Computational Studies of Epidermal Growth Factor Receptor: Docking Reliability, Three-Dimensional Quantitative Structure−Activity Relationship Analysis, and Virtual Screening Studies

Concettina La Motta, Stefania Sartini, Tiziano Tuccinardi,* Erika Nerini, Federico Da Settimo, and Adriano Martinelli

Dipartimento di Scienze Farmaceutiche, Università di Pisa, Via Bonanno 6, 56126 Pisa, Italy

Received July 8, 2008

An aberrant activity of the epidermal growth factor receptor (EGFR) has been shown to be related to many human cancers, such as breast and liver cancers, thus making EGFR an attractive target for antitumor drug discovery. In this study we evaluated the reliability of various kinds of docking software and procedures to predict the binding disposition of EGFR inhibitors. By application of the best procedure and use of more than 200 compounds, a receptor-based 3D-QSAR model for EGFR inhibition was developed. On the basis of the results obtained, the possibility of developing virtual screening studies was also evaluated. The VS procedure that proved to be the most reliable from a computational point of view was then used to filter the Maybridge database in order to identify new EGFR inhibitors. Enzymatic assays revealed that among the eight top-scoring compounds, seven proved to inhibit EGFR activity at a concentration of 100 μM, two of them exhibiting IC_{50} values in the low micromolar range and one in the nanomolar range. These results demonstrate the validity of the methodologies followed. Furthermore, the two low micromolar compounds may be considered as very interesting leads for the development of new EGFR inhibitors.

Introduction

Protein tyrosine kinases (PTKs[a]) represent a wide family of homologous enzymes, both transmembranous and cytoplasmic, that catalyze the transfer of the γ-phosphate group from ATP to a hydroxy group of selected tyrosine residues in target protein substrates. Under physiological conditions, tyrosine phosphorylation represents a fundamental signal transduction mechanism that proceeds in a hierarchically ordered sequence of protein−protein interactions, ensuring cross-talk between cells and regulating key aspects of cell life such as proliferation, differentiation, metabolism, and apoptosis.

Thus, an aberrant tyrosine kinase (TK) activity, leading to malfunctioning signal transduction pathways, gives rise to the development of numerous proliferative diseases, both malignant and nonmalignant, like cancers, atherosclerosis, psoriasis, and restenosis, as well as to the progression of a large number of inflammatory and autoimmune disorders. Actually, under specific conditions such as overexpression, activating point mutations, and loss of both inhibitory domains and negative regulators, PTKs can assume a hyperactive status, thus causing uncontrolled cell proliferation. That is why this class of proteins, with the signal pathways that they regulate, represents a highly attractive and widely explored target for drug development.[1−6]

Among the known PTKs, the most exhaustively discussed is undoubtedly the family of the epidermal growth factor receptor (EGFR or ErbB), which consists of four closely related members: EGFR (ErbB1 or HER1), ErbB2 (HER-2 or Neu), ErbB3 (HER-3), and ErbB4 (HER-4). They all share a common architecture, being made up of an extracellular domain, responsible for growth factor recognition and binding, a helical transmembrane segment, and an intracellular domain, where the tyrosine kinase activity resides. Different degrees of homology, regarding both the extracellular and the tyrosine kinase domains, accomplish signal specificity within the whole family. Different from the other ErbB members, the ErbB3 subtype lacks an intrinsic protein tyrosine kinase activity. Actually, crucial amino acid residues highly conserved throughout the protein kinases are altered in the catalytic domain of this receptor. Therefore, it works as a signaling entity only after heterodimerization with another ErbB receptor.[7]

Upon binding of peculiar growth factors, mainly epidermal growth factor, transforming growth factor-α (TGFα), amphiregulin (AR), epiregulin (EPR), neuregulins (NRG), and cytokines, to the extracellular domain of the inactive receptor, a dimerization of the ligand−receptor complex occurs. Through formation of either a homodimer or a heterodimer, with another member of the HER family, the receptor assumes an active form, thus triggering the autophosphorylation of the intracellular tyrosine domain. Phosphotyrosine residues become, in turn, attachment sites for downstream signaling proteins that, through additional phosphorylations, allow signal transduction from the cell surface to the nucleus, where it results in the intended effect.

ErbB receptors are expressed in a variety of tissues of epithelial, mesenchymal, and neuronal origin, where they play pivotal roles in development, proliferation, and differentiation.

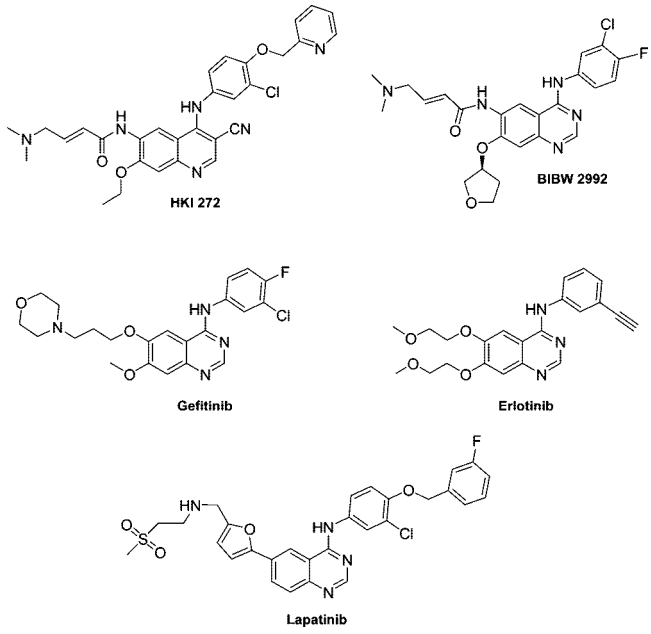
A deregulated activity of these kinds of receptors is the leading cause of a variety of solid human cancers, including non-small-cell lung, breast, bladder, ovarian, and head and neck carcinomas. Furthermore, as the ErbB family is involved not only in cell proliferation but also in a number of side processes regulating tumor progression, such as cell motility, cell adhesion, and angiogenesis, overexpression or deregulation of its activity

* To whom correspondence should be addressed. Phone: ++39 050 2219595. Fax: ++39 050 2219605. E-mail: tuccinardi@farm.unipi.it.

[a] Abbreviations: EGFR, epidermal growth factor receptor; VS, virtual screening; PTKs, protein tyrosine kinases; TK, tyrosine kinase; TGFα, growth factor-α; AR, amphiregulin; EPR, epiregulin; NRG, neuregulins; CS, conformational search; rmsd, root-mean-square deviation; KCS, kinase scoring function; SDEP, standard deviation of errors of prediction; EF, enrichment factor; PR, Precision·Recall; SRD, smart region definition; FFD, fractional factorial design.

10.1021/jm800829v CCC: $40.75
Published on Web 01/26/2009

Chart 1. EGFR Inhibitors



is generally associated with the malignancy of the disease and a poor prognosis.[8−10]

These receptors thus represent an excellent drug target, and the development of specific inhibitors has proved to be a promising therapeutic concept. Actually, patients with cancer can significantly profit from the development of this targeted therapy, especially when the activated ErbB pathway is shown to be the underlying cause of a given malignancy. Consequently, the search for specific inhibitors of the ErbB family is increasingly becoming a major pharmaceutical challenge. Today, most of the inhibitors under development are ATP-site directed compounds targeting mainly the EGFR and the ErbB-2 subtypes. Many structurally different small molecules have been described as potent reversible or irreversible inhibitors of these enzymes,[11−13] all showing various degrees of efficacy and selectivity against small panels of tyrosine kinases. Several of them, like (*E*)-*N*-{4-[3-chloro-4-(2-pyridinylmethoxy)anilino]-3-cyano-7-ethoxyquinolin-6-yl}-4-(dimethylamino)-2-butenamide[14] (HKI 272, Wyeth) and *N*-[4-(3-chloro-4-fluorophenyl)amino]-7-[[(3*S*)-tetrahydro-3-furanyl]oxy]-6-quinazolinyl]-4-(dimethylamino)-2-butenamide[15,16] (BIBW 2992, Boehringer Ingelheim), are being used in clinical studies and show encouraging anticancer activity. A few examples, like gefitinib,[17] erlotinib[18] and lapatinib[19] are currently marketed for the treatment of certain types of solid tumors sustained by mutant or overexpressed ErbB1 or ErbB2, such as breast, lung, and colon-rectal cancer (Chart 1).

However, the number of effective EGFR inhibitors is constantly expanding. Different methods are followed to obtain novel compounds, but the most fruitful one is undoubtedly computer screening.

Many docking,[20−23] QSAR,[24−26] and 3D-QSAR studies[27−33] of EGFR ligands have been widely published over the past 10 years, and by use of the docking approach, successful virtual screening (VS) studies have been developed.[20−23] From these studies came out that the utilization of a docking procedure able to predict with a low degree of error the binding pose of newly designed inhibitors, together with the possibility of in silico predicting their EGFR inhibition activity, should indeed favor the design of new compounds. As shown in Figure 1, in the present study, we evaluated the reliability of various kind of docking software and procedures to predict the binding disposition of EGFR inhibitors, and after using more than 200 compounds, a receptor- and a ligand-based 3D-QSAR model of EGFR was developed. Finally, we evaluated the possibility of using a mixed ligand/receptor-based approach for developing a VS method capable of finding new EGFR inhibitors.

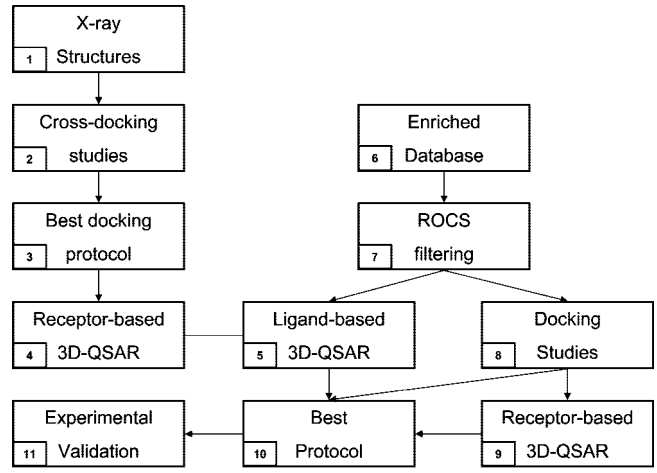


Figure 1. Schematic diagram of the computational workflow.

Results and Discussion

The first step of the study concerned the evaluation of the reliability of various kind of docking software and procedures for the prediction of the binding disposition of EGFR inhibitors into the crystal structure of the protein (see the second point of Figure 1). At present, 7 wild-type and 8 mutated EGFR complexed with inhibitors are deposited in the Protein Data Bank for a total of 15 EGFR TK complexes. In order to extend the analysis to a larger number of TK−inhibitor complexes, using the Blastp server, a sequence similarity search was carried out so that the docking analysis was also performed on the 27 TK−inhibitor complexes with the highest percentage of sequence homology with respect to the EGFR catalytic domain.

Therefore, a total of 42 kinase inhibitors (15 of them were EGFR inhibitors) were docked into their crystal structures through Autodock 3.0,[34] Autodock 4.0,[34,35] Gold 3.0.1,[36] Dock 6.0,[37] and Fred 2.1[38] software after an extensive conformational search (CS) and using a total of 14 procedures. The docking results were evaluated through a comparison of the found docked positions of the ligand and the experimental ones: as a measure of docking reliability, the root-mean-square deviation (rmsd) between the positions of heavy atoms of the ligand in the calculated and experimental structures was taken into account.

