Published in final edited form as:
J Comput Aided Mol Des. 2010 December ; 24(12): 971–991. doi:10.1007/s10822-010-9390-0.

Small molecule correctors of F508del-CFTR discovered by structure-based virtual screening

Ori Kalid,
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel; Department of Biochemistry, George S. Wise Faculty of Life Sciences, Tel-Aviv University, 69978 Ramat Aviv, Israel

Martin Mense,
EPIX Pharmaceuticals Inc, 4 Maguire Road, Lexington, MA 02421, USA

Sharon Fischman,
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Alina Shitrit,
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Hermann Bihler,
EPIX Pharmaceuticals Inc, 4 Maguire Road, Lexington, MA 02421, USA

Efrat Ben-Zeev,
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Nili Schutz,
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Nicoletta Pedemonte,
Advanced Biotechnology Center, L.go R. Benzi 10, 16132 Genoa, Italy

Philip J. Thomas,
Departments of Physiology and Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Robert J. Bridges,
Departments of Physiology and Biophysics, Rosalind Franklin University of Medicine and Science, Chicago Medical School, 3333 Green Bay Road, North Chicago, IL 60064, USA

Correspondence to: Ori Kalid, orikalid@gmail.com.

Present Address: M. Mense, Cystic Fibrosis Foundation Therapeutics, 529 Main Street—Suite 208, Charlestown, MA 02129, USA
Present Address: S. Fischman, Biolojic Design LTD, Tel Aviv, Israel
Present Address: H. Bihler, CFRx LLC, 529 Main Street—Suite 208, Charlestown, MA 02129, USA
Present Address: A. Shitrit, E. Ben-Zeev, N. Schutz, Dynamix Pharmaceuticals Ltd, 13 Gad Feinstein Rd, 3rd Level, Suite 30, 76385 Rehovot, Israel
Present Address: D. R. Wetmore, Emerald BioStructures, 7869 NE Day Rd W, Bainbridge Island, WA 98110, USA
Present Address: Y. Marantz, Teva Pharmaceuticals Industries LTD, Global Inovative R&D, P. O Box 8077, 42504 Netanya, Israel
Present Address: H. Senderowitz, Department of Chemistry, Bar-Ilan University, 52900 Ramat Gan, Israel

Electronic supplementary material The online version of this article (doi:10.1007/s10822-010-9390-0) contains supplementary material, which is available to authorized users.

Diana R. Wetmore,
Cystic Fibrosis Foundation Therapeutics, 6931 Arlington Road, Suite 200, Bethesda, MD 20814, USA

Yael Marantz, and
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Hanoch Senderowitz
EPIX Pharmaceuticals Ltd, 3 Hayetzira Street, 52521 Ramat Gan, Israel

Ori Kalid: orikalid@gmail.com

Abstract

Folding correctors of F508del-CFTR were discovered by in silico structure-based screening utilizing homology models of CFTR. The intracellular segment of CFTR was modeled and three cavities were identified at inter-domain interfaces: (1) Interface between the two Nucleotide Binding Domains (NBDs); (2) Interface between NBD1 and Intracellular Loop (ICL) 4, in the region of the F508 deletion; (3) multi-domain interface between NBD1:2:ICL1:2:4. We hypothesized that compounds binding at these interfaces may improve the stability of the protein, potentially affecting the folding yield or surface stability. In silico structure-based screening was performed at the putative binding-sites and a total of 496 candidate compounds from all three sites were tested in functional assays. A total of 15 compounds, representing diverse chemotypes, were identified as F508del folding correctors. This corresponds to a 3% hit rate, ~tenfold higher than hit rates obtained in corresponding high-throughput screening campaigns. The same binding sites also yielded potentiators and, most notably, compounds with a dual corrector-potentiator activity (dual-acting). Compounds harboring both activity types may prove to be better leads for the development of CF therapeutics than either pure correctors or pure potentiators. To the best of our knowledge this is the first report of structure-based discovery of CFTR modulators.

Keywords

Cystic fibrosis; CFTR; Structure-based virtual screening; Correctors; Potentiators; Modulators; F508; Docking; Homology modeling; YFP; Ussing chamber; Chemical chaperones; Pharmacological chaperones

Introduction

Cystic Fibrosis (CF) is the most common lethal autosomal recessive disorder in Caucasian population, affecting approximately 30,000 people in the United States and ~70,000 worldwide. In the United States, an additional ten million, or about one in every 31 Americans, are carriers of the defective CF gene, but do not express the disease. CF occurs in approximately one of every 3,200 live Caucasian births (http://www.health.state.ny.us/diseases/cystic_fibrosis/faq.htm) and approximately 1,000 new cases of CF are diagnosed each year, usually (>70%) by the age of two (http://www.cff.org/AboutCF/). While there is yet no cure for CF, aggressive treatment including mucus thinners, antibiotics, anti-inflammatories and bronchodilators along with physical therapy and proper nutritional repletion, can lengthen and improve the quality of life of CF patients. Consequently, the

currently predicted median age of survival for CF patients is just over 37 (http://www.cff.org/AboutCF/).

The cystic fibrosis transmembrane conductance regulator (CFTR) glycoprotein, a chloride channel essential in the apical membrane of epithelial cells for ion and fluid homeostasis, is a member of the ATP-binding cassette (ABC) protein family. However, it is functionally distinct, being the only member of this family known to function as an ion channel. Nevertheless, CFTR architecture is similar to other ABC transporters (Fig. 1), comprising two membrane spanning domains (MSDs), each linked, through intracellular loops (ICLs) to a nucleotide binding domain (NBD). Uniquely to CFTR these two repeated motifs are connected by a structurally disordered regulatory domain (R-domain), which requires phosphorylation by PKA in order to allow for channel opening [1]. Thus, channel gating depends on the extent of R domain phosphorylation at multiple sites, reflecting the balance between protein kinases and the phosphatases acting on these sites. Gating is further regulated by association of the NBD domains and the binding and hydrolysis of the ATP ligands, however this mechanism is not fully understood [2].

At the molecular level, CF is caused by mutations in the CFTR gene. Currently, over 1,400 disease causing mutations have been identified that vary in the severity of the resulting disease [3] (http://www.cff.org/AboutCF/). However, the F508del mutation is by far the most common one; at least one allele is found in ~90% of the patient population [3]. The F508del mutation nearly abolishes correct cellular processing of CFTR, probably by reducing the folding efficiency of NBD1 as well as disrupting inter-domain contacts required for the stability of the wt protein [4–6]. Most of the mutant protein is targeted to endoplasmic reticulum-associated degradation (ERAD) and the few mutant channels that make it to the plasma membrane are characterized by a reduced open probability and thermal instability and are rapidly endocytosed and targeted for lysosomal degradation [2].

As chloride conductance by CFTR depends on the number of channels in the cell membrane as well as their gating, mainly two complementary approaches are being pursued as potential treatment for F508del-associated CF. The first involves the development of small molecule ‘correctors’, namely, compounds that would act as chemical chaperones to correct the folding defect of F508del-CFTR, thereby increasing the number of CFTR channels that reach the cell membrane. In principle there are several intervention points within the cellular quality control mechanisms where small molecule interference might lead to improved trafficking of F508del-CFTR to the cell surface, such as the ER associated degradation (ERAD pathway), endocytic trafficking etc., each of which contributes to the low surface expression of F508del-CFTR [2]. However, these mechanisms are non-specific and may lead to undesirable side-effects. In contrast, an “ideal” mechanism of action for correctors would be through direct interaction with the mutant protein and subsequent stabilization of the correctly folded state throughout the maturation and surface trafficking pathways. The viability of such a mechanism of action is supported by recent reports of compounds correcting disease-associated misfolding of GPCRs and enzymes by direct binding to the mutant proteins [7]. The second strategy for increasing F508del-CFTR conductance involves the development of ‘activators’ or ‘potentiators’, namely compounds that correct the defective gating of mutant channels already at the cell surface. While such compounds may

act indirectly by increasing cAMP production or inhibition of regulatory enzymes, many reports have ruled out these non-specific mechanisms suggesting direct interaction with the mutant CFTR [8–15].

Efforts to develop both types of therapeutics have been ongoing for over a decade [16–18]. One of the early high throughput screening (HTS) campaigns for CFTR modulators was described by Van Goor et al. of Vertex Pharmaceuticals, who reported the screening of over 160,000 compounds for F508del-CFTR correction and of over 120,000 compounds for F508del-CFTR potentiation [19]. In this work, six structurally distinct correctors were identified, the most potent was a quinazoline compound, VRT-422, which was further optimized to yield VRT-325. However, both quinazoline compounds were also found to correct the cellular processing of the G601S mutant hERG channel and thus may operate via a target in the biogenesis pathway shared by both proteins. The potentiator screen led to the discovery of the pyrazole compound, VRT-532, as a potent potentiator. Later studies have shown that this compound may in fact be a dual acting corrector-potentiator [20] which may interact directly with the CFTR protein [21]. Development at Vertex eventually resulted in the potentiator VX-770 (Phase-II and Phase-III clinical trials for CF associated with the F508del and G551D mutations, respectively, are ongoing, http://clinicaltrials.gov), and the corrector VX-809 (Phase-IIa clinical trial completed, http://clinicaltrials.gov and http://www.vrtx.com/current-projects/drug-candidates/VX-809.html). The success of the Vertex compounds supports the notion that increasing chloride conductance has a positive effect on CF patients.

In a study by Pedemonte et al. [22] high throughput screening of 150,000 chemically diverse compounds and of more than 1,500 analogs of active compounds yielded several classes of F508del-CFTR correctors. Biochemical studies suggested a mechanism of action involving improved F508del-CFTR folding at the ER as well as increased stability at the cell surface. A class of bisaminomethylbithiazoles, including corrector-4a, was found to improve F508del folding through a potentially specific mechanism [22].

Both screens described above were performed using functional assays. Pedemonte et al. [22] assayed iodide flux in F508del-CFTR—transfected epithelial cells using a fluorescent halide indicator while Van-Goor et al. [19] used a voltage-sensitive assay based on the change in fluorescence resonance energy transfer between a membrane-soluble voltage sensitive dye and a plasma membrane-localized fluorescent coumarin-linked phospholipid in NIH-3T3 cells.

In contrast, Carlile and coworkers developed a HTS assay that does not rely on a functional read-out of CFTR-related ion conductance but rather on CFTR surface expression measured by immunodetection [23]. Several hits were identified in a high throughput screen of over 40,000 compounds, including several compounds previously reported to affect CFTR function or trafficking [24, 25]. One of the correctors identified in the original screen was the PDE5 inhibitor, sildenafil, and a purchased analog of that compound, KM11060, which showed even greater corrector activity. Interestingly, Sildenafil and even more so KM11060 are chemically similar to VRT-325 and VRT-422, which may suggest a similar mechanism of action for these compounds (Fig. 2).

Much progress has been made in the identification and development of both F508del-CFTR correctors and potentiators. However, the exact mechanism of action of the compounds under development has not been elucidated. In particular, there is no clear strategy for designing compounds that act by direct interaction with CFTR, mostly due to the lack of a high resolution 3D structure of the full length protein. However, advances in molecular modeling of CFTR, especially following the disclosure of the Sav1866 crystal structure [26] may provide templates for structure-based design of CFTR modulators. Two models of full length *wt*-CFTR, as well as models of the nucleotide binding domains, have recently been reported. Both Serohijos et al. [27], and Mornon et al. [28] reported models based on Sav1866 which are in agreement with a large volume of experimental data. Serohijos et al. highlight the NBD1:ICL4 interface as a potential target for CFTR modulators and Mornon et al. specifically propose this region as a potential site of action for small molecule correctors based on their modeling of the F508del mutant CFTR. Moran et al. [29] and Huang et al. [30] modeled the NBD1:2 dimer and predicted binding sites for CFTR potentiators. However, to the best of our knowledge, none of these models have been used for structure-based discovery of CFTR modulators.

In order to provide a structural framework for the development of CFTR therapeutics, we have modeled the 3D structure of full length CFTR in its open (i.e., conducting) state [31]. Of particular relevance to drug discovery are the intracellular (i.e., cytosolic) regions of both *wt*- and F508del-CFTR. These regions are comprised of the intracellular loops and the two nucleotide binding domains (NBD1 and NBD2). While putative binding sites for CFTR modulators may in principle also exist within the transmembrane (TM) domain [32], ligand binding to such sites was not considered in the present work. This is supported by the fact that, to the best of our knowledge, a CFTR correction/potentiation effect for any of the numerous CFTR porebinding open channel blockers has not been demonstrated to date. Thus, several models of the CFTR intracellular domains (both *wt* and F508del) were developed independently of the full length model mentioned above, using different modeling and refinement protocols as described below. The resulting models were explored for potential binding sites. Experimental evidence suggests that the F508del mutation reduces the folding efficiency of NBD1 and also interferes with the interactions of NBD1 with other CFTR domains within the cytosolic region [4–6]. Consequently, inter-domain interfaces involving NBD1 were analyzed for potential binding sites, under the assumption that small molecules binding at these sites may stabilize essential interactions between CFTR domains during the folding process consequently acting as CFTR correctors, which were the main focus of the present study. Three distinct sites were identified in the models and each was subjected to in silico screening of commercially available compound libraries using an array of proprietary and commercial computational tools. A set of 100–200 compounds were prioritized from each screening campaign then purchased and tested in cell-based functional assays. The details of this process as well as the properties of several validated hit compounds are reported below.

Results

Modeling and structure-based screening

Our models of the CFTR cytoplasmic region have evolved over time, and each model has been used for in silico screening of new putative binding sites for CFTR correctors. The first model was developed prior to the publication of the Sav1866 crystal structure and was limited to the isolated NBD1:2 dimer. Following the disclosure of the Sav1866 structure, new models were developed that also included the ICLs, modeled by homology to Sav1866 ("Methods"). In silico screening of these models was performed according to the general paradigm described in Fig. 3. However, the exact methodologies varied due to specific challenges associated with each putative binding site.

The following sections describe each of the models along with the respective putative binding sites used for screening, followed by details of the screening results including examples of validated hit compounds.

