, Schlessinger J, Kong XP (2000) Proc Natl Acad Sci USA 97:7732
4. Jiang X, Gurel O, Mendiaz EA, Stearns GW, Clogston CL, Lu HS, Osslund TD, Syed RS, Langley KE, Hendrickson WA (2000) EMBO J 19:3192
5. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28:235
6. Fletcher S, Hamilton AD (2006) J R Soc Interface 3:215
7. http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html
8. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE III, DeBolt S, Ferguson DC, Seibel G, Kollman PA (1995) Comp Phys Commun 91:1
9. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) J Comput Aided Mol Des 11:425
10. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225
11. Schrödinger Inc., Portland, OR, USA, 1998–2001
12. Mohamadi F, Richard NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, Chang G, Hendrickson T, Still WC (1990) J Comput Chem 11:440
13. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112:6127
14. Alcaro S, Gasparrini F, Incani O, Mecucci S, Misiti D, Pierini M, Villani C (2000) J Comp Chem 21:515
15. MarvinSketch 4.1.5 - logD Plugin; ChemAxon Kft. Máramaros köz 3/a, Budapest, 1037 Hungary, http://www.chemaxon.com/marvin
16. Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) J Chem Inf Comput Sci 29:163
17. Csizmadia F, Tsantili-Kakoulidou A, Panderi I, Darvas F (1997) J Pharm Sci 86:865
18. Bouchard G, Carrupt PA, Testa B, Gobry V, Girault HH (2001) Pharm Res 18:702
19. DeLano Scientific, San Carlos, CA, USA, 2006
20. Smith CD, Glickman JF, Chang KJ (1988) Biochem Biophys Res Commun 156:1250
21. Roskoski R Jr (2003) Biochem Biophys Res Commun 309:709
22. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Adv Drug Delivery Rev 46:3
23. Sun J, Blaskovich MA, Jove R, Livingston SK, Coppola D, Sebti SM (2005) Oncogene 24:3236
24. Wright LF, Hopwood DA (1976) J Gen Microbiol 96:289