As reported in Figure 2, using the average rmsd obtained for each docking procedure as a measure of the reliability of the methods applied, the Gold software seemed to be the most suitable, since it showed the lowest average rmsd value. In particular, the use of ChemScore fitness function modified by the kinase scoring function (KCS) gave the best results (rmsd = 2.7 Å) and was the only method that gave an average rmsd lower than 3.0 Å.

Automated Docking and 3D-QSAR Analysis. The above-described results indicated that among all the procedures, the use of ChemScore fitness function modified by the kinase scoring function of the Gold software was the one that showed the most reliable ligand disposition inside the used TKs.

By use of this procedure, 206 EGFR inhibitors[39−49] (see Table 1 in the Supporting Information) were docked into the binding site of EGFR (1XKK[50] PDB code). These compounds were tested for EGFR inhibitory activity in substrate phosphorylation

RMSD < 1.5 | 1.5 < RMSD < 2.5 | 2.5 < RMSD < 4.0 | RMSD > 4.0

Protein / Ligand

TK Type	Align Score	PDB Code	AUTODOCK 3.0 Gasteiger charge First Cluster	AUTODOCK 3.0 Gasteiger charge Most pop. Cluster	AUTODOCK 3.0 AM1-bcc charge First Cluster	AUTODOCK 3.0 AM1-bcc charge Most pop. Cluster	DOCK 6 RMSD	GOLD GS RMSD	GOLD CS RMSD	GOLD KCS RMSD	AUTODOCK 4.0 Gasteiger charge First Cluster	AUTODOCK 4.0 Gasteiger charge Most pop. Cluster	AUTODOCK 4.0 AM1-bcc charge First Cluster	AUTODOCK 4.0 AM1-bcc charge Most pop. Cluster	FRED Cgauss2 RMSD	FRED Cscore RMSD	FRED PLP RMSD	FRED Sgauss RMSD	FRED Sscore RMSD	FRED Global RMSD
EGFR	/	1XKK	11.9	11.8	10.0	10.0	8.0	3.2	4.5	3.5	3.3	3.0	3.0	1.3	3.7	4.9	10.2	3.6	4.1	3.7
EGFR	/	1M17	2.1	2.0	5.0	5.0	6.4	6.0	1.8	2.1	4.9	4.9	4.8	8.8	1.9	2.3	8.0	1.7	8.0	1.1
M_EGFR	/	2GS7	4.8	4.9	8.9	8.9	10.1	3.4	3.5	3.0	3.8	2.6	6.9	5.2	4.0	5.4	2.3	8.5	4.3	4.3
EGFR	/	2ITY	4.0	4.0	8.1	8.1	5.0	7.3	6.2	3.5	7.2	7.2	6.6	6.6	7.3	4.0	7.4	7.6	7.3	4.0
M_EGFR	/	2ITZ	3.9	8.9	3.8	8.7	1.7	2.4	3.6	3.4	8.3	8.2	6.5	6.5	7.5	2.7	6.1	8.0	5.5	6.1
M_EGFR	/	2ITO	4.8	4.8	3.9	9.0	5.8	8.1	4.9	8.3	8.7	8.7	9.4	8.7	9.3	6.2	6.1	5.7	9.8	5.6
EGFR	/	2J6M	2.0	3.2	2.0	3.6	1.3	1.9	1.8	1.7	1.5	1.4	1.5	10.4	1.6	11.0	9.7	1.6	9.4	1.6
M_EGFR	/	2ITT	1.7	1.7	1.7	4.1	1.9	2.1	2.7	1.6	1.3	11.0	2.3	10.6	9.3	1.9	10.3	1.7	9.3	9.3
M_EGFR	/	2ITP	1.6	1.6	2.8	3.0	3.1	1.3	11.1	1.7	1.4	2.1	10.7	10.7	1.4	1.4	11.1	1.3	12.8	11.1
EGFR	/	2ITW	6.3	5.2	6.6	4.8	2.6	5.3	3.7	3.8	5.8	4.3	6.1	4.9	6.0	5.8	6.3	6.3	6.3	6.3
M_EGFR	/	2ITU	6.2	4.2	6.2	4.2	3.1	4.0	3.8	3.7	5.9	3.7	5.8	4.0	2.8	6.1	6.8	6.7	6.8	6.7
M_EGFR	/	2ITQ	1.4	6.1	1.4	1.4	0.9	1.4	1.3	1.2	1.3	1.3	1.4	1.4	1.5	5.8	1.5	5.3	1.5	1.5
EGFR	/	2ITX	5.2	5.0	5.3	5.3	4.3	6.0	5.8	5.6	9.1	4.9	10.0	10.0	3.5	4.5	4.2	6.0	2.9	4.6
M_EGFR	/	2ITV	9.7	9.7	9.6	9.6	4.9	4.9	5.2	4.1	9.9	9.9	11.5	11.5	3.4	4.0	3.4	6.6	4.8	4.8
M_EGFR	/	2ITN	1.4	6.0	7.8	6.2	7.2	4.7	4.9	2.8	10.2	10.2	10.7	10.1	4.7	4.6	2.5	5.9	4.5	4.7
ACK1	212	1U54	7.5	5.5	4.5	5.0	4.3	5.6	5.5	4.0	6.4	4.4	8.2	7.3	4.9	3.8	4.7	5.9	4.9	5.2
KSYK	203	1XBC	1.4	1.4	1.2	1.2	0.9	1.1	0.6	0.7	1.2	1.2	1.2	1.2	1.3	5.4	1.3	1.3	1.3	1.3
ECK TK	201	1MQB	8.0	8.3	7.5	7.3	11.9	9.7	3.8	5.1	7.5	7.5	8.1	9.7	2.4	4.0	2.5	8.4	3.4	2.8
FAK1	200	1MP8	5.5	5.6	2.2	2.2	2.1	2.6	2.6	1.5	2.1	2.1	7.1	7.1	2.7	2.6	2.5	3.8	4.0	2.7
ZAP-70	195	1U59	1.3	5.3	1.3	1.3	1.0	1.2	0.6	0.6	1.2	1.2	1.1	1.1	1.3	1.3	1.3	1.3	1.3	1.3
ABL1	195	1FPU	11.5	6.8	11.5	6.8	9.6	1.1	1.2	1.3	11.5	11.5	11.7	3.7	1.8	1.8	1.8	1.7	1.8	1.8
ABL1	193	1OPK	1.7	1.7	1.6	1.6	1.4	1.7	2.4	2.4	1.5	1.5	1.5	1.5	0.7	2.3	0.7	0.7	2.3	0.7
FGFR1	174	2FGI	2.3	10.8	3.7	10.5	6.1	1.8	2.8	3.3	2.9	1.6	1.6	2.3	2.0	10.6	2.0	3.3	2.1	3.3
JAK3	173	1YVJ	1.4	1.4	1.5	1.5	1.2	1.2	0.6	0.5	1.4	1.4	1.3	5.4	3.4	3.4	3.0	1.4	3.0	3.4
FGFR2	172	1OEC	3.2	3.2	3.1	3.1	4.7	2.3	3.0	3.0	2.0	2.0	1.8	4.8	6.7	7.1	7.1	9.1	6.5	6.5
JAK2	162	2B7A	0.7	0.7	0.7	4.0	0.4	0.9	0.5	1.1	0.7	0.7	0.6	0.6	1.7	6.0	5.5	1.1	4.3	5.5
CSK	162	1BYG	1.3	0.6	0.6	0.6	0.3	0.6	0.9	0.7	0.3	0.3	0.4	0.4	1.9	3.8	3.8	7.8	2.0	3.8
VGFR2	157	1YWN	0.8	0.8	0.8	0.8	0.4	1.1	0.7	0.8	0.6	0.6	0.7	0.7	0.6	0.6	0.6	10.5	0.6	0.6
HCK	157	1QCF	5.0	0.5	5.0	0.9	7.3	1.1	1.0	1.4	0.6	0.6	0.5	0.5	1.0	1.0	1.0	4.0	1.0	1.0
SRC	157	1Y57	3.3	3.3	3.3	3.3	4.2	3.5	3.3	3.4	3.5	3.5	3.2	3.2	6.2	5.0	4.6	5.3	5.3	5.3
SRC	157	2SRC	4.9	4.9	3.1	3.0	2.9	4.8	4.7	5.0	4.4	4.4	6.1	3.3	3.3	5.3	2.2	8.8	6.5	3.7
SRC	156	1YOL	1.9	8.7	1.9	1.9	4.5	2.5	2.9	2.2	1.9	1.9	10.2	1.9	4.7	9.9	8.6	11.6	8.6	8.6
INSR	155	1I44	4.8	4.8	3.0	4.7	2.3	7.6	3.0	2.5	3.3	3.3	7.7	7.7	3.5	4.2	2.6	5.7	7.5	3.1
SRC	154	1YOM	1.0	6.6	1.4	6.4	1.7	1.8	6.2	6.3	1.6	1.6	6.6	6.6	2.2	2.4	1.9	1.8	1.8	1.5
IGF1R	154	1JQH	8.2	8.2	2.9	5.3	4.7	6.4	5.0	4.8	3.0	2.6	6.0	6.7	6.3	2.0	2.3	9.2	8.5	2.9
SRC	154	1KSW	4.4	4.4	4.7	5.5	1.5	4.4	4.2	4.1	2.5	6.4	2.8	2.8	4.1	3.2	3.3	5.6	3.6	4.2
LCK	154	1QPJ	1.1	1.1	1.0	1.0	0.2	0.4	1.0	1.0	0.9	0.9	0.6	0.6	2.0	2.8	2.1	2.0	2.1	2.0
KIT	153	1T46	12.5	6.4	12.5	6.3	9.4	1.3	1.5	1.4	1.3	1.3	12.7	6.2	1.6	1.6	1.6	12.6	12.7	12.0
LCK	150	1QPD	1.9	1.9	1.9	0.6	0.5	0.5	0.5	0.6	2.3	2.3	2.3	0.6	2.9	2.9	2.5	2.5	2.5	2.5
HCK	148	2HCK	6.6	2.5	6.5	2.3	2.4	6.9	2.9	2.7	6.9	2.6	2.7	2.7	2.6	7.2	3.4	7.0	6.2	3.7
VGFR2	139	1Y6B	5.7	7.6	7.7	5.7	4.2	1.1	5.0	1.4	4.5	4.5	7.1	7.1	6.1	8.8	9.0	7.5	9.0	9.0
PDPK1	74	2BIY	3.9	3.9	4.0	4.0	11.0	2.4	1.6	1.4	10.0	1.6	10.0	10.0	2.7	9.6	2.5	4.4	2.6	4.7
AVERAGE RMSD			4.3	4.7	4.3	4.5	4.0	3.3	3.2	2.7	4.0	3.7	5.3	5.2	3.5	4.5	4.3	5.3	5.1	4.2
STANDARD DEVIATION			3.1	2.9	3.1	2.8	3.1	2.4	2.1	1.7	3.2	3.1	3.7	3.5	2.2	2.6	2.9	3.1	3.1	2.7

Figure 2. Matrix of rmsds obtained by the docking studies. The second column indicates the score of sequence similarity between the corresponding TK and human EGFR.