Model of the NBD1:NBD2 dimer and putative binding site at the dimer interface—CFTR is believed to adopt its conducting state when the NBDs are dimerized [2, 33, 34]. Binding at the NBD1:2 interface and consequent dimer stabilization has been proposed as a possible mechanism of action for CFTR potentiators [29]. We hypothesize that stabilization of the NBD1:2 dimer may also lead to an overall fold stabilization and thereby such molecules may also act as correctors.

A model of the wt NBD1:2 dimer was developed as described in "Methods" and was found consistent with available crosslinking data [33–35]. Cavity mapping with Sybyl MOLCAD [36] revealed a pocket at the NBD1:NBD2 dimer interface. Slight rotamer adjustments were manually performed to increase cavity size making it more suitable for small molecule binding. This was not unanticipated as the dimer model was refined by MD in its "apo" form. The resulting site was mostly linear and largely composed of the following CFTR residues: Q493, S495, S573, F575, S605, and Y577 from NBD1, ATP from NBD1, H1348, A1374, H1375, and R1403 from NBD2 (Fig. 4). In order to evaluate whether the NBD1:2 interface site could accommodate correctors described in the literature, a representative set of known modulators was docked into this putative binding site ("Methods"). Interestingly, while there is no experimental evidence suggesting this as their site of action, a set of known modulators described by Hadida et al. [37] displayed both shape complementarity and specific hydrogen bonding interactions with the dimer interface site. Since these compounds seemed to utilize most of the potential binding site interactions, we decided to use a representative as a "molecular probe" for optimizing binding site conformation in preparation for screening (generating a hypothetical "co-crystal" structure). Thus, compound 20 from Hadida et al. [37] was re-docked into the interface site using an Induced-Fit Docking protocol [38] which enables structural adaptation of the protein in response to ligand binding. The resulting 1.1 Å all-atom RMSD compared to the initial conformation reflects the minor adjustments required for obtaining improved interactions.

The optimized model was subsequently subjected to virtual screening (VS) as described in "Methods". "Screening results" are described below.

First generation model of the complete intracellular domain and putative binding site—Following the disclosure of the Sav1866 crystal structure [26], the NBD1:2 dimer model was extended to include the entire predicted cytoplasmic domain of CFTR. Sav1866 is a 12TM bacterial transporter belonging to the ABC transporters subfamily B, which belongs to the same ABC transporters superfamily as CFTR. However, in contrast to CFTR which consists of a single polypeptide chain, Sav1866 is a dimer of two identical MSD:NBD polypeptide chains. Sav1866 bares a relatively low sequence identity to CFTR (18% sequence identity between the Sav1866 monomer and either MSD1-NBD1 or MSD2-NBD2,<14% in the TM regions) and the outward-facing conformation of its transmembrane helices does not agree with the expected TM conformation of CFTR in its conducting state [39–41]. Still, the structure of the intracellular domains of Sav1866, namely, the ICLs and NBDs, provides a reasonable template for modeling the corresponding regions of CFTR. In addition to the improved sequence identity in the isolated intracellular region (~23%), the choice of Sav1866 as a template is supported by the observed head-to-tail configuration of the Sav1866 NBD dimer, which is also expected for the conducting state of CFTR [34], as well as by the "domain swapped" interaction between Sav1866 ICLs and NBDs which is also expected for CFTR based on crosslinking data [27, 35].

A model of the intracellular domain of *wt* CFTR was developed as described in "Methods" and was found largely consistent with available crosslinking data [27, 33–35]. This model reveals a cluster of aromatic interactions at the NBD1:ICL4 interface, mediated by F508, F1068, F1074, Y1073, and W496. Since the F508del mutation has been suggested to disrupt inter-domain contacts in CFTR [4–6], this model may not be suitable for predicting interaction sites for F508del correctors. Thus, a model of the F508del mutant was similarly developed ("Methods"). A comparison of the wt and F508del models reveals that the NBD1:ICL4 aromatic interaction network is disrupted in the mutant (Fig. 5), resulting in the formation of a cavity at the NBD1:ICL4 interface. This cavity (Fig. 6a) represents a structural deformation of the NBD1:ICL4 interface, which may be related to the proposed destabilization of inter-domain contacts by the F508 deletion. We hypothesize that a small molecule binding at this cavity may lead to correction of CFTR misfolding by effectively mimicking the interactions of the missing F508 side chain.

In support of this hypothesis, it has been demonstrated that the double mutant F508del/R1070W, predicted by our model to improve hydrophobic packing in this region (Fig. 7), improves the trafficking of F508del-CFTR (Philip J. Thomas, personal communication).

The putative F508del binding site at the NBD1:ICL4 interface is composed of the following CFTR residues: R1070, Y1073, F1068, and F1074 from ICL4, R516, Y563, and M498, from NBD1 (Fig. 6b). As for the NBD1:2 interface site, the F508del cavity was probed by docking a representative set of known modulators ("Methods"). However, in this case none of these compounds docked well. To further evaluate the potential of this cavity as binding site for small molecules prior to screening, a random library of 1,000 drug-like compounds from the EPIX database was docked into the binding site and protein–ligand interactions were visually inspected. Encouragingly, results indicated that the F508del site could accommodate diverse drug-like molecules. Subsequent in silico screening was performed as described in "Methods". Screening results are described below.

Second generation model of the intracellular domain and putative binding sites—Following the completion of the F508del site screening, an updated model of the intracellular domain was generated by incorporating the following modifications: a more recent crystal structure of NBD1, inclusion of crystallographic water molecules, minor adjustments to the sequence alignment with Sav1866, and application of a modified refinement protocol ("Methods"). A binding site search using Schrödinger SiteMap [42] exposed a pocket at the multi-domain interface between NBD1, NBD2, ILC1, ICL2 and ICL4 (Fig. 8). Since the loss of inter-domain contacts has been proposed as a source for the F508del folding defect [4–6], this pocket also seems a reasonable interaction site for small molecule correctors.

The multi-domain interface site may be divided into three sub-sites (Fig. 8). The upper portion is composed of residues from four protein domains: I172, I177, and G178 from ICL1, S263 from ICL2, D1341 and F1294 from NBD2 and K1060 from ICL4; The middle part of the site is a narrow aromatic cage formed by three residues, F494 from NBD1, W496 from NBD1 and W1063 from ICL4; The lower section is in contact with the NBD1 ATP molecule, crystallographic water molecules and two residues from ICL1, D173 and S169. This site was also probed by known modulators ("Methods"), and similar to what was observed in the F508del site, no favorable interactions were observed for any of them. Similar to the F508del site, a random library of 1,000 drug-like compounds from the EPIX database was docked to the multi-domain interface site in order to evaluate potential interactions with diverse chemotypes, and visual inspection of the docking results indicated a potential for favorable interactions with diverse drug-like molecules. Virtual screening was performed as described in "Methods" and the results are discussed below.

In vitro screening

In the absence of a direct binding assay to CFTR, compounds were assessed by functional assays. Following virtual screening, a total of 496 compounds were selected for in vitro testing, as described in "Methods". To rapidly evaluate this large set of compounds, in vitro screening commenced with a high throughput iodide flux corrector assay [22] performed in the laboratory of Professor Luis Galietta (Advanced Biotechnology Center, Genova, Italy). Compound activity was assessed at 10 μM in Fisher Rat Thyroid (FRT) cells and in A549 cells, a human alveolar basal epithelial cancer cell line, both stably expressing F508del-CFTR. In either cell line, a compound producing activity greater than three standard deviations (>3 SD) above vehicle levels in each of duplicate measurements was considered a hit. Hit compounds were then tested to confirm functional activity in an Ussing chamber short circuit current (I_{SC}) corrector assay with F508del-CFTR FRT cells at ChanTest (Cleveland, OH) or at EPIX. Compounds that yielded a statistically significant increase in I_{SC} in the combined response to forskolin, IBMX and genistein ("end current") were considered as confirmed corrector hits. Compounds yielding a significantly increased forskolin response compared to vehicle while having little or no effect on the end current were identified as potentiator candidates (see "Data analysis" section in "Methods"). Compounds of interest were further tested for potentiator activity by acute addition in the Ussing chamber assay. Tables 1, 2, and 3 and Figs. 9, 10, and 11 present several confirmed hits from the three binding sites. Following hit confirmation, a small set of chemical analogs

(usually around 30) of each promising hit was purchased in order to generate initial SAR and possibly identify more potent compounds.

Screening results

NBD1:2 interface site—Two hundred and five compounds from the NBD1:NBD2 binding site screening were purchased and tested in the high throughput iodide flux corrector assay [22]. Ten compounds displayed activity >3 SDs above vehicle in the F508del-CFTR FRT cell line and 10 compounds showed activity>3 SDs above vehicle in the F508del-CFTR A549 cell line. Only one compound was found active in both cell lines. This lack of congruence may be accounted for by the use of single concentration screening and differences in sensitivity between cells lines, and is typically observed for other screening campaigns in the Galietta laboratory (Luis Galietta, personal communication). For the purpose of the screening activity that is the subject of this work, an initial hit is defined as a compound that displayed activity >3 SDs above vehicle in either the FRT or the A549 cells. Of the 20 combined initial hits from both cell types, eight compounds were found to have statistically significant corrector activity in the Ussing chamber short circuit current (I_{SC}) corrector assay and were thus considered as confirmed corrector hits (none of these were identified as hits in the A549 cell line). This corresponds to a confirmed hit rate of 3.9% for the NBD1:2 interface site (8/205). Five additional compounds were identified as potentiators after acute addition in the Ussing chamber assay. None of the active compounds were structurally similar to the reference modulator used for binding-site optimization ("Methods").

Seven of the hits, including both correctors and potentiators, were subjected to an analog search from commercial sources, and of those, four yielded additional analogs with corrector activity in the FRT Ussing assay (EPX-106224, EPX-106260, EPX-106303 and EPX-106047). Two of the primary screening hits subjected to analog search, EPX-106303 and EPX-106047 (Fig. 9; Table 1) showed strong potentiation activity that in the case of EPX-106047 was paired with weak corrector activity (at 10 μM EPX-106047 showed a 36% improvement in the total current compared to vehicle). Interestingly, these compounds yielded the two most efficacious corrector analogs, EPX-108380 and EPX-108361 (Fig. 9; Table 1). In fact, EPX-108380, an analog of EPX-106303, was found to be a dual acting corrector-potentiator. When tested at 10 μM, this compound reached 48% of the efficacy of 6 μM Corrector 4a, corresponding to a maximum current approximately 50% greater than vehicle. This was coupled to potentiation activity independently confirmed in an acute cumulative dose response titration indicating an EC_{50} below 1 μM. In contrast, the potentiation activity of EPX-106047 was mostly lost in its analog EPX-108361, achieving at 10 μM a 40% improvement in total current over vehicle. Structures and activities of the confirmed corrector hits, the most efficacious analogs and the potentiator EPX-106303 are presented in Fig. 9 and Table 1.

F508del site at the NBD1:ICL4 interface—One hundred and one compounds from the F508del site screening were tested in the high throughput corrector assay. Fourteen initial hits were identified, of which five compounds were validated as correctors in Ussing chamber, corresponding to a confirmed hit rate of 5.0% (5/101). Three additional

compounds had confirmed potentiator activity when added acutely in the Ussing assay. Two of the confirmed hits, EPX-106428 and EPX-106403, showed significant corrector activity and yielded additional analogs with corrector activity. EPX-106428 produced currents ~26% above vehicle when tested at 10 μM. A number of commercially available analogs of this compound showed increased potency, most notably EPX-106817 and EPX-106779. In the FRT Ussing assay, 10 μM EPX-106817 increased F508del current threefold over vehicle, obtaining 50% of cold-conditioning control (low temperature rescue), in which cells expressing F508del-CFTR are incubated at 27 °C resulting in partial correction of the folding defect [22]. This compound had an EC50 value between 3–10 μM. EPX-106779 was even more active, increasing current threefold over vehicle at 10 μM and reaching 87% of cold-conditioning control, with an EC_{50} ~ 5 μM. The second primary hit, EPX-106403, increased F508del current by ~ 40% compared to vehicle when tested at 10 μM. However, none of the tested analogs of this compound showed improved activity. Structures and activities of the five confirmed hits as well as most efficacious analogs are presented in Fig. 10 and Table 2. Ussing chamber traces for compounds EPX-106817, EPX-106779 and EPX-106209 are provided as supplementary information Figures S1–S3.

Multi-domain interface site (NBD1:2:ICL1:ICL2:ICL4)—One hundred and ninety compounds from the multi-domain interface site screening were tested in a high throughput corrector assay. Of the sixteen identified initial hits, two were validated as correctors in the Ussing chamber assay, corresponding to a hit rate of 1.1% (2/190), and two more compounds were confirmed as potentiators.

The more potent of the two confirmed corrector hits was EPX-107860. When tested in the FRT Ussing assay at 10 μM, this compound increased F508del current to almost twofold the current obtained from vehicle-treated cells, equivalent to ~84% of the correction obtained with 6 μM corrector 4a. Ussing chamber testing of compounds from this binding site also included acute compound addition at 10 μM following initial activation of CFTR with forskolin as a first indication of potentiation activity (Table 3). A significant response to this acute addition was observed for the second confirmed corrector, EPX-107979, and potentiator activity was independently confirmed in an acute cumulative dose response titration. At 10 μM, the forskolin current was sevenfold over vehicle and at the same time the current after forskolin, IBMX and genistein was only twofold greater than that from vehicle-treated cells, effectively corresponding to no genistein response. These are the hallmark activity characteristics of a dual acting compound as opposed to a pure corrector (see “Data analysis” section in “Methods”).

EPX-107979 yielded four analogs with significant corrector activity and seven others characterized as potentiators. Analogs of EPX-107860 included four additional compounds with significant corrector activity and more than ten additional potentiators. Still, EPX-107860 remained the most efficacious commercially available compound obtained from this binding site. Structures and activities of four validated hits, two correctors and two potentiators, are presented in Fig. 11 and Table 3. Ussing chamber traces for compounds EPX-107860 and EPX-107979 are provided as supplementary information Figures S4–S5.

Discussion

The hypothesis that increasing chloride conductance through mutated CFTR channels may result in clinical benefits to CF patients has been at least partially validated by the VX-770 and VX-809 clinical trial results obtained to date. However, in the absence of a high-resolution experimental 3D structure of CFTR, the development of CFTR modulators has had to rely on classical medicinal chemistry approaches leading to a relatively slow development cycle.