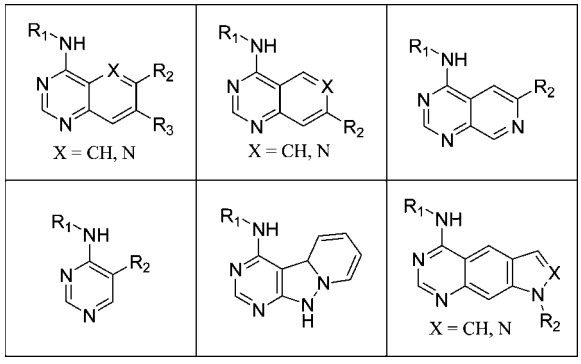


Figure 3. General structures of the EGFR inhibitors used for the development of the 3D-QSAR models.

assays, and their inhibitory activities spanned about 5 orders of magnitude (the inhibition data ranged from 0.1 nM to 10 μM, expressed as IC_{50}). As shown in Figure 3, they were characterized by six different central scaffolds.

The docking results were then scored by means of the GoldScore, ChemScore, and ChemScore fitness function modified by the kinase scoring function, but the correlation between the calculated free energy of binding and the experimental one was not good, since the quadratic correlation (R^2) showed low values ($R^2 < 0.4$). Therefore, with the aim of obtaining quantitative results, the docking poses were used as a receptor-based alignment for a 3D-QSAR study (see point 4 of Figure 1).

The reliability of the 3D-QSAR model was characterized by its correlation coefficient (r^2), predictive correlation coefficient (q^2), and cross-validated standard deviation of errors of prediction ($SDEP_{CV}$).

Compounds of the original set were divided into a training set and a test set in accordance with the usual guidelines. In

Table 1. Statistical Results of the 3D-QSAR Model

Vars	PC	r^2	q^2	$SDEP_{CV}$	$SDEP_{TS}$
1576	2	0.81	0.71	0.56	0.67
1576	3	0.87	0.76	0.51	0.59
1576	4	0.91	0.76	0.51	0.63

detail, both the training and the test sets should contain compounds in such a way as to maximize their structural diversity (that is, compounds belonging to the training set and to the test set should be representative of the molecular diversity of all the compounds under study) and uniformly span over the whole range of activity. As a result, the training set was made up of 142 compounds, while an external validation set (test set) of 64 molecules was selected to test the predictive power of the model.

Table 1 presents the main data of the 3D-QSAR analysis. The second PLS component accounted for 81% of variance and was predictive ($q^2 = 0.71$, SDEP of the external test set = 0.67). The third PLS component improved both the fitting and the predictive ability of the model ($r^2 = 0.87$, $q^2 = 0.76$, and the SDEP of the test set = 0.59), whereas the fourth PLS component did not offer any further predictivity improvement. Thus, the model's optimal dimensionality was given by three components (see Figure 4).

One important feature of 3D-QSAR analysis is the graphic representation of the model, usually used to make its interpretation easier. In the Golpe program[51] there are several options for displaying the final model. Among these, the PLS pseudocoefficient and the activity contribution plots are very useful. The PLS coefficient plot makes it possible to visualize favorable and unfavorable interactions between the probes and the molecules under study, while the activity contribution plot is

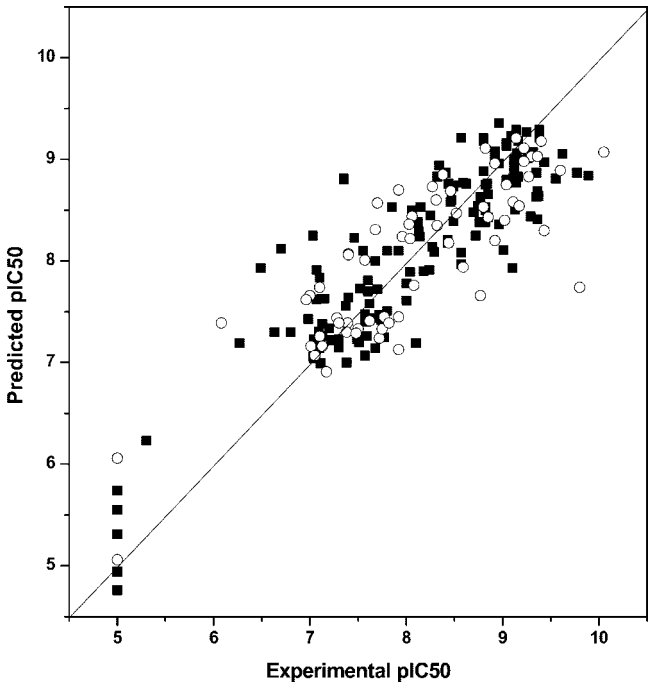


Figure 4. Plot of the receptor-based 3D-QSAR model. Experimental/predicted pIC_{50} is reported. The training set is represented as ■ and the test set as ○.

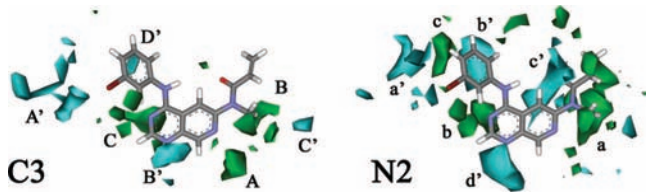


Figure 5. Negative (cyan) and positive (green) regions of the PLS coefficient plot obtained with the C3 and N2 probes.

different for each molecule within the training set, displaying spatial regions that are individually important for the selected molecule.

Figure 5 illustrates the PLS coefficient plots of the model for the C3 and N2 probes. A high affinity 4-anilinopyrido[3,4-*d*]pyrimidine (compound **107** of Table 1 in the Supporting Information, IC_{50} = 0.17 nM) is also reported.

For the C3 probe, there are three regions (A−C) with positive values (green polyhedrons) in which a favorable interaction between a substituent and the probe results in an increase in activity, whereas an unfavorable interaction between a substituent and the probe results in a decrease in activity. There are also four negative regions (A′−D′, cyan polyhedrons) in which a favorable interaction between a substituent and the probe results in a decrease in activity, whereas an unfavorable interaction between a substituent and the probe results in an increase in activity. As shown in Figure 6, regarding the N2 probe, there are four principal regions (a′−d′) with negative values (cyan polyhedrons) in which a favorable interaction between a substituent and the probe results in an increase in activity, whereas an unfavorable interaction between a substituent and the probe results in a decrease in activity, and there are three positive regions (green polyhedrons) where a favorable interaction between a substituent and the probe determines a decrease in activity.

It is also interesting to note that some of the positive and negative PLS regions for the N2 probe corresponded to those

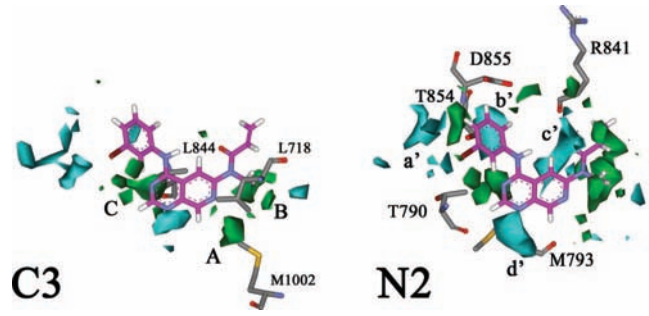


Figure 6. Negative (cyan) and positive (green) regions of the PLS coefficient plot obtained with the C3 and N2 probes superimposed on the human EGFR binding site.

Table 2. Statistical Results of the 3D-QSAR Model

Vars	PC	r^2	q^2	$SDEP_{CV}$	$SDEP_{TS}$
1427	2	0.82	0.68	0.59	0.75
1427	3	0.86	0.74	0.53	0.75
1427	4	0.89	0.74	0.53	0.76

of C3. In particular two positive regions (a, b) for the N2 probe corresponded to the positive regions B and C for the C3 probe and three negative regions (a′, b′, and d′) corresponded to the A′, D′, and B′ negative regions for the C3 probe.

As the alignment of the ligands was performed using the structures docked into the X-ray EGFR structure, it would be useful to check for matching between the EGFR protein and the 3D-QSAR maps. In Figure 6, the binding site of the EGFR overlaps with the positive 3D PLS coefficient maps of the C3 probe and negative maps of the N2 probe. For the C3 probe, regions A, B, and C are in the proximity to M1002, L718, and L844, respectively, thus suggesting an important role for these residues. For the N2 probe, the electrostatic surface a′ corresponds to T790, region b′ is situated between T854 and D855, and region c′ corresponds to the carbonyl oxygen of R841, while region d′ corresponds to the backbone nitrogen of M793.

Ligand-Based 3D-QSAR Analysis. The results obtained for the receptor-based 3D-QSAR study were then compared with a ligand-based model (see point 5 of Figure 1). The alignment of the molecules was developed using ROCS.[52] This software, starting from a reference structure, compares the shape of the molecules using a smooth Gaussian function. The alignment was developed using Lapatinib[50] as a reference structure.

The 3D-QSAR model was developed using the C3 and N2 probes, and as displayed in Table 2 and Figure 7, the third component of the ligand-based model showed a very similar correlation coefficient (r^2 = 0.86) and predictive correlation coefficient (q^2 = 0.74) with respect to the receptor-based 3D-QSAR model (r^2 = 0.87 and q^2 = 0.76, respectively); however, the ligand-based 3D-QSAR model possessed a worse ability for the prediction of the activity of the external test set, as it showed a higher value of the SDEP for the external test ($SDEP_{TS}$ = 0.75, whereas $SDEP_{TS}$ = 0.59 for the receptor-based 3D-QSAR model).

Virtual Screening Studies. The Maybridge database[53] was enriched with 48 known high affinity and 30 low affinity EGFR inhibitors that were not included in the 3D-QSAR studies (see Tables 2 and 3 in the Supporting Information). Conformational ensemble generations were developed using Omega 2.1.0,[54] and the database thus obtained was submitted to a shape-similarity filtering using ROCS 2.2[52] (see point 7 of Figure 1). Lapatinib was used as the reference compound for the shape-similarity search. This compound had already been used for the alignment

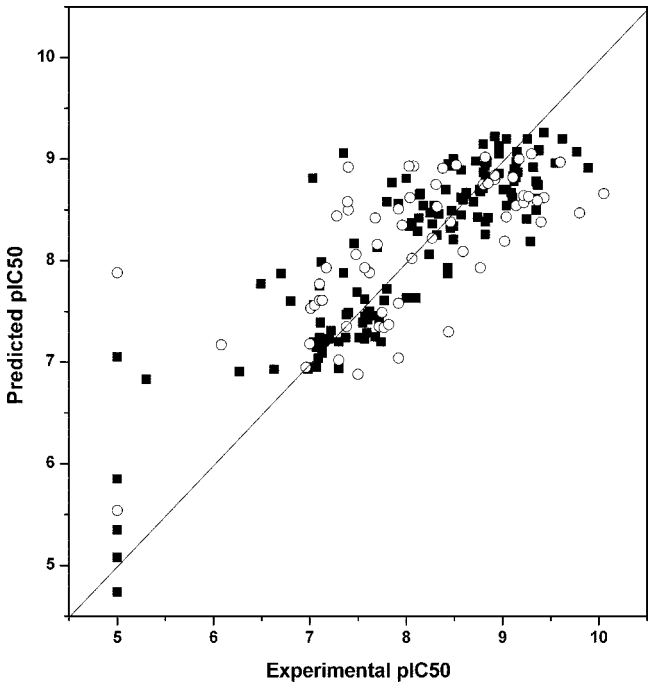


Figure 7. Plot of the ligand-based 3D-QSAR model. Experimental/predicted pIC_{50} is reported. The training set is represented as ■ and the test set as ○.

of molecules for building the ligand-based 3D-QSAR model. As already reported by Wood and co-workers lapatinib is bound to an inactive-like conformation of EGFR that is different from the active-like structure bound by the other common EGFR inhibitors.[50] However, lapatinib has a very slow off-rate from the purified intracellular domains of EGFR that correlates with a prolonged down-regulation of receptor tyrosine phosphorilation in tumor cells. Therefore, the development of new compounds showing a behavior similar to lapatininb could be very interesting. Furthermore, lapatininb also seems to be suitable for a shape-based virtual screening analysis, since it is characterized by large dimensions, occupies most of the binding site, and shows the main key interactions with the protein.

As shown in the Table 4 reported in the Supporting Information, this first step eliminated a very high percentage of the original 79 526 compounds, since only 0.3% of them passed to the next step.

Of the compounds with known EGFR inhibitory activity, all the high affinity ones were accepted and only one low affinity compound was eliminated. Hence, this step made a very important contribution to reduce the number of compounds to be studied, but it did not discriminate among compounds that possessed an activity that differed by at least 3 orders of magnitude.

The filtered database was then subjected to docking studies. This step could be carried out by means of the GOLD software using the kinase scoring function, as it showed the best results; however, for a wide analysis and a further test of the reliability of the applied procedure, all the tested docking procedure, carried out using Autodock 3.0,[34] Autodock 4.0,[34,35] Gold 3.0.1,[36] Dock 6.0,[37] and Fred 2.1[38] software, was applied on the filtered database as a qualitative filter term. It is well-known from the analysis of X-ray complexes between kinases and inhibitors interacting in the ATP binding site[55−58] and from pharmacophoric[59,60] and docking studies[61] that the backbone nitrogen of residue 793, which corresponds to a methionine in the EGFR, plays a role in inhibitor binding, as it forms a hydrogen bond with N-1 of the ATP adenine and also with the most high affinity kinase inhibitors. As a result, in the docking step all the compounds that did not form an H-bond with the backbone nitrogen of M793 were rejected.

Table 3. Summary of the Receptor-Based 3D-QSAR Rescoring of the Docking Results Obtained by Means of the Gold Software Using the Chemscore Fitness Function Modified by the Kinase Scoring Function[a]

MayBridge database	high affinity inhibitors	low affinity inhibitors
	GOLD KCS	
54	47	7
	3DQSAR	
26	43	4
EF	976	
REC	0.9	
PRC	0.9	
PR	0.8	

[a] EF = Enrichment Factor. REC = Recall. PRC = Precision. PR = Precision·Recall.

The docking results were evaluated in terms of enrichment factor (EF) and Precision·Recall (PR) (see the Experimental Section for a brief description of these measurements). As shown in Table 4 of the Supporting Information, from analysis of the EF results, the Dock procedure showed the best value (1035) and therefore seemed to be the best docking program for developing the VS study. With this software, only 15 of the 48 high affinity compounds were selected, though one low affinity compound was selected and only eight compounds of the Maybridge database were proposed as high affinity compounds.

For the PR term, the Gold software using the KCS fitness function was the best screening docking software (PR = 0.9). By this method, 47 of the 48 EGFR high affinity inhibitors were selected, and 7 low affinity compounds and 54 compounds of the Maybridge database were proposed as high affinity compounds.

In conclusion, for the VS applications, Dock (EF = 1035, PR = 0.3) and Gold using the KCS fitness function (EF = 721, PR = 0.9) seemed to be the two best approaches.

As reported above, we constructed a receptor-based 3D-QSAR model using as an alignment tool the docking poses obtained by means of the KCS method of the Gold software. The docking results of the VS study obtained using the KCS procedure were thus filtered using the 3D-QSAR model; in particular, this filter was applied considering as high affinity compounds the ligands that the 3D-QSAR model predicted with an IC_{50} better than 20 nM. As shown in Table 3, 43 of the 48 known high affinity EGFR inhibitors were selected, 4 low affinity ligands were selected, and 26 compounds of the Maybridge database were proposed as high affinity compounds. These results determined an increase in the EF value (EF = 976) similar to that obtained with the Dock software, with the maintenance of a good PR value (PR = 0.8). This analysis indicated that the best VS procedure consisted of the use of Gold with the KCS fitness function followed by the rescoring of the results through the receptor-based 3D-QSAR model.

As a further analysis, we also used Dock software as an alignment tool for the generation of receptor-based 3D-QSAR study; however, as shown in Table 5 in the Supporting Information section, the best 3D-QSAR model did not show good statistical results, as it showed a q^2 of 0.48 and an SDEP for the external test set of 1.08.

In this case the 3D-QSAR was used as a filter for the docking results. Then, in order to provide a comparison between the VS results from the docking and 3D-QSAR approaches inde-

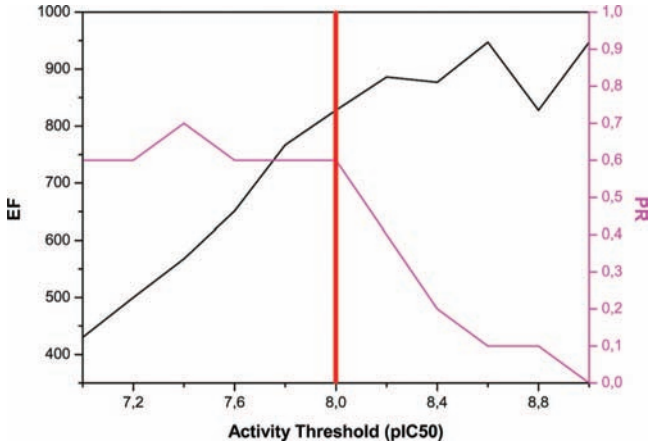


Figure 8. Statistical results of the ligand-based 3D-QSAR virtual screening. The red line indicates the activity threshold that resulted in the best EF and PR values.

pendent from one another, the enriched filtered database resulting from the ROCS analysis was subjected to the ligand-based 3D-QSAR model activity prediction.

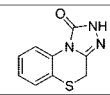
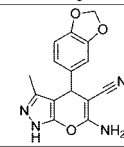
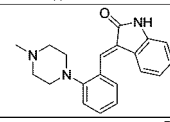
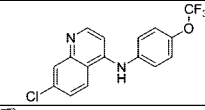
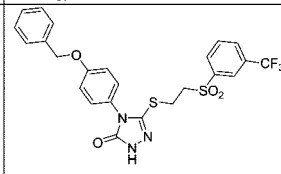
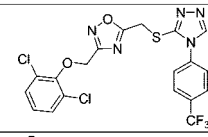
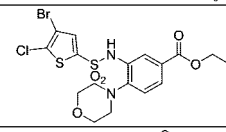
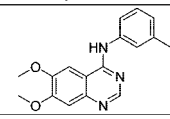
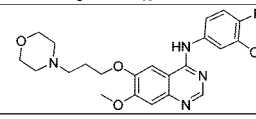
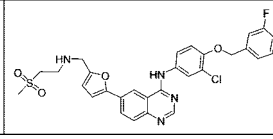
Figure 8 shows the statistical results. The best EF was obtained taking into account only those compounds with a predicted IC_{50} better than 2.5 nM; however, this threshold resulted in a PR value of 0.1. From the analysis of Figure 8, the best results in terms of EF and PR were obtained taking into account all the compounds with a predicted IC_{50} better than 10 nM (see the red line in Figure 8) that resulted in an EF of 828 and a PR value of 0.6.

These statistical results highlighted that the reliability of docking-based VS studies were comparable with the ligand-based protocol; however, the combination of the two approaches gave the best results.

In order to test from an experimental point of view the reliability of the best VS protocol obtained using the combination of Gold with the KCS fitness function followed by the rescoring of the results through the receptor-based 3D-QSAR model, the 26 compounds of the Maybridge database that were predicted as high affinity compounds by the 3D-QSAR model were visually checked, and among them 10 compounds were purchased from Maybridge (see point 11 of Figure 1). Two of them were discarded, as their NMR spectra clearly showed impurities. The remaining eight were subjected to an EGFR inhibition assay (see the Experimental Section for details). As indicated in Table 4, all the compounds tested except one showed EGFR inhibitory activity at a concentration of 100 μM, yielding an experimental hit ratio of 87.5%. Interestingly, the VS protocol succeeded in filtering a compound that showed an IC_{50} of 4.6 nM (compound **8**, GK03474[53]). However, it was not a surprising result because this compound was very similar to other already published kinase inhibitors.[62]

As shown in Figure 9A, the quinazoline ring of compound **8** formed an H-bond with M793, the two methoxy groups were directed toward the solvent accessible region of the binding site, and the 4-*m*-tolylamino substituent mainly interacted with L788 and T790, thus overlapping with the *N*-aryl substituent of lapatinib. Among the other compounds tested, two of them proved to be very interesting, as they showed a low micromolar activity and belong to two completely new chemical classes of kinase inhibitors. Compound **7** (HTS02741[53]) possessed an IC_{50} of 11.1 μM. As shown in Figure 9B, the 4-morpholino group formed an H-bond with M793, the ethylbenzoate central nucleus mimed

Table 4. Structure and EGFR Activity of the Compounds Tested

	Compd. Structure	Code	Inhibition %[a] (100μM)	IC_{50}[b] (μM)
1		HTS12932	48	
2		HTS07491	44	
3		SEW05754	55	
4		KM11074	n.a.[c]	
5		HTS02876	53	
6		HTS00213	77	6.30 (5.10-7.40)
7		HTS02741	80	11.1 (9.30-13.2)
8		GK03474	100	0.0046 (0.0038-0.0051)
		Gefitinib	100	0.022 (0.019-0.025)
		Lapatinib	100	0.017 (0.019-0.014)

[a] Percent inhibition observed at 100 μM concentration of the test compounds. Values are the average from at least two independent dose−response curves. Variations were generally within ±15%. [b] IC_{50} (95% CL) values represent the concentration required to produce 50% enzyme inhibition. Values are the average from at least two independent dose−response curves. [c] Not active. No inhibition was observed up to 100 μM of the test compound.

the substituted furan ring of lapatinib, and the 3-(4-bromo-5-chlorothiophene-2-sulfonamido) group was directed toward the core of the binding site, overlapping with the binding disposition of the *N*-aryl substituent of lapatinib. Finally, the triazole group of compound **6** (HTS00213[53]), which possessed an IC_{50} of 6.3 μM, formed an H-bond with M793, and the *p*-trifluoromethylphenyl group was directed toward the solvent accessible region of the EGFR binding site partially overlapping the furanic ring of lapatinib. The oxadiazole ring formed an H-bond with K745 and T854 and allowed the insertion of the (2,6-dichlorophenoxy)methyl group into the pocket mainly delimited by M766, C775, L777, T854, F856,

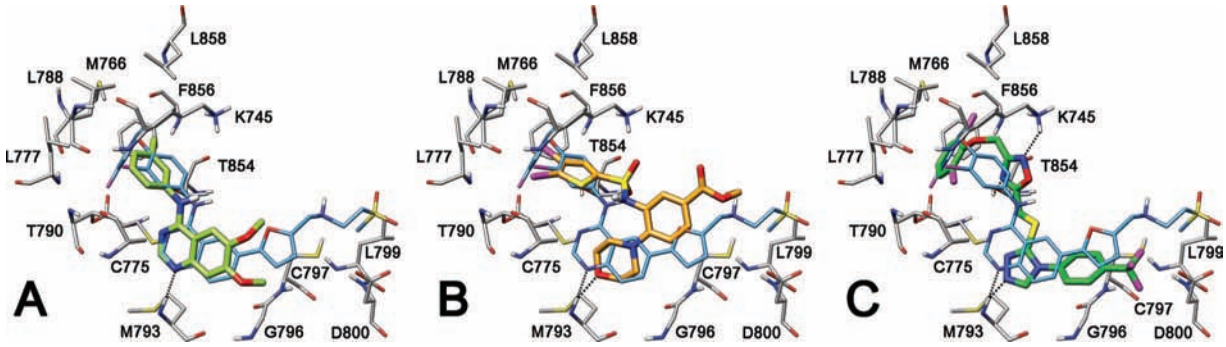


Figure 9. Docking of compounds **8** (A), **7** (B), and **6** (C) into the human EGFR binding site. Lapatinib (cyan) is reported.

and L858, in which the 3-chlorobenzyloxy group of lapatinib also interacted.