In recognition of the potential benefit of rational drug design approaches, we have developed 3D models of the cytoplasmic domains of *wt* and F508del-CFTR to enable structure-based discovery of small molecule correctors. It is generally accepted that the deletion of F508 alters critical inter-domain interaction and consequently only a small fraction of F508del-CFTR channels assume a fold that is permissive for ER exit. It would be anticipated that small molecules that bind to and stabilize inter-domain regions of CFTR may stabilize the permissive structure of F508del-CFTR thus promoting increased maturation of the mutant protein. Hence, our screening efforts were focused on cavities at inter-domain interfaces: NBD1:NBD2, NBD1:ICL4 and NBD1:NBD2:ICL1:ICL2:ICL4.

The three putative binding sites were subjected to separate in silico screening campaigns and a total of 496 compounds were selected for in vitro testing. In silico screening of each of the binding sites gave rise to hits representing multiple scaffolds. From a drug development perspective, identifying multiple, diverse hits is beneficial as it provides multiple options for hit-to-lead development and lead optimization. Hit rates of 1.1–5.0% were obtained for validated corrector hits (Table 4). For modulators, including both correctors and potentiators, hit rates ranged from 2.1 to 7.9% (Table 4) (since a corrector assay involving pre-incubation and subsequent compound washout was used for primary screening, the potentiator hit rate may be an underestimate of their actual percentage among the tested compounds). These can be contrasted with hit rates obtained in high throughput screening campaigns. For example, Pedemonte et al. reported 45 primary corrector hits identified by a high-throughput screening of a library containing 150,000 diverse drug-like compounds, corresponding to a hit rate of 0.03% which may further decline upon hit verification (numbers following verification were not reported) [22]. Van Goor et al. reported the experimental screening of 160,000 compounds in a correction assay and of 120,000 compounds in a potentiation assay leading to hit rates of 0.11 and 0.23% for CFTR correctors and potentiators, respectively [19]. Carlile et al. have reported the screening of 2,000 compounds for CFTR correction yielding a validated hit rate of 0.25% [23]. Thus, hit rates reported in this work are ~tenfold higher than hit rates obtained in high-throughput screening campaigns, lending some credibility to the present structure-based in silico screening approach for the identification of CFTR correctors.

Two primary screening hits from the NBD1:2 interface site, EPX-106303 and EPX-106047 (Fig. 9; Table 1) showed strong potentiation activity coupled with weak corrector activity which was significantly improved in the commercially available analogs EPX-108380 and EPX-108361. In contrast, two of the primary hits from the F508del site, EPX-106428 and EPX-106403, already showed promising corrector activity. EPX-106428 yielded a number

of commercially available analogs with increased potency, most notably EPX-106817 and EPX-106779 (Fig. 10; Table 2). The multi-domain interface site produced two confirmed correctors, EPX-107860 and EPX-107979, which was characterized as dual acting.

Interestingly, all the binding sites subjected to screening yielded CFTR potentiators as well as correctors. Moreover, several of the chemical series (e.g., EPX-106403, EPX-107860, and EPX-106303) were found to harbor the potential for both types of activities, with small chemical modifications independently modulating the level of correction and potentiation. Notably, several compounds were found to have significant dual corrector-potentiator activity, especially EPX-107979 which was identified from the multi-domain interface site. One could argue that this is due to the fact that the CFTR model used for screening was constructed to represent the conducting state of the channel and that stabilizing this state by direct binding of small molecules may increase the open probability of the channel (potentiation), improve the stability of the protein (potentially affecting the folding yield or surface stability of the protein, i.e., correction), or both. Due to their synergistic activity, dual acting compounds may prove to be superior to pure correctors or pure potentiators as leads for the development of CF therapeutics.

Analogs of several primary hits from the various sites were potent enough to be detected in Band-C Western blots, as shown in Fig. 12 for EPX-106817 (Band-C refers to the band corresponding to the complex-glycosylated form of the CFTR protein). This observation supports the strategy behind the in-silico methods described here, namely that small molecule ligands that are predicted to bind to F508del-CFTR would stabilize and lead to improved maturation of the protein.

None of the primary hits or their commercially available close analogs was potent enough to show corrector activity in human Bronchial Epithelial (hBE) cells [19] (data not shown). However, compared to functional assays in FRT cells, hBE cell assays require a higher level of corrector activity in order to obtain a positive signal. It should not be generally expected that the required leap in activity compared to primary hits would be attained with an initial set of commercially available analogs. It remains to be seen whether further development of these scaffolds would yield compounds active in hBE cells.

Taken together, the high confirmed hit ratio (in comparison to traditional cell based screening methods) for each of the three putative binding sites, and the experimental evidence for F508del-CFTR maturation exemplified in Fig. 12 provide confidence in the ability of the approach described herein to identify promising biologically active small molecule modulators of F508del-CFTR chloride channel activity.

Since methods for evaluating direct binding to the full length CFTR or its F508del mutant are currently unavailable, in vitro screening was performed using functional assays. Consequently, it is possible that some of the identified correctors may not bind directly to CFTR but may utilize alternative correction mechanisms including: (1) changes to the chaperone network, e.g., increased expression of rate-limiting chaperones [43], (2) reduction of the ER retention rate of F508del-CFTR, e.g., by lowering the ER calcium concentration [44], (3) reduction of the rate of endocytosis of F508del-CFTR [45] and (4)

phosphodiesterase(PDE)5 inhibition [23–25]. Similarly, “off-target” mechanisms of action have also been suggested for potentiators. One potential mechanism of action is inhibition of PDEs, which down regulates channel activity through enzymatic degradation of cAMP [9, 15, 46]. However, there is evidence that some PDE inhibitors potentiate CFTR through a non-PDE related mechanism, possibly by direct interaction with CFTR [14]. An alternative mechanism is the inhibition of phosphatases leading to increased phosphorylation of the CFTR R-domain [46]. A potentially direct, non PDE- or phosphoatase-dependent, mechanism of action has been suggested for several classes of potentiators [8, 9, 12, 13, 15]. Some non-specific potentiators that inhibit PDE4 may be identified in Ussing chamber assays if they substantially decrease the IBMX effect on I_{SC}. It is important to note that while the rationale for all of the proposed mechanisms of correction and potentiation is sound, none have been fully validated.

Despite the success of our structure-based approach for identifying F508del correctors we would like to point out several potential inherent limitations including firstly, low model resolution; Our CFTR models rely on homology to a bacterial ABC transporter for which a crystal structure is available. Due to the relatively low target-template sequence identity, the quality of the resulting model may not be optimal. Nevertheless, our models are largely consistent with experimental data and resemble two other CFTR homology models recently published [27, 28]. A second limitation is that of model suitability for identifying CFTR correctors; virtual screening was performed under the assumption that the modeled binding sites occur in the fully-folded mature conformation of CFTR. Thus, correctors identified by our screening process are expected to act by stabilizing the fully folded protein, which is not necessarily the optimal mechanism of action. It is conceivable that stabilization of intermediate states along the folding pathway, rather than of the end point conformation may provide a better approach to assist in CFTR folding. However, modeling such intermediate states is currently not feasible. A third limitation is related to the location of binding sites, as all binding sites used for screening are located at inter-domain interfaces that are believed to be dynamic [27, 35], and consequently not optimally adjusted for small molecule binding (as are the binding sites for small molecule activators of cellular receptor, for example). This issue may have been partly addressed by the use of several models and the conformational exploration by MD. Introducing protein flexibility into the docking process may be beneficial for such binding sites, but was not part of the screening procedures reported herein. Finally, as mentioned above, direct binding of known CFTR modulators to CFTR has not been demonstrated unequivocally. In the absence of such data, it is difficult to conclusively validate and optimize in silico methodologies for structure-based identification of F508del correctors.

Conclusions and future direction

In this work we presented the structure-based discovery of small molecule modulators of F508del-CFTR, including correctors, potentiators and dual-acting compounds. In-silico screening directed at three putative binding sites yielded hit rates between 1 and 5%, ~tenfold higher than corresponding HTS campaigns, and all binding sites gave rise to multiple hits representing diverse chemotypes. The most active compounds obtained from the current screening campaign originated from a putative binding site created by a local

disruption of inter-domain interactions in the vicinity of the missing F508 side-chain. EPX-106817 had an EC_{50} between 3 and 10 μM, and in FRT cells at 10 μM obtained 50% of cold-conditioning control, whereas EPX-106779 obtained 87% of cold-conditioning control at 10 μM and had an EC_{50} ~ 5 μM.

Methods

Molecular modeling

Construction of the NBD1:2 dimer model—The NBD1 model (residues 388–643) was based on the 1XMI crystal structure of NBD1 from human CFTR [47]. Solubilizing mutations and the F508A mutation present in 1XMI were mutated back to the wild-type sequence using Prime [48]. The Regulatory Extension (RE) helix, which interferes with dimer assembly, was removed from the structure. NBD2 was modeled by homology to the 1XEF structure of Haemolysin B NBD [49] (Haemolysin B NBD shares 25% sequence identity and 53% sequence similarity with human NBD2). While NBD2 is relatively similar to NBD1 (24% sequence identity, and 46% sequence similarity), it is predicted to have a partly alpha-helical insertion relative to NBD1 (residues 1,275–1,285) which is present in Haemolysin B, making the latter a more suitable template for modeling NBD2. The NBD1:NBD2 dimer was assembled in the canonical head-to-tail orientation observed in the crystal structure of the Haemolysin B NBD homodimer, which likely represents the dimer conformation present in the open channel state of the CFTR [34]. Initial relaxation was performed using Prime side chain refinement, followed by a stochastic dynamics relaxation of interface residues and subsequent refinement of the entire model. Dynamics was carried out in MacroModel using the OPLS_2005 force field and the Generalized-Born Surface-Area (GB/SA) implicit solvent model. The refinement protocol included an initial energy minimization followed by a 200 ps heating stage, 200 ps equilibration and 1 ns production phase.

This model was constructed prior to the release of the F508del NBD1 crystal structure (2BBT). Encouragingly, the limited conformational effect of the F508del mutation observed in 2BBT, which is localized and distant from the interface region, suggests that modeling results would not have been significantly different.

Construction of the first model of the full cytoplasmic domain—This model of the cytoplasmic domain of CFTR includes the ICLs and the NBD1:NBD2 dimer (residue numbers: 148–194(ICL1); 245–300(ICL2); 363–387 (intracellular helical extension of TM6 and connection to NBD1); 388–643(NBD1); 939–985(ICL3); 1,038–1,092(ICL4); 1,155–1,167(intracellular helical extension of TM12); 1,207–1,444(NBD2)). ICLs were modeled by homology to the crystal structure of Sav1866 (2HYD) using Prime. The 1XMI crystal structure of NBD1 was used for modeling the wt protein and the 2BBT crystal structure of human F508del NBD1 was used for modeling the mutant CFTR. In both NBD1 structures, solubilizing mutations present in the crystal structures were mutated back to *wt* and refined using Prime. As in the NBD1:2 dimer model, NBD2 was modeled using Haemolysin B as a template. However, the NBD2 model was updated to include three water molecules from the Haemolysin B crystal (1XEF), based on literature suggesting their involvement in ATP

hydrolysis [49]. In addition, H1402 was rotated towards the bound ATP molecule into an orientation similar to that observed in 1L2T, an NBD structure from the bacterial ABC transporter MJ0976 (this catalytic histidine corresponds to an alanine in 1XEF). NBD1 and NBD2 were superimposed onto the corresponding NBDs of Sav1866 to obtain their orientation relative to ICLs. Prime side chain refinement was performed on this construct except for residues that are either involved in ATP hydrolysis or those that form interactions with Mg^{2+}, ATP and bound water (Q493, W401, Y1219, Q1291 H1402, E1371, K464, K1250, S466, and Q552). The model was then subjected to energy minimization followed by stochastic dynamics simulations in MacroModel [48] using the OPLS_2005 force field. A distance dependent dielectric of 4 was used to roughly simulate solvent effects. Backbone atoms, as well as bound Mg^{2+}, ATP and water molecules were restrained using harmonic positional constraints that were gradually lifted during dynamics in 20 ps intervals, following an initial 100 ps heating phase, allowing for gradual relaxation of the structure. Following relaxation, the system was subjected to a final 500 ps MD run without any constraint on backbone atoms, Mg^{2+}, ATP, and water molecules. To preserve structural integrity in the implicit environment, distances between atoms of residues contacting the Mg^{2+}, ATP and water molecules, as well as phi-psi dihedral angles of helical segments were constrained throughout the entire run (except for proline residues and two residues C-terminal and N-terminal to each proline). The resulting model of F508del-CFTR, largely consistent with crosslinking data [27, 33–35], was used for screening.