A similarity search for compounds **6** and **7** against the already published kinase ligands revealed that there are no kinase compounds with a similarity score greater than 65 (score = 100 means that the two compounds are identical), and this fact encouraged us to investigate both compounds further.

Conclusions

In this study, with EGFR as a case study, we tested the possibility of developing TK receptor-based 3D-QSAR models. First of all we analyzed the ability of the docking software to predict the binding disposition of known ligands. Then the best procedure was used to dock 206 inhibitors and the best poses were used as an alignment tool for the development of a 3D-QSAR model. The results obtained allowed us to extract both qualitative and quantitative information about the ligand−receptor interactions. 3D-QSAR analysis and in particular superimposition of the PLS coefficient surfaces and the EGFR structure revealed that residues L718, L844, and M1002 showed important lipophilic functions while T790, M793, R841, T854, and D855 might be very important for their electrostatic interactions. Finally, the analysis performed suggested to us the possibility of combining the different methods used in this study in order to develop a VS approach capable of extracting new EGFR inhibitors from ligand databases. To test the VS approaches, we used an enriched Maybridge database, to which 78 known EGFR inhibitors with an IC_{50} lower than 10 nM or up to 5 μM were added. The VS protocol based on molecular shape comparison was carried out by means of the ROCS program, which allows a fast matching of the molecular volumes. The screening results were then subjected to docking studies using various methods and software, applying a qualitative filter term. The best docking programs for developing the VS study seemed to be the Gold software using the KCS scoring function and applying the receptor-based 3D-QSAR model. The reliability of this VS protocol was supported by good values of EF and PR. Furthermore, in order to verify the reliability of the procedure experimentally, eight compounds of the Maybridge database predicted as high affinity compounds by the VS study were tested for their inhibitory activity against EGFR. All the compounds tested except one showed EGFR inhibitory activity at a concentration of 100 μM, with two compounds showing a low micromolar activity and one showing a nanomolar activity.

The two low micromolar compounds may be considered as very interesting leads for the development of new high affinity EGFR inhibitors, and further studies will be carried out. Taken together, these findings suggest that the techniques thus optimized may be suitable for the design of new EGFR inhibitors, and the results obtained encourage us to also apply this method to other TKs, including the systems for which X-ray data are not available.

Experimental Section

Kinase−Inhibitor Complex Structures and Docking. A total of 15 wild and mutated EGFR structures, available from the Protein Data Bank,[63] were used in this work (see Figure 2). Moreover, in order to enhance the number of kinase complexes used for testing the docking procedures, the primary sequence of the catalytic domain of human EGFR, obtained from the SWISS-PROT protein sequence database,[64] was subjected to a sequence similarity search carried out by means of Blastp.[65] All the kinase proteins with an alignment score greater than 74 were taken into account; as a result, the 27 kinase−inhibitor X-ray structures possessing the highest sequence similarity with respect to the catalytic domain of human EGFR were included in the docking set (see Figure 2).

Hydrogen atoms were added by means of Maestro 7.5.[66] The ligands were extracted from the X-ray complexes and were then subjected to a CS of 1000 steps in a water environment (using the generalized Born/surface-area model) by means of Macromodel.[67] The algorithm used was the Monte Carlo method with the MMFFs force field and a distance-dependent dielectric constant of 1.0. The ligands were then minimized using the conjugated gradient method until a convergence value of 0.05 kcal/(Å·mol), using the same force field and parameters as for the conformational search. The ligands minimized in this way were docked into their corresponding proteins by means of Autodock 3.0,[34] Autodock 4.0,[34,35] Gold 3.0.1,[36] Dock 6.0,[37] and Fred 2.1[38] software.

The docking results were evaluated through a comparison between the found docked positions of the ligand and the experimental ones. As a measure of docking reliability, the rmsd between the positions of heavy atoms of the ligand in the calculated and experimental structures was taken into account.

As already reported by Shewchuk and co-workers,[68] also for the 4-anilinoquinazoline EGFR inhibitors, X-ray stuctures highlighted the presence of an ordered water that bridged the N3 position of the anilinoquinazoline ring to Thr766. However, since we also take into account compounds with different central scaffolds, we did not take into account the presence of this water molecule.

Docking Procedures. Autodock 3.0 and 4.0. Autodock Tools[69] was used in order to identify the torsion angles in the ligands, add the solvent model, and assign the Kollman atomic charges to the protein. The ligand charge was calculated using the AM1-BCC and Gasteiger method, as implemented in the Antechamber suite of Amber 9.[70] The regions of interest used by Autodock were defined by considering lapatinib (PDB code 1XKK[50]) as the central group; in particular, a grid of 50, 46, and 58 points in the *x*, *y*, and *z* directions was constructed centered on the center of the mass of this inhibitor. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the energetic map calculations.

By use of the Lamarckian genetic algorithm, the docked compounds were subjected to 100 runs of the Autodock search, using 500 000 steps of energy evaluation and the default values of the other parameters. Cluster analysis was performed on the results using an rms tolerance of 1.0 Å.

The docking pose analysis was carried out for the first pose of the first cluster and the first pose of the most populated cluster.

Gold 3.0.1. The region of interest used by Gold was defined in such a manner that it contained the residues that stayed within 10 Å from the ligand in the X-ray structures; the "allow early termination" command was deactivated, while the possibility for the ligand to flip ring corners was activated. For all the other parameters, the Gold default ones were used, and the ligands were submitted to 30 genetic algorithm runs.

Three docking analyses were carried out. In the first two cases, the two fitness functions implemented in Gold, i.e., GoldScore and ChemScore, were used. In the third one, we used the kinase scoring function implemented in Gold (applicable using ChemScore as fitness function), which calculates a contribution for weak CH···O interactions.

Dock 6.0. The molecular surface of the binding site was calculated by means of the MS program,[37] generating the Connolly surface with a probe with a radius of 1.4 Å; the points of the surface and the vectors normal to it were used by the Sphgen program[37] to build a set of spheres, with radii varying from 1.4 to 4 Å that describe, from a stereoelectronic point of view, the negative image of the site. The first set of spheres automatically generated and clustered by the program was used to represent the site. For each ligand, Dock 6.0 calculated 500 orientations; of these, the best grid scored was taken into consideration.

Fred 2.15. Fred requires a set of input conformers for each ligand. The conformers were generated by Omega 2.1.0[54] (Open Eye Scientific Software). We applied the following modifications to the default setting of Omega: the maximum number of output conformers was set to 10 000, and the rmsd value below which two conformations were considered to be similar was set at 0.5 Å. Fred docking roughly consists of two steps: shape fitting and optimization. During shape fitting, the ligand is placed into a 0.5 Å resolution grid box encompassing all active-site atoms (including hydrogens) using a smooth Gaussian potential. A series of three optimization filters are then processed, which consist of refinement of the positions of hydroxyl hydrogen atoms of the ligand, rigid body optimization, and optimization of the ligand pose in the dihedral angle space. In the last optimization step, the Chemgauss2 scoring function was used. After the docking calculation the poses returned were scored independently by the five available scoring functions (Chemgauss2, Chemscore, Plp, Screenscore, Shapegauss) and by a combination of them.

Receptor-Based 3D-QSAR Model. Alignment of the Molecules. The 206 compounds were docked inside the 1XKK[50] EGFR binding site using the KCS scoring function of Gold and also using Dock software. For each ligand, the best docked structure was chosen, and this receptor-based alignment was used for further studies.

Data Set. The Golpe program[51] was used to define a 3D-QSAR model, using GRID interaction fields[71] as descriptors (see below). The training set was composed of 142 compounds, characterized by affinity values spanning about 5 orders of magnitude. Similarly, the 64 compounds belonging to the test set showed a pIC_{50} affinity value ranging from 5.0 to 10 and were uniformly distributed along the activity range. The generation of the training and test sets was automatically developed using an in-house program. Using a shape-based algorithm, the program is able to select a series of compounds with a broad activity range, limiting the number of compounds that are similarly active and structurally similar.

Probe Selection. The Grid program[71] was used to describe the previously superimposed molecular structure. Interaction energies between the selected probes and each molecule were calculated using a grid spacing of 1 Å. Initially, many probes were tested and the preliminary PLS analyses suggested that the use of a combination of C3 (corresponding to a methyl group) and N2 (corresponding to neutral flat NH_2) probes best described the system.

Variable Selection. The MIFs of the training set were imported into the Golpe program; it is well-known that many of the variables derived from Grid analysis may be considered as noise, which decreases the quality of the model. For this reason, variable selection was operated by zeroing absolute values smaller than 0.06 kcal/mol and removing variables with a standard deviation below 0.1. Furthermore, variables that exhibited only two values and had a skewed distribution were also removed.

The smart region definition (SRD) algorithm[72] was applied with 10% of the active variables as the number of seed (selected in the PLS weights space), a critical distance cutoff of 2.5 Å, and a collapsing distance cutoff of 4.0 Å. The groups were then used in the fractional factorial design (FFD) procedure. FFD selection was applied twice, until the r^2 and q^2 values did not increase significantly, using the cross-validation routine with five random sets of compounds.

Ligand-Based 3D-QSAR Model. The alignment of the compounds was developed using the ROCS 2.2 software,[52] which is a shape-similarity method based on the Tanimoto-like overlap of volumes. The alignment was developed using the combo score, which combines the Tanimoto shape score with the color score for the appropriate overlap of groups with similar properties (donor, acceptor, hydrophobe, cation, anion, and ring). All the other parameters were used as ROCS default values. The maximum number of conformations per molecule was ~1000, generated using Omega 2.1.0.[54] Default parameters were used with the following exception: the rms parameter (which sets the minimum root mean square Cartesian distance below which two conformers are duplicates) was set to 0.5 (default = 0.8).

Lapatinib was extracted from the X-ray complex with EGFR (1XKK[50] PDB code) and was used as the reference structure for the alignment. For the data set, probe, and variable selection, the same approach and parameters applied for the receptor-based 3D-QSAR study were used.

Virtual Screening Protocol. The Maybridge database,[53] consisting of 79 448 compounds, was enriched with 48 known EGFR inhibitors with an IC_{50} better than 10 nM [73−80] and 30 compounds with an IC_{50} EGFR inhibitor activity worse than 5 μM.[81−85] These compounds were not included in the previous docking and 3D-QSAR studies, and their activity was tested by using various methods.

The conformation ensemble was generated using Omega 2.1.0.[54] Default parameters were used with the following exception: the maxconfs parameter (which sets the maximum number of conformations to be generated) was set to 10 000 (default = 20000).

The resulting enriched database generated by Omega was processed with ROCS 2.2 software.[52] Our studies showed that this software could be used as a ligand-based alignment tool for developing EGFR 3D-QSAR models; furthermore, ROCS is capable of processing 600−800 comparisons per second, making it possible to search multiconformer representations of corporate collections with a low time-consuming profile.

As a reference structure for shape comparison, the quinazoline derivative lapatinib extracted from the 1XKK X-ray structure[50] was chosen. It had already been used for the ligand-based 3D-QSAR study. All the compounds that showed a combo score higher than 0.50 were selected. The analysis of the ROCS alignment for building the ligand-based 3D-QSAR model revealed that all the EGFR inhibitors showed a combo score value higher than 0.5 with respect to lapatinib.

The filtered database was then subjected to docking studies. All the previously analyzed kinds of docking software (Autodock 3.0 and 4.0, Dock 6.0, Gold 3.0.1, and Fred 2.15) were used, applying the procedures described in the "Docking Procedures" section. In order to reduce the database further, at this step a qualitative filter term was applied; in the docking step, all the compounds that did not form an H-bond with the backbone nitrogen of M793 were rejected.

The results obtained using Gold with the KCS scoring function were further filtered using the receptor-based 3D-QSAR model developed.

The VS results were evaluated using four different performance measures: Recall, Precision, Enrichment, and the product of Recall and Precision.[86,87]

The Recall value for positives describes the ratio of correctly classified members of a data set. It is defined by

$$\text{Recall} = t_p/(t_p + f_n)$$

where t_p is the number of high affinity compounds not rejected (true positives) and f_n is the number of high affinity compounds rejected during the VS filtering (false negatives).

Precision measures the percentage of known high affinity compounds among all the compounds with known activity:

$$\text{Precision} = t_p/(t_p + f_p)$$

where f_p is the number of low affinity compounds selected by the VS protocol (false positives).

Enrichment factor (EF) measures the enrichment of the method compared with random selection:

$$\text{EF} = [t_p/(t_p + f_n)](\text{NC}_{tot}/\text{NC})$$

where NC_{tot} is the total number of molecules of the database ($\text{NC}_{tot} = 79448 + 48 + 30$) and NC is the total number of compounds obtained at the end of the VS protocol.

The Precision term does not tell us anything about the absolute number of predicted high affinity compounds taken into account using the Recall term. In order to take both terms into consideration, the product of precision and recall (PR) was used.

$$\text{PR} = (\text{Precision})(\text{Recall})$$

Biology. Materials and Methods. Ten compounds selected by VS were purchased from Maybridge.[53] All compounds were subjected to NMR analysis prior to testing for their inhibitory activity, to verify their correspondence to the chemical structures shown in Table 4. Two compounds out of 10 were discarded because of their impurity. Gefitinib was synthesized by our laboratory following a previously reported procedure.[88] Human recombinant protein tyrosine kinase EGFR (ErbB1) and protein tyrosine kinase assay kit, Omnia Tyr human recombinant kit 7, were from Invitrogen.[89] All solvents were from Sigma-Aldrich.

Tyrosine Kinase Assays. EGFR enzyme assays were performed in 96-well microtiter plates using the Omnia Tyr human recombinant kit 7 (Biosource), in accordance with the manufacturer's protocol.

The kinase activity was determined fluorimetrically by monitoring for 30 min the increase in fluorescence resulting from phosphorylation of a peptide substrate, carrying the fluorophore 8-hydroxy-5-(*N*,*N*-dimethylsulfonamido)-2-methylquinoline,[90] catalyzed by EGFR in the presence of ATP. EGFR was assayed at 30 °C in a reaction mixture containing 5 μL of tyrosine kinase reaction buffer (composed of 20 mM HEPES, pH 7.4, 10 mM $MgCl_2$, and 0.3 mM $CaCl_2$), 5 μL of tyrosine kinase substrate, 5 μL of 1 mM ATP, 5 μL of 1 mM DTT, 25 μL of ultrapure water, and 5 μL of 3 mU/μL EGFR in a total volume of 50 μL. All the above reagents, except EGFR, were incubated at 30 °C for 5 min. EGFR was then added to start the reaction, which was monitored with the fluorescence meter Victor3 PerkinElmer at 360 nm (excitation filter) and 485 nm (emission filter). Kinase activity was calculated from a linear least-squares fit of the data for fluorescence intensity versus time.

Enzymatic Inhibition. The inhibitory activity of the selected compounds against EGFR was assayed by adding 5 μL of the inhibitor solution to the reaction mixture described above. All the products were dissolved in 100% DMSO and diluted to the appropriate concentrations with tyrosine kinase reaction buffer, provided by the kit. The final concentration of DMSO in assay solutions never exceeded 1% and proved to have no effects on EGFR activity. The inhibitory effect of the products was routinely estimated at a concentration of 10^{-4} M. Those compounds found to be active were then tested at additional concentrations between 10^{-5} and 10^{-10} M. For a proper comparison gefitinib ($IC_{50} = 0.023$ μM [17]) and lapatinib ($IC_{50} = 0.0108$ μM [19]) were employed as the reference standards. The determination of the IC_{50} values was performed by linear regression analysis of the log-dose response curve, which was generated using at least four concentrations of the inhibitor causing an inhibition between 20% and 80%, with two replicates at each concentration. The 95% confidence limits (95% CL) were calculated from t values for $n - 2$, where n is the total number of determinations (see Table 4).

Similarity Search. The test compounds **6** and **7** were compared with published kinase ligands. This comparison was made by using the Similarity Search task of SciFinder scholar.[91] Similarity search locates structures that are similar to the query, based on a two-dimensional small molecule comparison using a Tanimoto similarity metric. The Tanimoto metric assigns a score based on structure descriptors, as follows:

$$\text{score} = (100 \times C)/[(\text{QS} + \text{FS}) - C]$$

where C is the number of descriptors that the query and answer set structures have in common, QS is the number of descriptors in the query structure, and FS is the number of descriptors in the answer set structure. A score of 100 means that two structures are identical.

Acknowledgment. Many thanks are due to Prof. Gabriele Cruciani and Prof. Sergio Clementi (Molecular Discovery and MIA srl) for the use of the GOLPE program in their chemometric laboratory (University of Perugia, Italy) and for having provided the GRID program. Figure 9 was produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH Grant P41 RR-01081).

Supporting Information Available: Actual versus predicted data of the receptor- and ligand-based 3D-QSAR model (Table 1), high affinity (Table 2) and low affinity (Table 3) EGFR inhibitors used for the enrichment of the Maybridge database, summary of the virtual screening results (Table 4), statistical results of the receptor-based 3D-QSAR model obtained using Dock software (Table 5), and ^{1}H NMR spectra (Table 6) and elemental analysis results for the purchased Maybridge compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hanks, S. K.; Quinn, A. M.; Hunter, T. The Protein Kinase Family: Conserved Features and Deduced Phylogeny of the Catalytic Domains. *Science* **1988**, *241*, 42–52.
(2) Ullrich, A.; Schlessinger, J. Signal Transduction by Receptors with Tyrosine Kinase Activity. *Cell* **1990**, *61*, 203–212.
(3) Levitzki, A.; Gazit, A. Tyrosine Kinase Inhibition: An Approach to Drug Development. *Science* **1995**, *267*, 1782–1788.
(4) Scapin, G. Structural Biology in Drug Design: Selective Protein Kinase Inhibitors. *Drug Discovery Today* **2002**, *7*, 601–611.
(5) Bafico, A.; Aaronson, S. A. Growth Factors and Signal Transduction in Cancer. *Cancer Med.* **2006**, *7*, 53–67.
(6) Bridges, A. J. Chemical Inhibitors of Protein Kinases. *Chem. Rev.* **2001**, *101*, 2541–2571.
(7) Holbro, T.; Hynes, N. E. ErbB Receptors: Directing Key Signaling Network Throughout Life. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 195–217.
(8) García-Echeverria, C.; Fabbro, D. Therapeutically Targeted Anticancer Agents: Inhibitors of Receptor Tyrosine Kinases. *Mini-Rev. Med. Chem.* **2004**, *4*, 273–283.
(9) Davies, D. E.; Chamberlin, S. G. Targeting the Epidermal Growth Factor Receptor for Therapy of Carcinomas. *Biochem. Pharmacol.* **1996**, *51*, 1101–1110.
(10) Woodburn, J. R. The Epidermal Growth Factor Receptor and its Inhibition in Cancer Therapy. *Pharmacol. Ther.* **1999**, *82*, 241–250.