Construction of the second model of the full cytoplasmic domain—This model of the cytoplasmic portion of F508del-CFTR (residue numbers: 148–194(ICL1); 245–300(ICL2); 359–387(intracellular helical extension of TM6 and connection to NBD1); 388–636(NBD1); 939–985(ICL3); 1,038–1,092(ICL4); 1,152–1,173 (intracellular helical extension of TM12); 1,207–1,444(NBD2)) was generated using a new refinement protocol, a more recent NBD1 crystal structure and a slightly altered sequence alignment for the intracellular extensions of TM6 and TM12, based on TM helix boundaries determined by our in house PREDICT™ program [50]. The change in sequence alignment relative to the first intracellular domain model did not affect ICLs participating in any of the putative binding sites located in the previous model and used for in silico screening. NBD2 was modeled as before but a protonated form of H1402 was used, based on the proposed catalytic mechanism in Haemolysin B [51]. F508del NBD1 coordinates were taken from a more recent crystal structure of the F508del-NBD1 without solubilizing mutations which was crystallized in a homodimer configuration (2PZF). Crystallographic water molecules were identified by comparison of several crystal structures of NBD1 and other NBDs (2PZF, 2PZE, 2PZG, 1XMI, 1XMJ, 1XEF, 2BBO, 2BBT, 2BBS, 1ROX and 1L2T). Water molecules that were conserved in several structures, had low B-factors, similar to their neighboring residues, and were not exposed to solvent, were considered as crystallographic water molecules [52]. The NBDs were oriented as in Sav1866. Prior to model refinement, a manual rotamer adjustment was made in order to alleviate a clash between Y577 and Q1291 at the NBD1:2 interface. The model was initially subjected to a rapid constrained minimization, designed to relieve initial clashes formed at inter-domain interfaces. MacroModel minimization was performed using the OPLS_2005 force field, a distance dependent dielectric of 4 and a gradient tolerance of 1.0 kJ/mol. Positional harmonic

constraints were applied to water oxygen atoms, Mg^{2+} and ATP; harmonic constraints were used to maintain the distances between the Mg^{2+}, ATP and water molecules and their interacting NBD residues. In addition, backbone phi-psi dihedral angles, as well as the two membrane-proximal turns of the ICL helices were constrained to prevent excessive unrealistic movements. This minimization was followed by Prime side chain refinement of all residues except for those interacting with Mg^{2+}, ATP and water molecules (Q493, W401, K464, S466, Q552, T460, Y1219, Q1291, H1402, E1371 and K1250). Subsequent refinement was performed in GB/SA solvent using the OPLS_2005 force field. Energy minimization was followed by a 100 ps heating stage and subsequent 1 ns stochastic dynamics refinement. In this simulation, all phi-psi dihedral constraints as well as positional constraints on Mg^{2+}, ATP and water molecules were lifted, allowing increased motion in the more realistic environment. Analysis of simulation snapshots revealed a potential binding site at the interface between NBDs and ICLs, as discussed in Results. However, this interface cavity showed significant flexibility during MD and the energy minimized conformation immediately preceding MD seemed the most suitable for screening. Thus, the structure following side chain refinement was subjected to an extended energy minimization using a more stringent convergence tolerance of 0.05 kj/mol and the resulting model was used for screening.

Docking known modulators into the model binding sites—Potential compatibility of known modulators with the various putative binding sites was evaluated by docking the following compounds into each of the putative binding sites using either GOLD [53, 54] or Glide [55]. Docked compounds included correctors from Pedemonte et al. [22] and Hirth et al. [56] and modulators from Hadida et al. [37]. GOLD docking was performed using 7–8 speedup settings and Glide was used in Standard Precision (SP) mode. Compounds were prepared for docking using either ligPrep [48] or as described in Bar-Haim et al. [57].

Library Preparation for screening—Prior to docking, compound libraries were prepared as described by Bar-Haim et al. [57].

NBD1:2-site screening protocol—Prior to docking experiments, the library of ~4-million commercially available compounds was reduced to 150,000 compounds in two steps. Initially, a library of 500,000 compounds was extracted using ranges of 1D chemical properties derived from: (1) An analysis of binding site properties (2) Lipinski Rule-of-Five [58] (3) Properties of known modulators from Hadida et al. [37] that were successfully docked and thus represent chemotypes which may fit this putative binding site (hence termed “reference modulators”). Molecular weight (MW) and number of rotatable bonds ranges were increased to improve diversity, taking into account the resultant library size. Final parameter ranges used for database filtering were: MW: 250–600, H-bond acceptors: 1–4, H-bond donors: 1–5, rotatable bonds: 3–7, aromatic rings: 1–3, positive charge centers: 0, negative charge centers: 0.

In the second stage, Structure Based Focusing (SBF [59]) was used to further filter the library using 3D information (interactions and shape) derived from the binding site including a shape similarity filter with a large tolerance value based on the van der Waals surface of the virtually co-crystallized probe compound (compound shape, size and binding

mode can vary considerably, as reflected by the diversity of the final set of molecules selected for in vitro testing). This procedure reduced the number of compounds from 500,000 to 150,000, a reasonable library size for docking.

The 150,000 compounds from the focused library were docked and scored using GOLD with 7–8 speedup settings. While the docking software produces multiple alternative binding modes for each ligand, the next stages of virtual screening require that a single binding mode be selected for each docked molecule. Since docking software generally do not identify the most “biologically relevant” binding mode as the best ranked solution, other methods for binding mode selection should be applied. In this particular case, we found empirically that binding modes forming the largest number of hydrogen bonds with binding site residues also made the most sense in terms of overall interactions. This was verified for a set of randomly selected library compound with favorable docking scores as well as the reference modulators. Thus, the GOLD H-bond score was used for binding mode selection and applied as a filter excluding from further analysis all compounds without sufficient H-bonds.

To enable further library filtration, compounds were re-scored with the following scoring functions: Accelrys: ligScore1 Drieding, ligScore2 Drieding, PLP1, PLP2, JAIN, PMF [59]; Tripos: D-Score, G-Score, ChemScore, PMF [36]. Of these, ligScore2 Dreiding ranked the reference modulators significantly better than the random library compounds. A cutoff value was selected for this score and used to filter the docked library down to 2,000 compounds.

The remaining 2,000 compounds were clustered based on molecular similarity as described below and Macro-Model eMBrAcE [48] scores were calculated and used for selecting cluster representatives. These representatives were analyzed by visual inspection in the binding site. Two hundred and five compounds were selected for in vitro testing following assessment of favorable interactions (e.g., H-bonds, aromatic interactions) versus unfavorable interactions (e.g., hydrophobic moieties in hydrophilic regions, charge repulsions), conformational strain, diversity and novelty. In addition, overly flexible molecules were deprioritized to avoid large entropic penalties upon binding. Cluster members of “interesting” compounds were also analyzed.

To make sure that our selected compounds were distinct from the reference modulators, molecular similarity distributions were analyzed for a set of compounds comprised of the representative modulators and all compounds selected for in vitro testing following virtual screening. The distribution of pairwise similarity values (see Molecular Similarity Analysis below) within the set of representative modulators was compared to the distribution within our set of selected compounds and to the distribution between the two sets. The result clearly shows that our selected compounds are diverse (EPIX-EPIX average Tanimoto 0.2 ± 0.08) and distinct from the reference modulators (EPIX-REF average Tanimoto 0.2 ± 0.07) which are relatively similar to one another (REF-REF average Tanimoto 0.5 ± 0.1). Moreover, our experimentally validated hits were structurally diverse and unrelated to the reference modulators.

F508del-site screening protocol—A focused library of 450,000 molecules was extracted from our in-house database of ~4-million compounds, using a combination of chemical property ranges derived from binding site analysis, and Structure-Based Focusing. The following ranges of chemical properties were used for database filtering: MW ≤ 350 (due to the relatively small size of the binding site), H-bond acceptors: ≤8, H-bond donors: ≤10, aromatic rings:>0, positive charge centers: 0, negative charge centers: ≤2.

Because this library, is still too large for efficient structure-based screening, we decided to use the Glide docking program in High Throughput Virtual Screening (HTVS) mode, which is designed for rapid albeit less accurate screening of large libraries, preceding docking with the more accurate SP mode. A plausible binding mode was selected based on a combination of Glide EModel score and the ability to occupy at least one of the two hydrophobic pockets observed in the binding site (Fig. 5). Compounds that did not have any binding pose occupying either one of the hydrophobic pockets were discarded, reducing library size to ~100,000 compounds. The remaining compounds were re-docked using Glide in SP mode. Binding mode selection and subsequent library filtration were conducted as before, resulting in a set of ~14,000 compounds. Visual inspection of the different MW and docking score classes revealed that molecules with either MW < 250 or Glide SP Score > −4 did not have significant interactions. These molecules were discarded, further reducing the library to merely 640 compounds.

The final set of molecules was clustered based on molecular similarity as described below. Cluster representatives were analyzed by visual inspection in the binding site as described above and 101 compounds were selected for in vitro testing following.

Multi-domain interface site screening protocol—A focused library of ~100,000 compounds was extracted from the EPIX database of ~4-million commercially available compounds based on a combination of binding site analysis, Lipinski rule of five, and the Veber rules [60]. The following property ranges were used for filtering: MW: 350–600, PSA: <140 $Å^2$, CLogP: 0–5, H-bond acceptors: 1–10, H-bond donors: 1–5, positive charge centers: 0, negative charge centers: 0, aromatic rings: 1–4, aliphatic rings: 0–3, total number of rings: ≤6.

Compounds were docked using GOLD with 7–8 speedup settings. Unlike the F508del site, there was no clear binding mode hypothesis and unlike the NBD1:2 site, H-bonds were not a dominant factor in this case. Thus, binding modes were selected based on a combination of the docking score and interactions found with residues in the binding site using our in-house BMA/SW [61] software. Compounds with low combined scores were excluded from further analysis. Additional filtration was performed by applying an empirical cutoff to the GOLD fitness score following visual inspection of the compounds in the binding site. ~25,000 compounds survived these two filtration steps and were re-scored using the following scoring functions: Acclerys: ligScore1 Drieding, ligScore2 Drieding, PLP1, PLP2, JAIN, PMF04 [62]; Tripos: D-Score, G-Score, ChemScore, PMF [36]; GOLD H-bond score and Gold ChemScore. Visual inspection of compounds from different score ranges could not identify cutoff values clearly separating molecules with “good” versus “bad” interactions, except for a general requirement for sufficient h-bonding interactions. Thus, compounds

with GOLD H-bond scores worse than −1 were eliminated and rank-by-vote consensus scoring [63] was performed using a cutoff value of one standard deviation below the mean for each of the following scores: GOLD ChemScore, GOLD GoldScore, PMF04, Ligscore2 Dreiding, PLP2, Jain. Tripos ChemScore and. Tripos G-Score and D-Score were excluded due to strong correlation with molecule size, which may lead to a high false positive rate. Ligscore1 and PLP1 were excluded due to strong correlation with the newer versions Ligscore2 and PLP2, respectively. Initially, compounds with a consensus score of zero were eliminated, resulting in a list of 7,080 compounds. These were separated into two classes: compounds occupying mostly the top part of the binding site and those also having significant interaction with the bottom part (Fig. 8). The two sets of molecules were individually clustered based on molecular similarity into 500 clusters each using Discovery Studio [62]. Cluster representatives were selected by maximal consensus score and 210 compounds were selected for in vitro testing as described above for the other screening campaigns.

Molecular similarity analysis and clustering—Molecular similarity was calculated using the Multilevel Neighborhoods of Atoms approach [64] implemented in-house. Clustering is based on the following procedure: (1) for a given Tanimoto cutoff value, all compounds are sorted by number of neighbors (2) the compound with the highest number of neighbors is selected as a reference molecule for the first stage of clustering (3) all compounds within cutoff distance of the reference molecule are grouped into a single cluster and the compound with the highest number of neighbors outside the cluster is selected as a reference compound for the next stage of clustering.

Biological assays

FRT cell tissue culture and ussing-chamber assays—Fisher Rat Thyroid (FRT) cells, stably transfected with F508del-CFTR, were obtained from Professor Luis Galietta. FRT cells were cultured according to published protocols [65]. FRT cell monolayers were grown on Snapwell filter inserts (Corning #3801, Corning, NY, USA) and were incubated with compound for 24 h at 37 °C before the Ussing chamber assay. Cells were subsequently transferred to the Physiologic Instruments (Physiologic Instruments, Inc., San Diego, CA, USA) Ussing recording chamber and superfused on the serosal side with a HEPES buffered physiological saline (HB-PS) with composition (in mM): NaCl, 137; KCl, 4.0; CaCl2, 1.8; MgCl2, 1; HEPES, 10; Glucose, 10; pH adjusted to 7.4 with NaOH. In order to create a transepithelial Cl– ion gradient, 10CF-PS (composition in mM: Na-gluconate, 137; KCl, 4; $CaCl_2$, 1.8; $MgCl_2$, 1; HEPES, 10; Glucose, 10; pH adjusted to 7.4 with N-methyl-D-glucamine) is applied at the mucosal side. A Physiologic Instruments VCC MC8-8S epithelial voltage clamp was used to record short circuit current (ISC). Inserts were voltage clamped at 0 mV and the assay was carried out at 27 °C. 10CF-PS solution (5 mL) was added to the mucosal side of the Snapwell filter and HB-PS solution (5 mL) was added to the serosal side of the Snapwell filter insert. After acquisition of at least 10 min of baseline current, agonists (final concentrations: 10 μM forskolin, 100 μM 3-isobutyl-1-methylxanthine [IBMX] and 20 μM genistein) and antagonist (final concentration: 20 μM CFTRinh-172) were applied sequentially and cumulatively at 10–15 min intervals for forskolin and IBMX, ~ 15 min intervals for genistein and CFTRinh-172, to both serosal and

mucosal epithelial surfaces. Agonists were added as 200–1000× stock solution to the solutions bathing the serosal and mucosal sides. Transepithelial resistance was monitored every 20 s with 10 mV voltage steps. In potentiator assays, appropriate volumes from 10 μM compound stock solution in DMSO were added to the mucosal 10CF-PS solution.

Reported FRT cell Ussing chamber data was obtained from assays in which FRT cells were incubated with compound in presence of 1% serum.

For CFTR biological activity assays, typical positive controls run would include either cold-conditioning control, in which cells expressing F508del-CFTR are incubated at 27 °C resulting in partial correction of the folding defect [22], or the reference CFTR corrector corrector-4a [22].

Data analysis—The “end current” after addition of all agonists, forskolin, IBMX, and genistein was considered a measure for correction according to: $I_{\text{macroscopic}} = i \times P_{\text{open}} \times N$, where i is the single channel current amplitude, P_{open} the open probability of a single channel, and N the total number of channels at the cell surface. For the end current to be a valid measure of correction it must be assumed that after all agonist additions, i.e., after the addition of genistein in the protocol above, F508del-CFTR channels assume a constant open probability regardless of the compound used for correction during the 24 h incubation. In this case the macroscopic current is then directly proportional to the total number of channels N. We have no ultimate proof for this assumption but we have consistently found that the current increase upon the addition of 20 μM genistein varies dependent on how much CFTR currents were potentiated prior to the genistein addition. Thus, in our hands the genistein response varies from ~40% of the total current for a straight corrector, e.g., corrector 4a, to virtually no genistein response at all if cells were incubated with a pure potentiator or a potent dual acting compound prior to the genistein addition. For a much reduced genistein responses see data for EPX-108380 (Table 1) or EPX-107979 (Table 3).

Western blots—HeLa cells were transiently transfected with pCMV-CFTR wt or F508del. 24 h post transfection, cells were butyrate-treated (0.5 mM) and shifted to 30 °C. Four hour following butyrate-treatment cells were treated with 20 μM compounds (in DMSO). For Western Blots, 24 h post butyrate-treatment, cells were harvested in RIPA Buffer and lysates run on 7% PAGE for Western Blot analysis of Band B and Band C with the UNC monoclonal NBD2-directed antibody 596 [66].

Assessment of cytotoxicity—An indication of compound cytotoxicity can be obtained from the Ussing chamber assay since the transepithelial resistance (TER) of a FRT cell monolayer will usually be lowered by cytotoxic compounds. Thus, any compound is potentially cytotoxic that after 24 h incubation significantly reduces the cell monolayer TER compared to cells incubated with vehicle only (DMSO). General cytotoxicity for compounds of interest may then be further investigated in a specific cytotoxicity assay: HepG2 cells (ATCC, Manassas, VA, USA) were grown according to the ATCC-recommended culture conditions. The ATCC recommended medium is made up with phenol-red free medium and was supplemented with 1 mM L-glutamine, penicillin (50 U/mL) and streptomycin (0.05 mg/mL). Cells were plated at 12,000 cells/well in a 96-well plate, and 2 days after plating

compound was added so that by serial solution a suitable concentration range is covered. After 48 h incubation cytotoxicity was measured according to the protocol for the Promega CellTiter 96® Aqueous Non-Radioactive Cell Proliferation Assay (Promega, Madison, WI, USA).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by Cystic Fibrosis Foundation Therapeutics. The authors would like to thank Luis Galietta and ChanTest, Inc. for their contribution to the in vitro screening presented in this paper. Special thanks to Melissa Ashlock, Julie Forman-Kay, David Gadsby, and Kevin Foskett for insightful discussions and advice. Excellent technical assistance was provided by Jan Harrington, Linda Coulson, Katelyn Cassidy and Linda Millen.

References

1. Gadsby DC, Nairn AC. Control of cftr channel gating by phosphorylation and nucleotide hydrolysis. Physiol Rev. 1999; 79(1 Suppl):S77–S107. [PubMed: 9922377]
2. Riordan JR. Cftr function and prospects for therapy. Annu Rev Biochem. 2008; 77:701–726. [PubMed: 18304008]
3. Davis PB. Cystic fibrosis since 1938. Am J Respir Crit Care Med. 2006; 173(5):475–482. [PubMed: 16126935]
4. Du K, Lukacs GL. Cooperative assembly and misfolding of cftr domains in vivo. Mol Biol Cell. 2009; 20(7):1903–1915. [PubMed: 19176754]
5. Du K, Sharma M, Lukacs GL. The deltaf508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of cftr. Nat Struct Mol Biol. 2005; 12(1):17–25. [PubMed: 15619635]
6. Thibodeau PH, Brautigam CA, Machius M, Thomas PJ. Side chain and backbone contributions of phe508 to cftr folding. Nat Struct Mol Biol. 2005; 12(1):10–16. [PubMed: 15619636]
7. Conn PM, Ulloa-Aguirre A, Ito J, Janovick JA. G protein-coupled receptor trafficking in health and disease: lessons learned to prepare for therapeutic mutant rescue in vivo. Pharmacol Rev. 2007; 59(3):225–250. [PubMed: 17878512]
8. Pedemonte N, Sonawane ND, Taddei A, Hu J, Zegarra-Moran O, Suen YF, Robins LI, Dicus CW, Willenbring D, Nantz MH, Kurth MJ, Galietta LJ, Verkman AS. Phenylglycine and sulfonamide correctors of defective delta f508 and g551d cystic fibrosis transmembrane conductance regulator chloride-channel gating. Mol Pharmacol. 2005; 67(5):1797–1807.10.1124/mol.105.010959 [PubMed: 15722457]
9. Becq F. On the discovery and development of cftr chloride channel activators. Curr Pharm Des. 2006; 12(4):471–484. [PubMed: 16472140]
10. Verkman AS, Lukacs GL, Galietta LJ. Cftr chloride channel drug discovery–inhibitors as antidiarrheals and activators for therapy of cystic fibrosis. Curr Pharm Des. 2006; 12(18):2235–2247. [PubMed: 16787252]
11. Springsteel MF, Galietta LJ, Ma T, By K, Berger GO, Yang H, Dicus CW, Choung W, Quan C, Shelat AA, Guy RK, Verkman AS, Kurth MJ, Nantz MH. Benzoflavone activators of the cystic fibrosis transmembrane conductance regulator: towards a pharmacophore model for the nucleotide-binding domain. Bioorg Med Chem. 2003; 11(18):4113–4120. [PubMed: 12927873]
12. Galietta LJ, Springsteel MF, Eda M, Niedzinski EJ, By K, Haddadin MJ, Kurth MJ, Nantz MH, Verkman AS. Novel cftr chloride channel activators identified by screening of combinatorial libraries based on flavone and benzoquinolizinium lead compounds. J Biol Chem. 2001; 276(23):19723–19728. [PubMed: 11262417]

13. Yang H, Shelat AA, Guy RK, Gopinath VS, Ma T, Du K, Lukacs GL, Taddei A, Folli C, Pedemonte N, Galietta LJ, Verkman AS. Nanomolar affinity small molecule correctors of defective delta f508-cftr chloride channel gating. J Biol Chem. 2003; 278(37):35079–35085. [PubMed: 12832418]
14. Chappe V, Mettey Y, Vierfond JM, Hanrahan JW, Gola M, Verrier B, Becq F. Structural basis for specificity and potency of xanthine derivatives as activators of the cftr chloride channel. Br J Pharmacol. 1998; 123(4):683–693. [PubMed: 9517388]
15. Ma T, Vetrivel L, Yang H, Pedemonte N, Zegarra-Moran O, Galietta LJ, Verkman AS. High-affinity activators of cystic fibrosis transmembrane conductance regulator (cftr) chloride conductance identified by high-throughput screening. J Biol Chem. 2002; 277(40):37235–37241. [PubMed: 12161441]
16. Hwang TC, Wang F, Yang IC, Reenstra WW. Genistein potentiates wild-type and delta f508-cftr channel activity. Am J Physiol. 1997; 273(3 Pt 1):C988–C998. [PubMed: 9316420]
17. Brown CR, Hong-Brown LQ, Biwersi J, Verkman AS, Welch WJ. Chemical chaperones correct the mutant phenotype of the delta f508 cystic fibrosis transmembrane conductance regulator protein. Cell Stress Chaperones. 1996; 1(2):117–125. [PubMed: 9222597]
18. Haws CM, Nepomuceno IB, Krouse ME, Wakelee H, Law T, Xia Y, Nguyen H, Wine JJ. Delta f508-cftr channels: kinetics, activation by forskolin, and potentiation by xanthines. Am J Physiol. 1996; 270(5 Pt 1):C1544–C1555. [PubMed: 8967457]
19. Van Goor F, Straley KS, Cao D, Gonzalez J, Hadida S, Hazlewood A, Joubran J, Knapp T, Makings LR, Miller M, Neuberger T, Olson E, Panchenko V, Rader J, Singh A, Stack JH, Tung R, Grootenhuis PD, Negulescu P. Rescue of deltaf508-cftr trafficking and gating in human cystic fibrosis airway primary cultures by small molecules. Am J Physiol Lung Cell Mol Physiol. 2006; 290(6):L1117–L1130. [PubMed: 16443646]
20. Wang Y, Bartlett MC, Loo TW, Clarke DM. Specific rescue of cystic fibrosis transmembrane conductance regulator processing mutants using pharmacological chaperones. Mol Pharmacol. 2006; 70(1):297–302. [PubMed: 16624886]
21. Wellhauser L, Chiaw PK, Pasyk S, Li C, Ramjeesingh M, Bear CE. A small-molecule modulator interacts directly with deltaphe508-cftr to modify its ATPase activity and conformational stability. Mol Pharmacol. 2009; 75(6):1430–1438. [PubMed: 19339490]
22. Pedemonte N, Lukacs GL, Du K, Caci E, Zegarra-Moran O, Galietta LJ, Verkman AS. Small-molecule correctors of defective deltaf508-cftr cellular processing identified by high-throughput screening. J Clin Invest. 2005; 115(9):2564–2571. [PubMed: 16127463]
23. Carlile GW, Robert R, Zhang D, Teske KA, Luo Y, Hanrahan JW, Thomas DY. Correctors of protein trafficking defects identified by a novel high-throughput screening assay. Chem-biochem. 2007; 8(9):1012–1020.
24. Dormer RL, Harris CM, Clark Z, Pereira MM, Doull IJ, Norez C, Becq F, McPherson MA. Sildenafil (viagra) corrects deltaf508-cftr location in nasal epithelial cells from patients with cystic fibrosis. Thorax. 2005; 60(1):55–59. [PubMed: 15618584]
25. Robert R, Carlile GW, Pavel C, Liu N, Anjos SM, Liao J, Luo Y, Zhang D, Thomas DY, Hanrahan JW. Structural analog of sildenafil identified as a novel corrector of the f508del-cftr trafficking defect. Mol Pharmacol. 2008; 73(2):478–489. [PubMed: 17975008]
26. Dawson RJ, Locher KP. Structure of a bacterial multidrug abc transporter. Nature. 2006; 443(7108):180–185. [PubMed: 16943773]
27. Serohijos AW, Hegedus T, Aleksandrov AA, He L, Cui L, Dokholyan NV, Riordan JR. Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the cftr 3d structure crucial to assembly and channel function. Proc Natl Acad Sci U S A. 2008; 105(9):3256–3261. [PubMed: 18305154]
28. Mornon JP, Lehn P, Callebaut I. Atomic model of human cystic fibrosis transmembrane conductance regulator: membrane-spanning domains and coupling interfaces. Cell Mol Life Sci. 2008; 65(16):2594–2612. [PubMed: 18597042]
29. Moran O, Galietta LJ, Zegarra-Moran O. Binding site of activators of the cystic fibrosis transmembrane conductance regulator in the nucleotide binding domains. Cell Mol Life Sci. 2005; 62(4):446–460. [PubMed: 15719171]

30. Huang SY, Bolser D, Liu HY, Hwang TC, Zou X. Molecular modeling of the heterodimer of human cftr's nucleotide-binding domains using a protein-protein docking approach. J Mol Graph Model. 2009; 27(7):822–828. [PubMed: 19167254]

31. Senderowitz H, Fischman S, Kalid O, Sela I, Shitrit A, Strajbl M, Marantz Y. Modeling the three-dimensional structure of cftr. Pediatr Pulmonol. 2007; 54

32. Wang Y, Loo TW, Bartlett MC, Clarke DM. Correctors promote maturation of cystic fibrosis transmembrane conductance regulator (cftr)-processing mutants by binding to the protein. J Biol Chem. 2007; 282(46):33247–33251. [PubMed: 17911111]

33. Mense M, Vergani P, White DM, Altberg G, Nairn AC, Gadsby DC. In vivo phosphorylation of cftr promotes formation of a nucleotide-binding domain heterodimer. EMBO J. 2006; 25(20): 4728–4739. [PubMed: 17036051]

34. Vergani P, Lockless SW, Nairn AC, Gadsby DC. Cftr channel opening by atp-driven tight dimerization of its nucleotide-binding domains. Nature. 2005; 433(7028):876–880. [PubMed: 15729345]

35. He L, Aleksandrov AA, Serohijos AW, Hegedus T, Aleksandrov LA, Cui L, Dokholyan NV, Riordan JR. Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (cftr) mediate regulation of channel gating. J Biol Chem. 2008; 283(39): 26383–26390. [PubMed: 18658148]

36. Sybyl Tripos Inc. Hanley Road, St. Louis, MO 63144: 1699.

37. Hadida RS, Hazelwood AR, Grootenhuis PDJ, Zhou J. Modulators of atp-binding cassette transporters. 2006

38. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem. 2006; 49(2):534–553. [PubMed: 16420040]

39. Linsdell P. Location of a common inhibitor binding site in the cytoplasmic vestibule of the cystic fibrosis transmembrane conductance regulator chloride channel pore. J Biol Chem. 2005; 280(10): 8945–8950.10.1074/jbc.M414354200 [PubMed: 15634668]

40. Linsdell P. Mechanism of chloride permeation in the cystic fibrosis transmembrane conductance regulator chloride channel. Exp Physiol. 2006; 91(1):123–129.10.1113/expphysiol.2005.031757 [PubMed: 16157656]

41. Smith SS, Liu X, Zhang ZR, Sun F, Kriewall TE, McCarty NA, Dawson DC. Cftr: covalent and noncovalent modification suggests a role for fixed charges in anion conduction. J Gen Physiol. 2001; 118(4):407–431. [PubMed: 11585852]

42. Halgren TA. Identifying and characterizing binding sites and assessing druggability. J Chem Inf Model. 2009; 49(2):377–389. [PubMed: 19434839]

43. Singh OV, Pollard HB, Zeitlin PL. Chemical rescue of deltaf508-cftr mimics genetic repair in cystic fibrosis bronchial epithelial cells. Mol Cell Proteomics. 2008; 7(6):1099–1110. [PubMed: 18285607]

44. Norez C, Antigny F, Becq F, Vandebrouck C. Maintaining low ca2+ level in the endoplasmic reticulum restores abnormal endogenous f508del-cftr trafficking in airway epithelial cells. Traffic. 2006; 7(5):562–573. [PubMed: 16643279]

45. Gentzsch M, Chang XB, Cui L, Wu Y, Ozols VV, Choudhury A, Pagano RE, Riordan JR. Endocytic trafficking routes of wild type and deltaf508 cystic fibrosis transmembrane conductance regulator. Mol Biol Cell. 2004; 15(6):2684–2696. [PubMed: 15075371]

46. Galietta LJV, Moran O. Identification of cftr activators and inhibitors: chance or design? Curr Opin Pharmacol. 2004; 4(5):497–503. [PubMed: 15351355]

47. Lewis HA, Zhao X, Wang C, Sauder JM, Rooney I, Noland BW, Lorimer D, Kearins MC, Conners K, Condon B, Maloney PC, Guggino WB, Hunt JF, Emtage S. Impact of the deltaf508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure. J Biol Chem. 2005; 280(2):1346–1353. [PubMed: 15528182]