(11) Burke, T. R. Protein Tyrosine Kinase Inhibitors. *Drugs Future* **1992**, *17*, 119–131.
(12) Fry, D. W. Protein Tyrosine Kinases as Therapeutic Targets in Cancer Chemotherapy and Recent Advances in the Development of New Inhibitors. *Exp. Opin. Invest. Drugs* **1994**, *3*, 577–595.
(13) Traxler, P.; Lydon, N. Recent Advances in Protein Tyrosine Kinase Inhibitors. *Drugs Future* **1995**, *20*, 1261–1274.
(14) Rabindran, S. K.; Discafani, C. M.; Rosfjord, E. C.; Baxter, M.; Floyd, M. B.; Golas, J.; Hallett, W. A.; Johnson, B. D.; Nilakantan, D.; OverBeek, E.; Reich, M. F.; Shen, R.; Shi, X.; Tsou, H. R.; Wang, Y. F.; Wissner, A. Antitumor Activity of HKI-272, an Orally Active, Irreversible Inhibitor of the Her-2 Tyrosine Kinase. *Cancer Res.* **2004**, *64*, 3958–3965.
(15) Li, D.; Ambrogio, L.; Shimamura, T.; Kubo, S.; Takahashi, M.; Chirieac, L. R.; Padera, R. F.; Shapiro, G. I.; Baum, A.; Himmelsbach, F.; Retting, W. J.; Meyerson, M.; Solca, F.; Greulich, H.; Wong, K. K. BIBW2992, an Irreversible EGFR/HER2 Inhibitor Highly Effective in Preclinical Lung Cancer Model. *Oncogene* **2008**, *27*, 4702–4711.
(16) Riely, G. J. Second-Generation Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. *J. Thorac. Oncol.* **2008**, *3*, S146–S149.
(17) Barker, A. J.; Gibson, K. H.; Grundy, W.; Godfrey, A. A.; Barlow, J. J.; Healy, M. P.; Woodburn, J. R.; Ashton, S. E.; Curry, B. J.; Scarlett, L.; Henthorn, L.; Richards, L. Studies Leading to the Identification of ZD1839 (Iressa): An Orally Active, Selective Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Targeted to the Treatment of Cancer. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1911–1914.
(18) Moyer, J. D.; Barbacci, E. G.; Iwata, K. K.; Arnold, L.; Boman, B.; Cunningham, A.; DiOrio, C.; Doty, J.; Morin, M. J.; Moyer, M. P.; Neveu, M.; Pollack, V. A.; Pustilnik, L. R.; Reynolds, M. M.; Sloan, D.; Theleman, A.; Miller, P. Induction of Apoptosis and Cell Cycle Arrest by CP-358,774, an Inhibitor of Epidermal Growth Factor Receptor Tyrosine Kinase. *Cancer Res.* **1997**, *57*, 4838–4848.
(19) Rusnak, D. W.; Lackey, K.; Affleck, K.; Wood, E. R.; Alligood, K. J.; Rhodes, N.; Keith, B. R.; Murray, D. M.; Glennon, K.; Knight, W. B.; Mullin, R. J.; Gilmer, T. M. The Effects of the Novel, Reversible Epidermal Growth Factor Receptor/ErbB-2 Tyrosine Kinase Inhibitor, GW2016, on the Growth of Human Normal and Tumor-Derived Cell Lines in Vitro and in Vivo. *Mol Cancer Ther.* **2001**, *1*, 85–94.
(20) Aparna, V.; Rambabu, G.; Panigrahi, S. K.; Sarma, J. A.; Desiraju, G. R. Virtual Screening of 4-Anilinoquinazoline Analogues as EGFR Kinase Inhibitors: Importance of Hydrogen Bonds in the Evaluation of Poses and Scoring Functions. *J. Chem. Inf. Model.* **2005**, *45*, 725–738.
(21) Cavasotto, C. N.; Ortiz, M. A.; Abagyan, R. A.; Piedrafita, F. J. In Silico Identification of Novel EGFR Inhibitors with Antiproliferative Activity Against Cancer Cells. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1969–1974.
(22) Abouzid, K.; Shouman, S. Design, Synthesis and in Vitro Antitumor Activity of 4-Aminoquinoline and 4-Aminoquinazoline Derivatives Targeting EGFR Tyrosine Kinase. *Bioorg. Med. Chem.* **2008**, *16*, 7543–7551.
(23) Gundla, R.; Kazemi, R.; Sanam, R.; Muttineni, R.; Sarma, J. A.; Dayam, R.; Neamati, N. Discovery of Novel Small-Molecule Inhibitors of Human Epidermal Growth Factor Receptor-2: Combined Ligand and Target-Based Approach. *J. Med. Chem.* **2008**, *51*, 3367–3377.
(24) Szantai-Kis, C.; Kovesdi, I.; Eros, D.; Banhegyi, P.; Ullrich, A.; Keri, G.; Orfi, L. Prediction Oriented QSAR Modelling of EGFR Inhibition. *Curr. Med. Chem.* **2006**, *13*, 277–287.
(25) Shi, W.; Shen, Q.; Kong, W.; Ye, B. QSAR Analysis of Tyrosine Kinase Inhibitor Using Modified Ant Colony Optimization and Multiple Linear Regression. *Eur. J. Med. Chem.* **2007**, *42*, 81–86.
(26) Deeb, O.; Clare, B. W. QSAR of Aromatic Substances: EGFR Inhibitory Activity of Quinazoline Analogues. *J. Enzyme Inhib. Med. Chem.* **2008**, *23*, 763–775.
(27) Peng, T.; Pei, J.; Zhou, J. 3D-QSAR and Receptor Modeling of Tyrosine Kinase Inhibitors with Flexible Atom Receptor Model (FLARM). *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 298–303.
(28) Kamath, S.; Buolamwini, J. K. Receptor-Guided Alignment-Based Comparative 3D-QSAR Studies of Benzylidene Malonitrile Tyrphostins as EGFR and HER-2 Kinase Inhibitors. *J. Med. Chem.* **2003**, *46*, 4657–4668.
(29) Assefa, H.; Kamath, S.; Buolamwini, J. K. 3D-QSAR and Docking Studies on 4-Anilinoquinazoline and 4-Anilinoquinoline Epidermal Growth Factor Receptor (EGFR) Tyrosine Kinase Inhibitors. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 475–493.
(30) Chen, G.; Luo, X.; Zhu, W.; Luo, C.; Liu, H.; Puah, C. M.; Chen, K.; Jiang, H. Elucidating Inhibitory Models of the Inhibitors of Epidermal Growth Factor Receptor by Docking and 3D-QSAR. *Bioorg. Med. Chem.* **2004**, *12*, 2409–2417.
(31) Pednekar, D. V.; Kelkar, M. A.; Pimple, S. R.; Akamanchi, K. G. 3D QSAR Studies of Inhibitors of Epidermal Growth Factor Receptor [EGFR] Using CoMFA and GFA Methodologies. *Med. Chem. Res.* **2004**, *13*, 605–618.
(32) Pinto-Bazurco, M.; Tsakovska, I.; Pajeva, I. QSAR and 3D QSAR of Inhibitors of the Epidermal Growth Factor Receptor. *Int. J. Quantum Chem.* **2006**, *106*, 1432–1444.
(33) San Juan, A. A. Towards Predictive Inhibitor Design for the EGFR Autophosphorylation Activity. *Eur. J. Med. Chem.* **2008**, *43*, 781–791.
(34) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(35) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152.
(36) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(37) *DOCK*, version 6.0; Molecular Design Institute, University of California: San Francisco, CA, 1998.
(38) McGann, M.; Almond, H.; Nicholls, A.; Grant, J. A.; Brown, F. Gaussian Docking Functions. *Biopolymers* **2003**, *68*, 76–90.
(39) Gaul, M. D.; Guo, Y.; Affleck, K.; Cockerill, G. S.; Gilmer, T. M.; Griffin, R. J.; Guntrip, S.; Keith, B. R.; Knight, W. B.; Mullin, R. J.; Murray, D. M.; Rusnak, D. W.; Smith, K.; Tadepalli, S.; Wood, E. R.; Lackey, K. Discovery and Biological Evaluation of Potent Dual ErbB-2/EGFR Tyrosine Kinase Inhibitors: 6-Thiazolylquinazolines. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 637–640.
(40) Zhang, Y. M.; Cockerill, S.; Guntrip, S. B.; Rusnak, D.; Smith, K.; Vanderwall, D.; Wood, E.; Lackey, K. Synthesis and SAR of Potent EGFR/erbB2 Dual Inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 111–114.
(41) Alberti, M. J.; Auten, E. P.; Lackey, K. E.; McDonald, O. B.; Wood, E. R.; Preugschat, F.; Cutler, G. J.; Kane-Carson, L.; Liu, W.; Jung, D. K. Discovery and in Vitro Evaluation of Potent Kinase Inhibitors: Pyrido[1′,2′:1,5]pyrazolo[3,4-*d*]pyrimidines. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3778–3781.
(42) Petrov, K. G.; Zhang, Y. M.; Carter, M.; Cockerill, G. S.; Dickerson, S.; Gauthier, C. A.; Guo, Y., Jr.; Rusnak, D. W.; Walker, A. L.; Wood, E. R.; Lackey, K. E. Optimization and SAR for Dual ErbB-1/ErbB-2 Tyrosine Kinase Inhibition in the 6-Furanylquinazoline Series. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4686–4691.
(43) Waterson, A. G.; Stevens, K. L.; Reno, M. J.; Zhang, Y. M.; Boros, E. E.; Bouvier, F.; Rastagar, A.; Uehling, D. E.; Dickerson, S. H.; Reep, B.; McDonald, O. B.; Wood, E. R.; Rusnak, D. W.; Alligood, K. J.; Rudolph, S. K. Alkynyl Pyrimidines as Dual EGFR/ErbB2 Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2419–2422.
(44) Smaill, J. B.; Rewcastle, G. W.; Loo, J. A.; Greis, K. D.; Chan, O. H.; Reyner, E. L.; Lipka, E.; Showalter, H. D.; Vincent, P. W.; Elliott, W. L.; Denny, W. A. Tyrosine Kinase Inhibitors. 17. Irreversible Inhibitors of the Epidermal Growth Factor Receptor: 4-(Phenylamino)quinazoline- and 4-(Phenylamino)pyrido[3,2-*d*]pyrimidine-6-acrylamides Bearing Additional Solubilizing Functions. *J. Med. Chem.* **2000**, *43*, 1380–1397.
(45) Smaill, J. B.; Showalter, H. D.; Zhou, H.; Bridges, A. J.; McNamara, D. J.; Fry, D. W.; Nelson, J. M.; Sherwood, V.; Vincent, P. W.; Roberts, B. J.; Elliott, W. L.; Denny, W. A. Tyrosine Kinase Inhibitors. 18. 6-Substituted 4-Anilinoquinazolines and 4-Anilinopyrido[3,4-*d*]pyrimidines as Soluble, Irreversible Inhibitors of the Epidermal Growth Factor Receptor. *J. Med. Chem.* **2001**, *44*, 429–440.
(46) Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trumpp-Kallmeyer, S.; Dobrusin, E. M. Specific, Irreversible Inactivation of the Epidermal Growth Factor Receptor and ErbB2, by a New Class of Tyrosine Kinase Inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12022–12027.
(47) Smaill, J. B.; Palmer, B. D.; Rewcastle, G. W.; Denny, W. A.; McNamara, D. J.; Dobrusin, E. M.; Bridges, A. J.; Zhou, H.; Showalter, H. D.; Winters, R. T.; Leopold, W. R.; Fry, D. W.; Nelson, J. M.; Slintak, V.; Elliot, W. L.; Roberts, B. J.; Vincent, P. W.; Patmore, S. J. Tyrosine Kinase Inhibitors. 15. 4-(Phenylamino)quinazoline and 4-(Phenylamino)pyrido[*d*]pyrimidine Acrylamides as Irreversible Inhibitors of the ATP Binding Site of the Epidermal Growth Factor Receptor. *J. Med. Chem.* **1999**, *42*, 1803–1815.
(48) Palmer, B. D.; Trumpp-Kallmeyer, S.; Fry, D. W.; Nelson, J. M.; Showalter, H. D.; Denny, W. A. Tyrosine Kinase Inhibitors. 11. Soluble Analogues of Pyrrolo- and Pyrazoloquinazolines as Epidermal Growth Factor Receptor Inhibitors: Synthesis, Biological Evaluation, and Modeling of the Mode of Binding. *J. Med. Chem.* **1997**, *40*, 1519–1529.
(49) Thompson, A. M.; Murray, D. K.; Elliott, W. L.; Fry, D. W.; Nelson, J. A.; Showalter, H. D.; Roberts, B. J.; Vincent, P. W.; Denny, W. A.

Tyrosine Kinase Inhibitors. 13. Structure−Activity Relationships for Soluble 7-Substituted 4-[(3-Bromophenyl)amino]pyrido[4,3-*d*]pyrimidines Designed as Inhibitors of the Tyrosine Kinase Activity of the Epidermal Growth Factor Receptor. *J. Med. Chem.* **1997**, *40*, 3915–3925.

(50) Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A.; Dickerson, S. H.; Ellis, B.; Pennisi, C.; Horne, E.; Lackey, K.; Alligood, K. J.; Rusnak, D. W.; Gilmer, T. M.; Shewchuk, L. A Unique Structure For Epidermal Growth Factor Receptor Bound to GW572016 (Lapatinib): Relationships among Protein Conformation, Inhibitor Off-Rate, and Receptor Activity in Tumor Cells. *Cancer Res.* **2004**, *64*, 6652–6659.

(51) *GOLPE*, version 4.5; Multivariate Infometric Analysis Srl. (Viale dei Castagni 16): Perugia, Italy, 1999.

(52) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. A Shape-Based 3-D Scaffold Hopping Method and Its Application to a Bacterial Protein−Protein Interaction. *J. Med. Chem.* **2005**, *48*, 1489–1495.

(53) Maybridge. http://www.maybridge.com.

(54) Boström, J.; Greenwood, J. R.; Gottfries, J. Assessing the Performance of OMEGA with Respect to Retrieving Bioactive Conformations. *J. Mol. Graphics Modell.* **2003**, *21*, 449–462.

(55) Mohammadi, M.; McMahon, G.; Sun, L.; Tang, C.; Hirth, P.; Yeh, B. K.; Hubbard, S. R.; Schlessinger, J. Structures of the Tyrosine Kinase Domain of Fibroblast Growth Factor Receptor in Complex with Inhibitors. *Science* **1997**, *276*, 955–960.