48. Schrodinger, llc, portland or. Schrodinger, LLC, Portland OR

49. Zaitseva J, Jenewein S, Oswald C, Jumpertz T, Holland IB, Schmitt L. A molecular understanding of the catalytic cycle of the nucleotide-binding domain of the abc transporter hlyb. Biochem Soc Trans. 2005; 33(Pt 5):990–995. [PubMed: 16246029]

50. Shacham S, Marantz Y, Bar-Haim S, Kalid O, Warshaviak D, Avisar N, Inbal B, Heifetz A, Fichman M, Topf M, Naor Z, Noiman S, Becker OM. Predict modeling and in-silico screening for g-protein coupled receptors. Proteins. 2004; 57(1):51–86. [PubMed: 15326594]
51. Hanekop N, Zaitseva J, Jenewein S, Holland IB, Schmitt L. Molecular insights into the mechanism of atp-hydrolysis by the nbd of the abc-transporter hlyb. FEBS Lett. 2006; 580(4):1036–1041. [PubMed: 16330029]
52. Koellner G, Kryger G, Millard CB, Silman I, Sussman JL, Steiner T. Active-site gorge and buried water molecules in crystal structures of acetylcholinesterase from torpedo californica. J Mol Biol. 2000; 296(2):713–735. [PubMed: 10669619]
53. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol. 1995; 245(1):43–53. [PubMed: 7823319]
54. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol. 1997; 267(3):727–748. [PubMed: 9126849]
55. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004; 47(7): 1739–1749. [PubMed: 15027865]
56. Hirth BH, Qiao S, Cuff LM, Cochran BM, Pregel MJ, Gregory JS, Sneddon SF, Kane JL Jr. Discovery of 1, 2, 3, 4-tetrahydroisoquinoline-3-carboxylic acid diamides that increase cftr mediated chloride transport. Bioorg Med Chem Lett. 2005; 15(8):2087–2091. [PubMed: 15808474]
57. Bar-Haim S, Aharon A, Ben-Moshe T, Marantz Y, Senderowitz H. Selex-cs: a new consensus scoring algorithm for hit discovery and lead optimization. J Chem Inf Model. 2009; 49(3):623–633. [PubMed: 19231809]
58. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 1997; 23(1–3):3–25.
59. Cerius2. 4.11. Accelrys, Inc.; San Diego: 2008.
60. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002; 45(12):2615–2623. [PubMed: 12036371]
61. Sela I, Golan G, Strajbl M, Rivenzon-Segal D, Bar-Haim S, Bloch I, Inbal B, Shitrit A, Ben-Zeev E, Fichman M, Markus Y, Marantz Y, Senderowitz H, Kalid O. G protein coupled receptors -in silico drug discovery and design. Curr Top Med Chem. 10(6):638–656. [PubMed: 20337589]
62. Discovery studio. Version 2.1. Accelrys, Inc; San Diego: 2008.
63. Wang R, Wang S. How does consensus scoring work for virtual library screening? An idealized computer experiment. J Chem Inf Comput Sci. 2001; 41(5):1422–1426. [PubMed: 11604043]
64. Filimonov D, Poroikov V, Borodina Y, Gloriozova T. Chemical similarity assessment through multilevel neighborhoods of atoms: definition and comparison with the other descriptors. J Chem Inf Comput Sci. 1999; 39(4):666–670.
65. Zegarra-Moran O, Romio L, Folli C, Caci E, Becq F, Vierfond JM, Mettey Y, Cabrini G, Fanen P, Galietta LJ. Correction of g551d-cftr transport defect in epithelial monolayers by genistein but not by cpx or mpb-07. Br J Pharmacol. 2002; 137(4):504–512. [PubMed: 12359632]
66. Gentzsch M, Choudhury A, Chang XB, Pagano RE, Riordan JR. Misassembled mutant deltaf508 cftr in the distal secretory pathway alters cellular lipid trafficking. J Cell Sci. 2007; 120(Pt 3):447–455. [PubMed: 17213331]

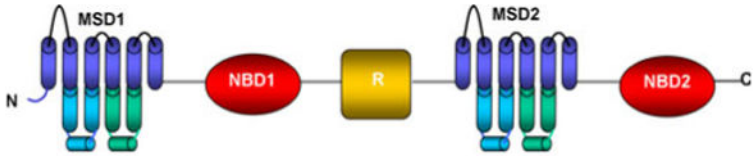

Fig. 1.

Modular organization of CFTR. CFTR is composed of two membrane spanning domains (MSD1 and MSD2), each comprised of six transmembrane segments connected by intracellular loops, as well as two nucleotide binding domains (NBD1 and NBD2) and a regulatory domain (R domain)

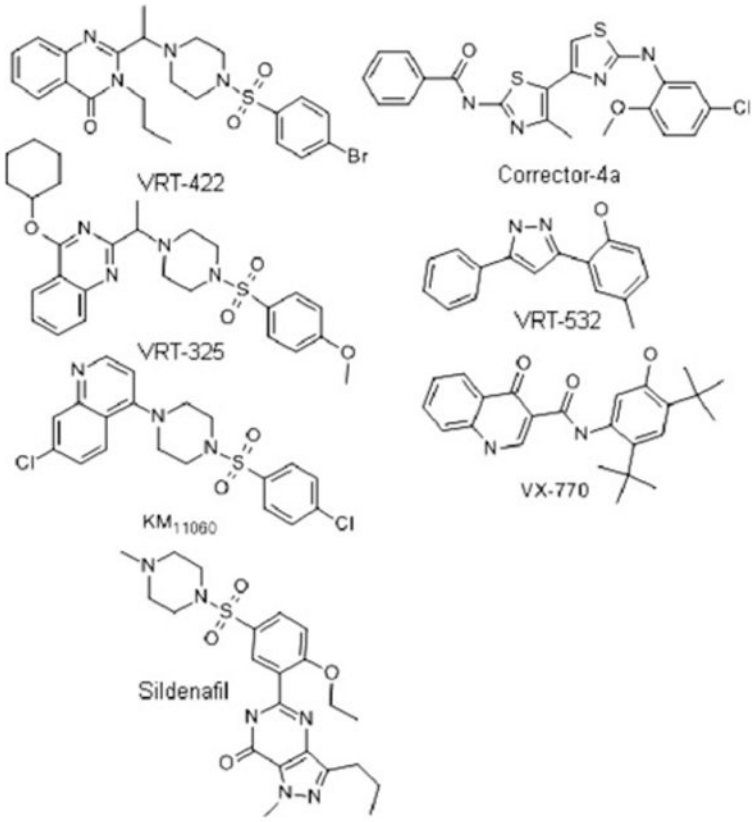

Fig. 2. Known CFTR modulators

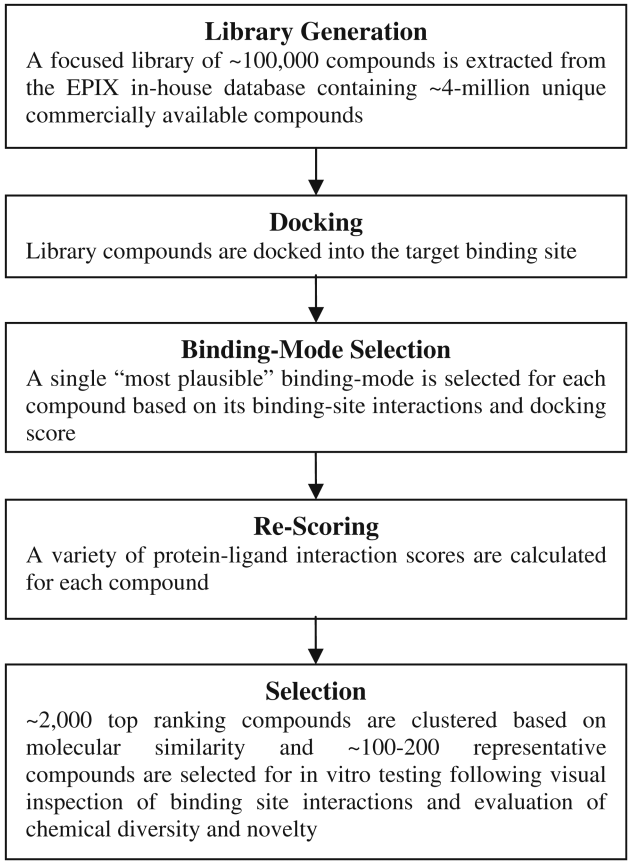

Fig. 3. Generic virtual screening workflow

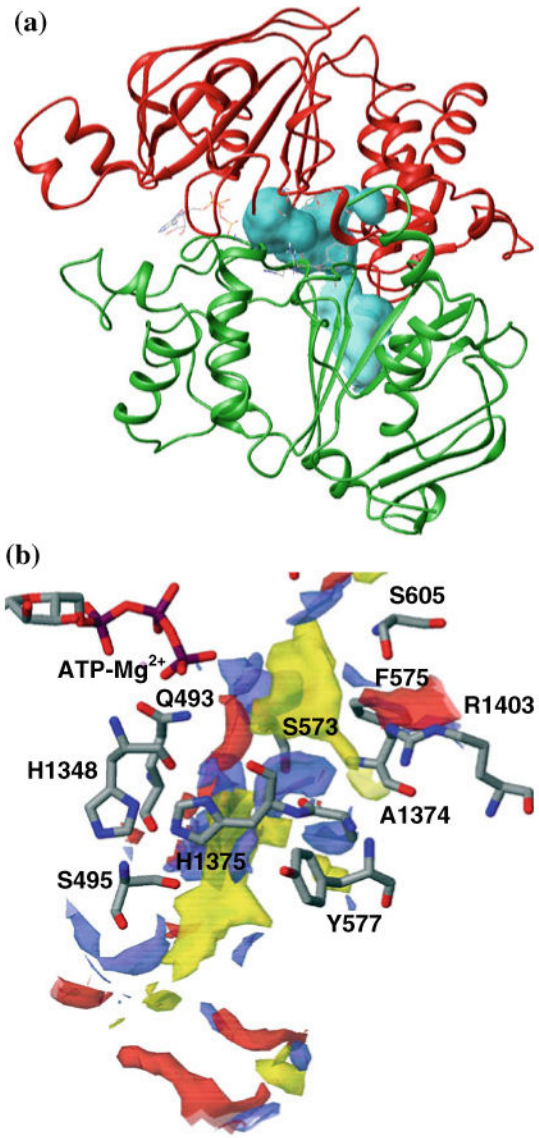

Fig. 4.

NBD1:NBD2 interface binding site. **a** MOLCAD [36] surface showing the shape and size of the interface-site. NBD1 is shown in *red*; NBD2 is shown in *green*. **b** Binding site residues (stick representation) and binding site interaction regions generated with SiteMap. *Yellow* hydrophobic region, *blue* H-bond donor region; and red: H-bond acceptor region

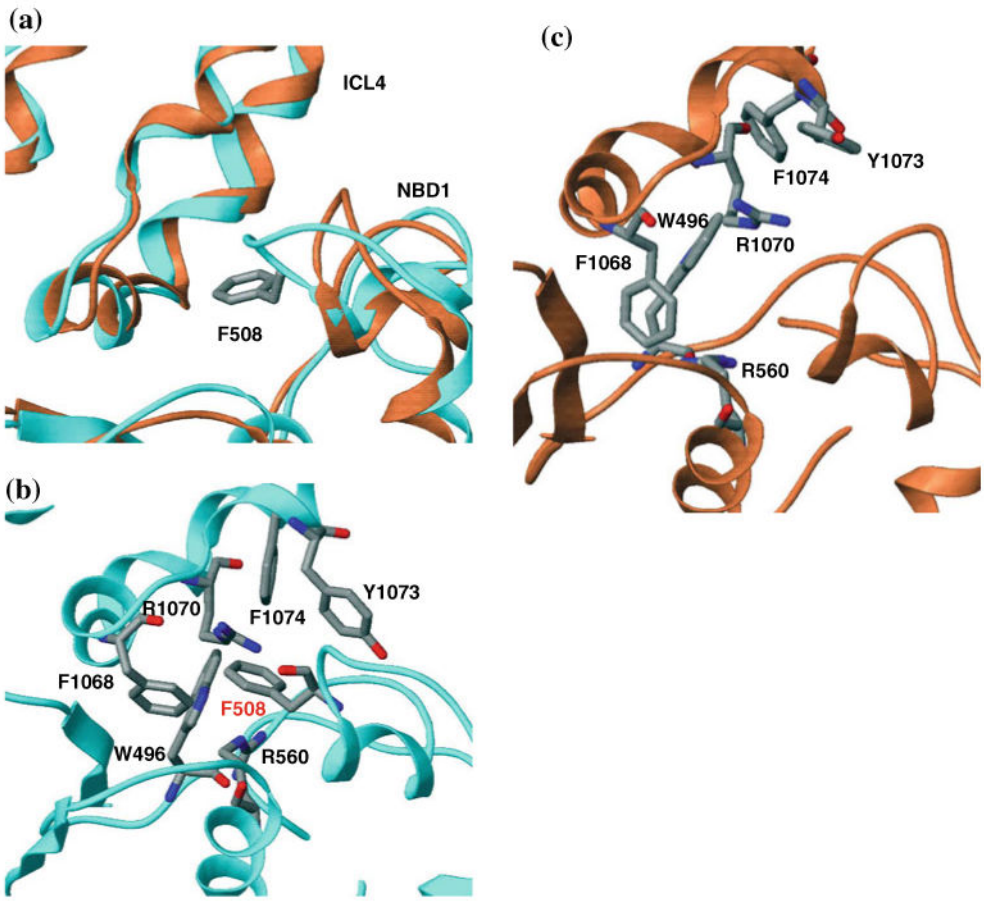

Fig. 5.