(56) Zhu, X.; Kim, J. L.; Newcomb, J. R.; Rose, P. E.; Stover, D. R.; Toledo, L. M.; Zhao, H.; Morgenstern, K. A. Structural Analysis of the Lymphocyte-Specific Kinase Lck in Complex with Non-Selective and Src Family Selective Kinase Inhibitors. *Structure* **1999**, *7*, 651–661.

(57) Nagar, B.; Bornmann, W. G.; Pellicena, P.; Schindler, T.; Veach, D. R.; Miller, W. T.; Clarkson, B.; Kuriyan, J. Crystal Structures of the Kinase Domain of c-Abl in Complex with the Small Molecule Inhibitors PD173955 and Imatinib (STI-571). *Cancer Res.* **2002**, *62*, 4236–4243.

(58) Stamos, J.; Sliwkowski, M. X.; Eigenbrot, C. Structure of the Epidermal Growth Factor Receptor Kinase Domain Alone and in Complex with a 4-Anilinoquinazoline Inhibitor. *J. Biol. Chem.* **2002**, *277*, 46265–46272.

(59) Traxler, P.; Green, J.; Mett, H.; Sequin, U.; Furet, P. Use of a Pharmacophore Model for the Design of EGFR Tyrosine Kinase Inhibitors: Isoflavones and 3-Phenyl-4(1*H*)-quinolones. *J. Med. Chem.* **1999**, *42*, 1018–1026.

(60) Traxler, P.; Furet, P. Strategies toward the Design of Novel and Selective Protein Tyrosine Kinase Inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195–206.

(61) Tuccinardi, T.; Manetti, F.; Schenone, S.; Martinelli, A.; Botta, M. Construction and Validation of a RET TK Catalytic Domain by Homology Modeling. *J. Chem. Inf. Model.* **2007**, *47*, 644–655.

(62) Gomes Franchini, K.; Abdalla Saad, M. J.; Rittner Neto, R.; Marin, R. M.; Rocco Aparecida, S. Preparation of Compounds Derived from 4-Anilinequinazolines with Adenosine Kinase Inhibitor Properties for Therapeutic Use. PCT Int. Appl. WO2005085213, 2005.

(63) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(64) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: The Proteomics Server for In-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788.

(65) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

(66) *Maestro*, version 7.5; Schrödinger Inc.: Portland, OR, 2005.

(67) *Macromodel*, version 8.5; Schrödinger Inc.: Portland, OR, 1999.

(68) Shewchuk, L.; Hassell, A.; Wisely, B.; Rocque, W.; Holmes, W.; Veal, J.; Kuyper, L. F. Binding Mode of the 4-Anilinoquinazoline Class of Protein Kinase Inhibitor: X-ray Crystallographic Studies of 4-Anilinoquinazolines Bound to Cyclin-Dependent Kinase 2 and p38 Kinase. *J. Med. Chem.* **2000**, *43*, 133–138.

(69) Sanner, M. F. Python: A Programming Language for Software Integration and Development. *J. Mol. Graphics Modell.* **1999**, *17*, 57–61.

(70) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER*, version 9; University of California: San Francisco, CA, 2006.

(71) *GRID*, version 22a; Molecular Discovery Ltd. (West Way House, Elms Parade): Oxford, U.K., 2004.

(72) Pastor, M.; Cruciani, G.; Clementi, S. Smart Region Definition: A New Way To Improve the Predictive Ability and Interpretability of Three-Dimensional Quantitative Structure−Activity Relationships. *J. Med. Chem.* **1997**, *40*, 1455–1464.

(73) Ballard, P.; Bradbury, R. H.; Harris, C. S.; Hennequin, L. F.; Hickinson, M.; Johnson, P. D.; Kettle, J. G.; Klinowska, T.; Leach, A. G.; Morgentin, R.; Pass, M.; Ogilvie, D. J.; Olivier, A.; Warin, N.; Williams, E. J. Inhibitors of Epidermal Growth Factor Receptor Tyrosine Kinase: Novel C-5 Substituted Anilinoquinazolines Designed To Target the Ribose Pocket. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1633–1637.

(74) Ballard, P.; Bradbury, R. H.; Harris, C. S.; Hennequin, L. F.; Hickinson, M.; Kettle, J. G.; Kendrew, J.; Klinowska, T.; Ogilvie, D. J.; Pearson, S. E.; Williams, E. J.; Wilson, I. Inhibitors of Epidermal Growth Factor Receptor Tyrosine Kinase: Optimisation of Potency and in Vivo Pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4908–4912.

(75) VanBrocklin, H. F.; Lim, J. K.; Coffing, S. L.; Hom, D. L.; Negash, K.; Ono, M. Y.; Gilmore, J. L.; Bryant, I.; Riese, D. J., 2nd. Anilinodialkoxyquinazolines: Screening Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors for Potential Tumor Imaging Probes. *J. Med. Chem.* **2005**, *48*, 7445–7456.

(76) Lee, Y. S.; Seo, S. H.; Yang, B. S.; Lee, J. Y. Synthesis and Biological Evaluation of Bis(methoxy methyl)-7,8-dihydro-[1,4]dioxino[2,3-*g*]quinazolines as EGFR tyrosine kinase inhibitors. *Arch. Pharm. (Weinheim, Ger.)* **2005**, *338*, 502–505.

(77) Wissner, A.; Hamann, P. R.; Nilakantan, R.; Greenberger, L. M.; Ye, F.; Rapuano, T. A.; Loganzo, F. Syntheses and EGFR Kinase Inhibitory Activity of 6-Substituted-4-anilino [1,7] and [1,8] Naphthyridine-3-carbonitriles. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1411–1416.

(78) Klutchko, S. R.; Zhou, H.; Winters, R. T.; Tran, T. P.; Bridges, A. J.; Althaus, I. W.; Amato, D. M.; Elliott, W. L.; Ellis, P. A.; Meade, M. A.; Roberts, B. J.; Fry, D. W.; Gonzales, A. J.; Harvey, P. J.; Nelson, J. M.; Sherwood, V.; Han, H. K.; Pace, G.; Smaill, J. B.; Denny, W. A.; Showalter, H. D. Tyrosine kinase Inhibitors. 19. 6-Alkynamides of 4-Anilinoquinazolines and 4-Anilinopyrido[3,4-*d*]pyrimidines as Irreversible Inhibitors of the ErbB Family of Tyrosine Kinase Receptors. *J. Med. Chem.* **2006**, *49*, 1475–1485.

(79) Hennequin, L. F.; Ballard, P.; Boyle, F. T.; Delouvrié, B.; Ellston, R. P.; Halsall, C. T.; Harris, C. S.; Hudson, K.; Kendrew, J.; Pease, J. E.; Ross, H. S.; Smith, P.; Vincent, J. L. Novel 4-Anilinoquinazolines with C-6 Carbon-Linked Side Chains: Synthesis and Structure−Activity Relationship of a Series of Potent, Orally Active, EGF Receptor Tyrosine Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2672–2676.

(80) Tsou, H. R.; Mamuya, N.; Johnson, B. D.; Reich, M. F.; Gruber, B. C.; Ye, F.; Nilakantan, R.; Shen, R.; Discafani, C.; DeBlanc, R.; Davis, R.; Koehn, F. E.; Greenberger, L. M.; Wang, Y. F.; Wissner, A. 6-Substituted-4-(3-bromophenylamino)quinazolines as Putative Irreversible Inhibitors of the Epidermal Growth Factor Receptor (EGFR) and Human Epidermal Growth Factor Receptor (HER-2) Tyrosine Kinases with Enhanced Antitumor Activity. *J. Med. Chem.* **2001**, *44*, 2719–2734.

(81) Smith, L.; Wong, W. C.; Kiselyov, A. S.; Burdzovic-Wizemann, S.; Mao, Y.; Xu, Y.; Duncton, M. A.; Kim, K.; Piatnitski, E. L.; Doody, J. F.; Wang, Y.; Rosler, R. L.; Milligan, D.; Columbus, J.; Balagtas, C.; Lee, S. P.; Konovalov, A.; Hadari, Y. R. Novel Tricyclic Azepine Derivatives: Biological Evaluation of Pyrimido[4,5-*b*]-1,4-benzoxazepines, Thiazepines, and Diazepines as Inhibitors of the Epidermal Growth Factor Receptor Tyrosine Kinase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5102–5106.

(82) Smith, L.; Piatnitski, E. L.; Kiselyov, A. S.; Ouyang, X.; Chen, X.; Burdzovic-Wizemann, S.; Xu, Y.; Pan, W.; Chen, X.; Wang, Y.; Rosler, R. L.; Patel, S. N.; Chiang, H. H.; Milligan, D. L.; Columbus, J.; Wong, W. C.; Doody, J. F.; Hadari, Y. R. Tricyclic Azepine Derivatives: Pyrimido[4,5-*b*]-1,4-benzoxazepines as a Novel Class of Epidermal Growth Factor Receptor Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1643–1646.

(83) Jin, Y.; Li, H. Y.; Lin, L. P.; Tan, J.; Ding, J.; Luo, X.; Long, Y. Q. Synthesis and Antitumor Evaluation of Novel 5-Substituted-4-hydroxy-8-nitroquinazolines as EGFR Signaling-Targeted Inhibitors. *Bioorg. Med. Chem.* **2005**, *13*, 5613–5622.

(84) Gangjee, A.; Yang, J.; Ihnat, M. A.; Kamat, S. Antiangiogenic and Antitumor Agents. Design, Synthesis, and Evaluation of Novel 2-Amino-4-(3-bromoanilino)-6-benzylsubstituted Pyrrolo[2,3-*d*]pyrimidines as Inhibitors of Receptor Tyrosine Kinases. *Bioorg. Med. Chem.* **2003**, *11*, 51555170].

(85) Wissner, A.; Overbeek, E.; Reich, M. F.; Floyd, M. B.; Johnson, B. D.; Mamuya, N.; Rosfjord, E. C.; Discafani, C.; Davis, R.; Shi, X.; Rabindran, S. K.; Gruber, B. C.; Ye, F.; Hallett, W. A.; Nilakantan, R.; Shen, R.; Wang, Y. F.; Greenberger, L. M.; Tsou, H. R. Synthesis and Structure−Activity Relationships of 6,7-Disubstituted 4-Anilinoquinoline-3-carbonitriles. The Design of an Orally Active, Irreversible

Inhibitor of the Tyrosine Kinase Activity of the Epidermal Growth Factor Receptor (EGFR) and the Human Epidermal Growth Factor Receptor-2 (HER-2). *J. Med. Chem.* **2003**, *46*, 49–63.
(86) Sperandio, O.; Andrieu, O.; Miteva, M. A.; Vo, M. Q.; Souaille, M.; Delfaud, F.; Villoutreix, B. O. MED-SuMoLig: A New Ligand-Based Screening Tool for Efficient Scaffold Hopping. *J. Chem. Inf. Model.* **2007**, *47*, 1097–1110.
(87) Lepp, Z.; Kinoshita, T.; Chuman, H. Screening for New Antidepressant Leads of Multiple Activities by Support Vector Machines. *J. Chem. Inf. Model.* **2006**, *46*, 158–167.
(88) Chandregowda, V.; Rao, G. V.; Reddy, G. C. Convergent Approach for Commercial Synthesis of Gefitinib and Erlotinib. *Org. Process Res. Dev.* **2007**, *11*, 813–816.
(89) http://www.invitrogen.com.
(90) Shults, M. D.; Imperiali, B. Versatile Fluorescence Probes of Protein Kinase Activity. *J. Am. Chem. Soc.* **2003**, *125*, 14248–14249.
(91) *SciFinder Scholar*, American Chemical Society: Washington, DC, 2007.

JM800829V