Comparison of the ICL4:NBD1 interaction network in the wt and F508del models of the CFTR intracellular domains. **a** Overlay of *wt* (cyan) and F508del (*orange*) models. The deletion of F508 and the resulting alteration of the loop conformation affect the interaction of NBD1 with ICL4. **b** The F508 network of aromatic interactions in the *wt* model. **c** Loss of aromatic interactions in the F508del model

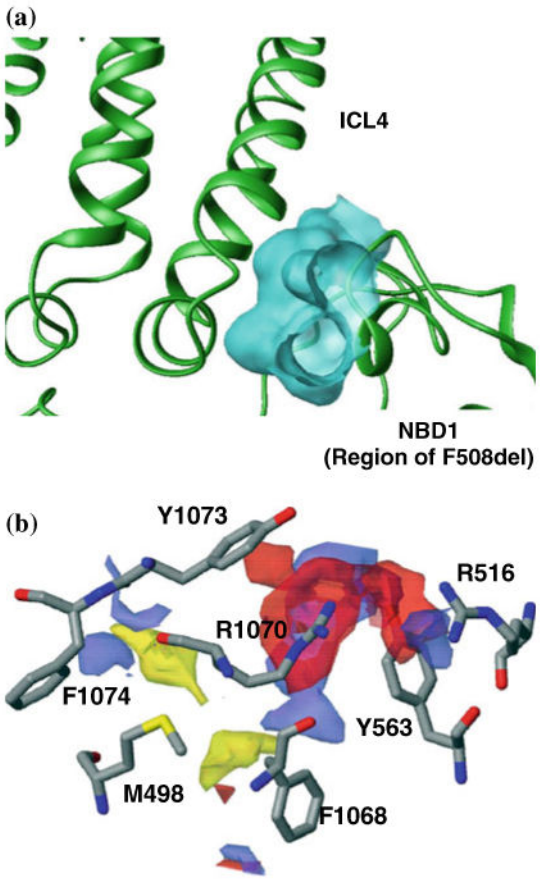

Fig. 6.

F508del binding site (NBD1:ICL4 interface). **a** MOLCAD surface of the interface site. **b** Binding site residues (stick representation) and binding site interaction regions generated with SiteMap. *Yellow* hydrophobic region, *blue* H-bond donor region, and *red* H-bond acceptor region

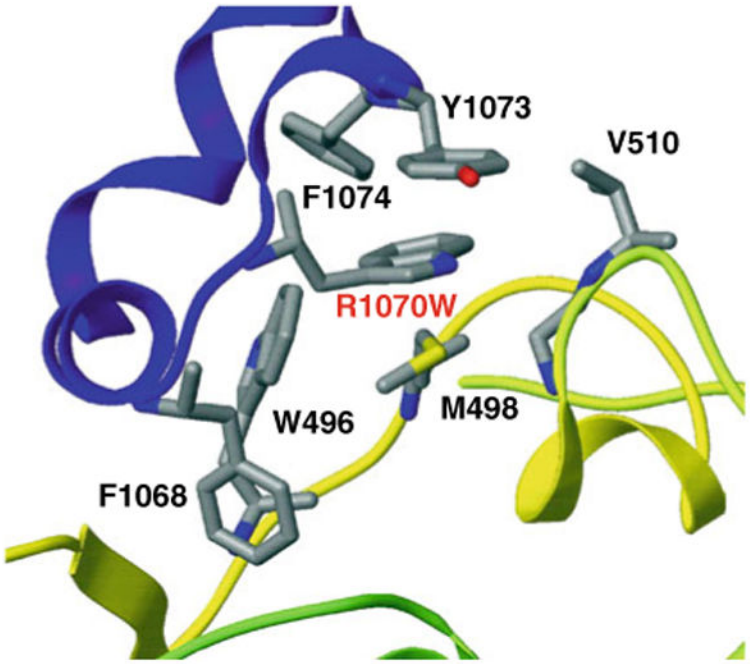

Fig. 7. In silico generation of the R1070W mutant in the F508del model suggests that a tryptophan in this position may restore interactions lost in the F508del mutant

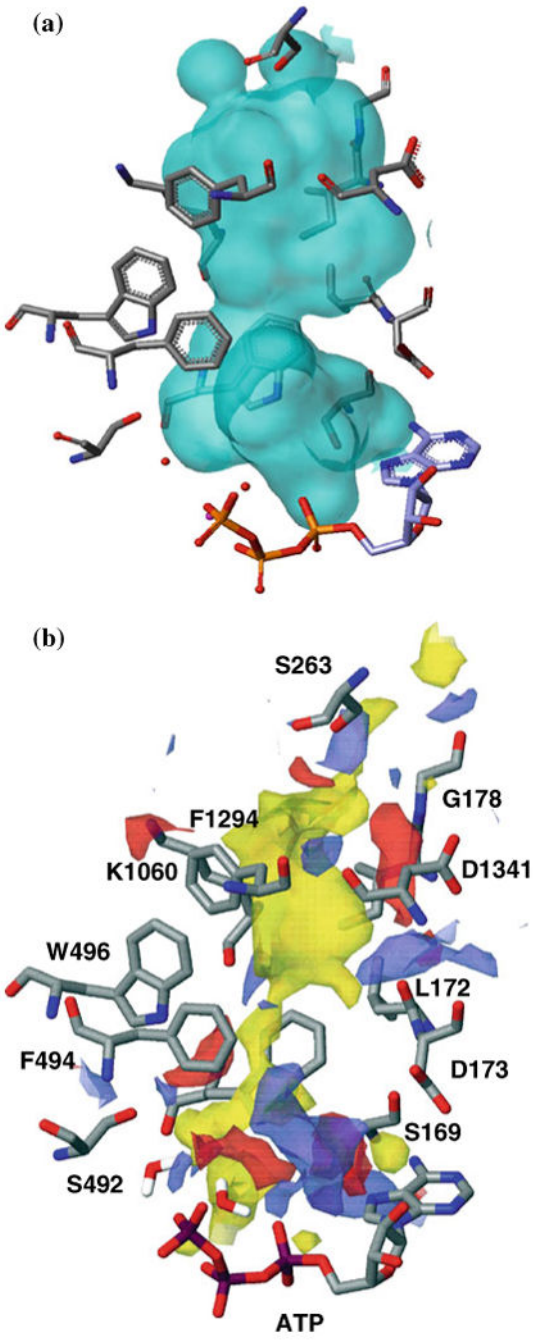

Fig. 8.

Multi-domain interface site. **a** MOLCAD surface of the interface site. **b** Binding site residues (stick representation) and binding site interaction regions generated with SiteMap. *Yellow* hydrophobic region, *blue* H-bond donor region; and *red* H-bond acceptor region

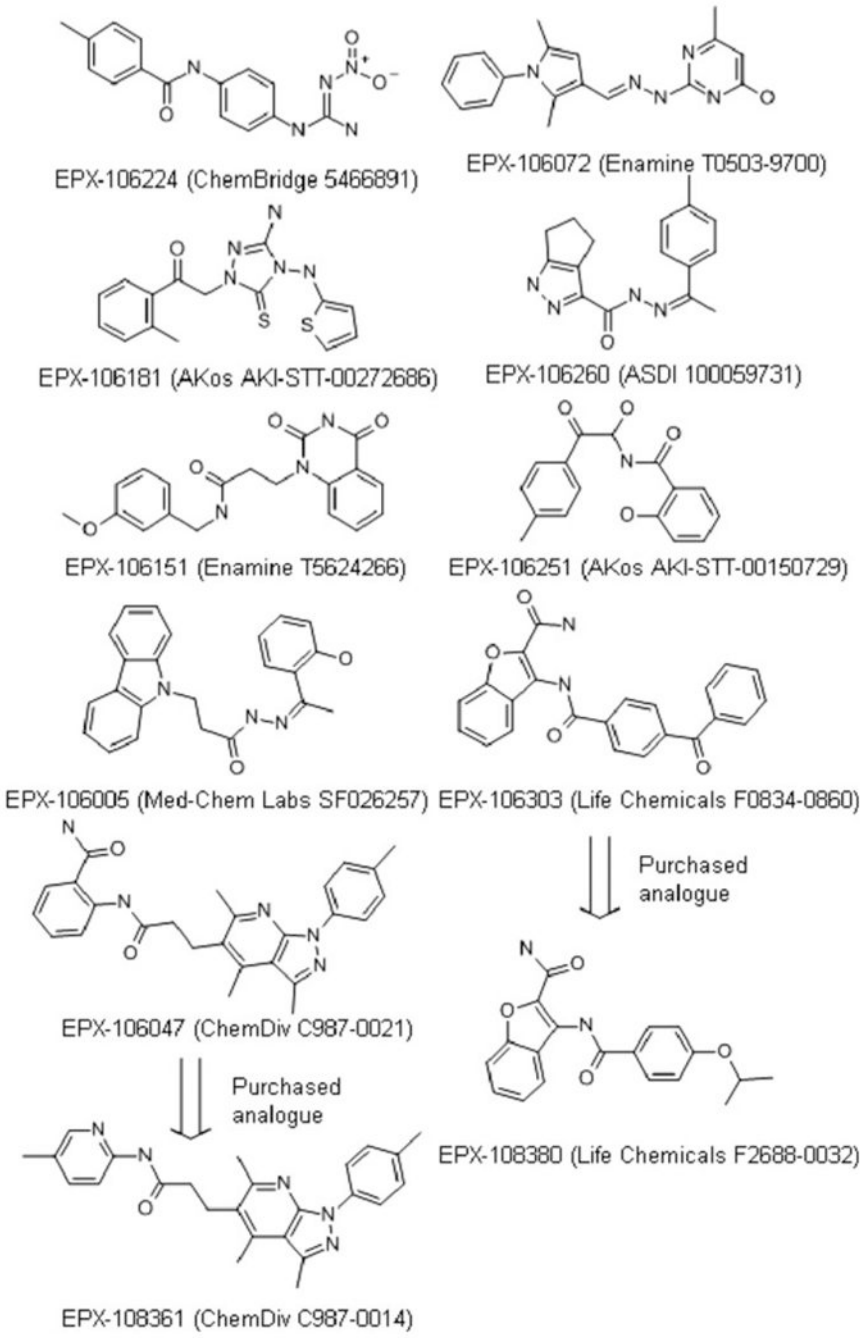

Fig. 9.

Hits from the NBD1:2 interface site and selected analogs. Vendor catalog numbers in parentheses

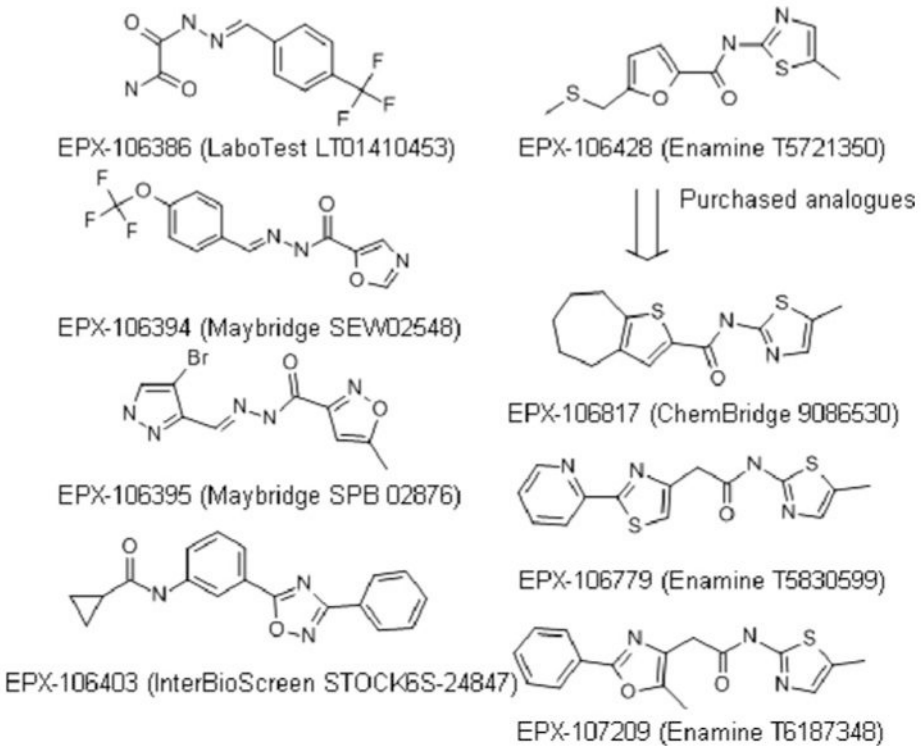

Fig. 10.

Hits from the F508del site and selected analogs. Vendor catalog numbers in parentheses

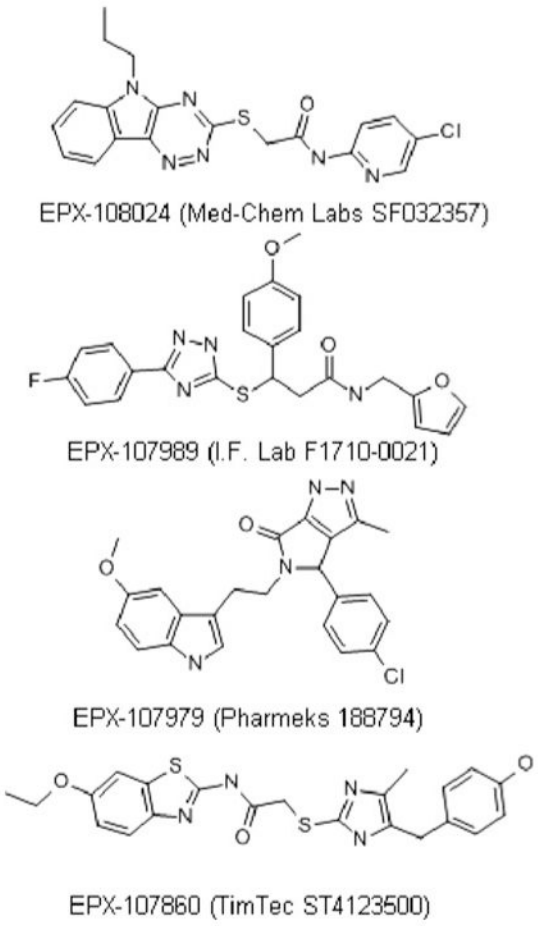

Fig. 11.

Hits from the multi-domain interface site and selected analogs. Vendor catalog numbers in *parentheses*

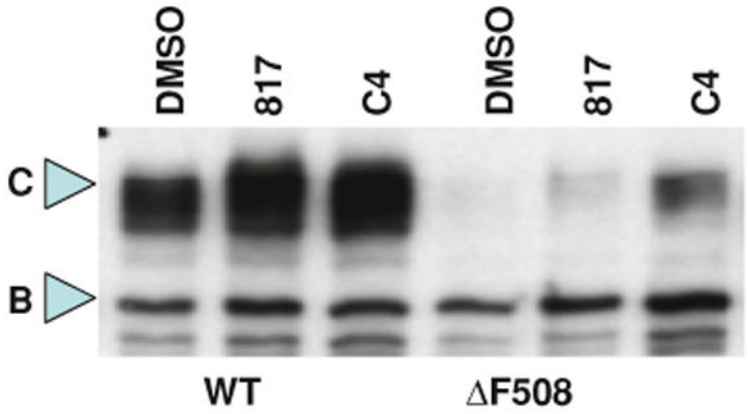

Fig. 12.

Band C Western blot analysis. HeLa cells were transiently transfected with wt CFTR or F508del-CFTR. One day after transfection cells were incubated at 30 °C with vehicle (DMSO), 20 μM EPX-106817 or 10 μM corrector 4a (C4) for 24 h after which cells were collected and whole cell lysates analyzed by Western blot. Compared to DMSO, EPX-106817 and C4 reproducibly increased the amount of Band C observed for both wt and F508del-CFTR transfected cells

J Comput Aided Mol Des. Author manuscript; available in PMC 2014 May 05.

Table 1

Activities of hits from the NBD1:2 interface site and selected analogs

Compound ID	Activity	Forskolin response	Acute addition of 10 μM compound	IBMX response	Genistein response	Total response
EPX-106224	Corrector	**−4.3 ± 0.3 (6)**		−8.4 ± 0.6 (6)	−10.2 ± 0.6 (6)	**−22.9 ± 1.2 (6)**
Vehicle		−2.3 ± 0.3 (4)		−7.3 ± 0.6 (4)	−8.7 ± 0.3 (4)	−18.4 ± 0.9 (4)
EPX-106181	Corrector	**−2.9 ± 0.2 (6)**		−8.5 ± 0.3 (6)	−11.8 ± 0.4 (6)	**−23.2 ± 0.6 (6)**
Vehicle		−2.0 ± 0.2 (4)		−8.1 ± 0.4 (4)	−10.8 ± 0.4 (4)	−21.0 ± 0.6 (4)
EPX-106151	Corrector	−3.0 ± 0.3 (6)		−7.2 ± 0.3 (6)	−10.9 ± 0.3 (6)	**−21.1 ± 0.7 (6)**
Vehicle		2.0 ± 0.2 (4)		−5.9 ± 0.5 (4)	−9.8 ± 0.4 (4)	−17.7 ± 1.1 (4)
EPX-106251	Corrector	−2.1 ± 0.1 (6)		**−7.6 ± 0.3 (6)**	−9.7 ± 0.3 (6)	**−19.4 ± 0.5 (6)**
Vehicle		1.6 ± 0.2 (4)		−6.7 ± 0.1 (4)	−9.3 ± 0.4 (4)	−17.6 ± 0.4 (4)
EPX-106072	Corrector	**−3.8 ± 0.3 (6)**		**−9.7 ± 0.3 (6)**	**−11.8 ± 0.3 (6)**	**−25.4 ± 0.6 (6)**
Vehicle		−2.0 ± 0.1 (4)		−6.6 ± 0.2 (4)	−9.4 ± 0.3 (4)	−18.0 ± 0.4 (4)
EPX-106047	Corrector	**−1.4 ± 0.2 (6)**		**−7.9 ± 0.5 (6)**	−11.2 ± 0.5 (6)	**−20.6 ± 1.0 (6)**
Vehicle		−0.7 ± 0.1 (4)		−4.2 ± 0.4 (4)	−10.2 ± 0.4 (4)	−15.1 ± 0.8 (4)
EPX-108361	Corrector	**−7.1 ± 0.3 (6)**	**−1.0 ± 0.1 (6)**	**−6.3 ± 0.3 (6)**	**−25.4 ± 1.0 (6)**	**−39.8 ± 1.3 (6)**
Vehicle		−4.1 ± 0.4 (3)	−0.2 ± 0.1 (3)	−4.2 ± 0.1 (3)	−19.8 ± 1.7 (3)	−28.4 ± 1.8 (3)
EPX-106005	Corrector	**−3.9 ± 0.5 (5)**		**−11.0 ± 0.8 (5)**	−9.6 ± 0.4 (5)	−24.5 ± 1.5 (5)
Vehicle		−2.2 ± 0.2 (4)		−8.6 ± 0.5 (4)	−9.4 ± 0.6 (4)	−20.2 ± 1.1 (4)
EPX-106260	Corrector	**−2.8 ± 0.2 (6)**		**−10.5 ± 0.3 (6)**	**−10.5 ± 0.3 (6)**	**−24.6 ± 0.4 (6)**
Vehicle		−1.6 ± 0.1 (4)		−8.0 ± 0.0 (4)	−8.0 ± 0.0 (4)	−17.9 ± 0.4 (4)
EPX-106303	Potentiator	**−4.0 ± 0.6 (6)**		**−9.6 ± 0.3 (6)**	*−8.2 ± 0.4 (6)*	−21.8 ± 0.6 (6)
Vehicle		−2.1 ± 0.2 (4)		−7.7 ± 0.2 (4)	−10.9 ± 0.4 (4)	−20.8 ± 0.8 (4)
EPX-108380	Dual acting	**−33.2 ± 1.1 (6)**	**−12.3 ± 0.8 (6)**	*−2.9 ± 0.2 (6)*	*−1.1 ± 0.0 (6)*	**−49.5 ± 1.6 (6)**
Vehicle		−5.1 ± 0.4 (4)	-0.2 ± 0.2 (4)	−7.3 ± 1.1 (4)	−20.3 ± 0.4 (4)	−32.9 ± 1.3 (4)

Numbers represent current densities in μA/cm^2 recorded in Ussing chamber assays from FRT cells after 24 h incubation with 10 μM compound. Bold numbers represent significantly ($p < 0.05$) increased response compared to response from vehicle-incubated cells. Italicized numbers represent significantly ($p < 0.05$) decreased response compared to vehicle-treated cells. Statistical significance was computed with Dunnett's test. Correctors must achieve greater total response than vehicle, whereas potentiators only need to significantly increase the forskolin response. In addition to incubation, the analogs EPX-108380 and EPX-108361, tested at a later stage, were acutely added following forskolin addition as an initial measure for potentiation which was later validated by acute addition without prior incubation. Vehicle control data is given for each compound because testing was performed on different days. Note that under comparable assay conditions low temperature correction yields current densities after forskolin, IBMX, and genistein of −78.5 ± 17.0 mA/cm^2 with a range of −51.1 to −111.3 mA/cm^2 (n = 23)

Table 2

Activities of hits from F508del site and selected analogs

Compound ID	Activity	Forskolin response	IBMX response	Genistein response	Total response
EPX-106386	Corrector	−1.1 ± 0.1 (6)	**−4.8 ± 0.2 (6)**	**−10.9 ± 0.2 (6)**	**−16.8 ± 0.4 (6)**
Vehicle		−0.9 ± 0.2 (4)	−3.7 ± 0.3 (4)	−9.5 ± 0.4 (4)	−14.1 ± 0.9 (4)
EPX-106394	Corrector	**−5.1 ± 0.6 (6)**	**−9.8 ± 0.3 (6)**	−11.4 ± 0.3 (6)	**−26.2 ± 0.9 (6)**
EPX-106395	Corrector	−4.5 ± 0.3 (6)	**−10.4 ± 0.5 (6)**	−11.9 ± 0.3 (6)	**−26.9 ± 0.7 (6)**
Vehicle		−2.2 ± 0.8 (4)	−6.9 ± 0.8 (4)	−11.4 ± 0.4 (4)	−20.5 ± 1.5 (4)
EPX-106403	Corrector	**−8.3 ± 0.4 (6)**	**−10.7 ± 0.5 (6)**	**−13.2 ± 0.7 (6)**	**−32.2 ± 1.4 (6)**
Vehicle		−6.4 ± 0.3 (4)	−6.1 ± 0.1 (4)	−10.5 ± 0.4 (4)	−23.1 ± 0.7 (4)
EPX-106428	Corrector	−1.5 ± 0.1 (6)	**−5.3 ± 0.2 (6)**	**−12.0 ± 0.5 (6)**	**−18.8 ± 0.7 (6)**
Vehicle		−1.3 ± 0.1 (4)	−4.0 ± 0.2 (4)	−9.6 ± 0.2 (4)	−14.9 ± 0.5 (4)
EPX-106817	Corrector	**−24.4 ± 2.8 (6)**	**−12.9 ± 0.8 (6)**	**−18.4 ± 1.2 (6)**	**−55.7 ± 4.0 (6)**
Vehicle		−5.7 ± 0.6 (4)	−5.7 ± 0.6 (4)	−8.2 ± 0.6 (4)	−19.7 ± 1.8 (4)
EPX-106779	Corrector	**−15.4 ± 0.9 (6)**	**−24.6 ± 1.9 (6)**	**−28.2 ± 1.6 (6)**	**−68.2 ± 2.9 (6)**
Vehicle		−5.1 ± 0.4 (4)	−8.0 ± 0.7 (4)	−10.2 ± 0.1 (4)	−23.3 ± 0.6 (4)
EPX-107209	Corrector	**−8.0 ± 0.5 (6)**	**−20.1 ± 1.2 (6)**	**−11.8 ± 0.5 (6)**	**−40.0 ± 2.2 (6)**
Vehicle		−3.4 ± 0.0 (4)	−4.4 ± 0.2 (4)	−6.0 ± 0.2 (4)	−13.8 ± 0.4 (4)

Numbers represent current densities in μA/cm^2 recorded in Ussing chamber assays from FRT cells after 24 h incubation with 10 μM compound. Bold numbers represent significantly ($p < 0.05$) increased response compared to response from vehicle-incubated cells. Statistical significance was computed with Dunnett's test. Correctors must achieve greater total response than vehicle, whereas potentiators only need to significantly increase the forskolin response. Vehicle control data is given for each compound because testing was performed on different days (Compounds 394 and 395 were tested on the same day and share the same control). Note that under comparable assay conditions low temperature correction yields current densities after forskolin, IBMX, and genistein of −78.5 ± 17.0 mA/cm^2 with a range of −51.1 to -111.3 mA/cm^2 (n = 23)

Table 3

Activities of hits from the multi-domain interface site

Compound ID	Activity	Forskolin response	Acute addition of 10 μM compound	IBMX response	Genistein response	Total response
EPX-108024	Potentiator	−3.5 ± 0.5 (6)	**−1.7 ± 0.3 (6)**	−5.5 ± 0.4 (6)	−16.8 ± 0.9 (6)	−27.5 ± 1.4 (6)
EPX-107989	Potentiator	−5.4 ± 0.4 (6)	**−1.4 ± 0.2 (6)**	−5.2 ± 0.4 (6)	−16.5 ± 1.6 (6)	−28.6 ± 2.3 (6)
6 μM corrector 4a		**−8.5 ± 0.6 (2)**	0.0 ± 0.0 (2)	**−12.6 ± 1.8 (2)**	**−36.1 ± 6.2 (2)**	**−57.3 ± 7.4 (2)**
Vehicle		−3.3 ± 0.2 (4)	−0.2 ± 0.1 (4)	−5.6 ± 0.4 (4)	−17.9 ± 0.9 (4)	−27.0 ± 1.1(4)
EPX-107860	Corrector	**−14.0 ± 0.9 (6)**	−0.4 ± 0.1 (6)	**−11.0 ± 1.0 (6)**	**−26.8 ± 1.6 (6)**	**−52.2 ± 3.0 (6)**
6 μM corrector 4a		**−10.0 ± 1.4 (2)**	−0.5 ± 0.3 (2)	**−17.9 ± 0.7 (2)**	**−28.5 ± 3.8 (2)**	**−56.9 ± 1.3 (2)**
Vehicle		−5.0 ± 1.2 (4)	−0.2 ± 0.1 (4)	−8.9 ± 0.8 (4)	−14.2 ± 2.7 (4)	−28.4 ± 1.4 (4)
EPX-107979	Dual acting	**−23.0 ± 0.7 (6)**	**−1.0 ± 0.2 (6)**	**−5.5 ± 0.2 (6)**	*−4.2 ± 0.1 (6)*	**−33.8 ± 0.8 (6)**
6 μM corrector 4a		**−12.6 ± 0.6 (2)**	−0.2 ± 0.2 (2)	**−17.1 ± 0.6 (2)**	**−37.7 ± 0.1(2)**	**−67.6 ± 0.0(2)**
Vehicle		−3.5 ± 0.4(4)	−0.3 ± 0.1(4)	−4.8 ± 0.1(4)	−13.4 ± 0.9(4)	−22.0 ± 1.3(4)

Numbers represent current densities in μA/cm^2 recorded in Ussing chamber assays from FRT cells after 24 h incubation with 10 μM compound. Bold numbers represent significantly ($p < 0.05$) increased response compared to response from vehicle-incubated cells. Italicized numbers represent significantly ($p < 0.05$) decreased response compared to vehicle-treated cells. Statistical significance was computed with Dunnett's test. Correctors must achieve greater total response than vehicle, whereas potentiators only need to significantly increase the forskolin response. In addition to incubation, compounds were acutely added following forskolin addition as an initial measure for potentiation which was later validated by acute addition without prior incubation. EPX-107979 achieved a significantly increased total response paired with a significant response to acute addition, thus it is classified as a dual-acting corrector-potentiator. Relevant vehicle and corrector 4a control data is provided below each compound. Responses from vehicle-treated cells varied because testing was performed on different days. Note that under comparable assay conditions low temperature correction yields current densities after forskolin, IBMX, and genistein of −78.5 ± 17.0 mA/cm^2 with a range of −51.1 to −111.3 mA/cm^2 (n = 23)

Table 4

Screening hit statistics for structure-based corrector screens

Binding site screened	Number of compounds tested	Number of validated corrector hits	Number of validated potentiator hits	Corrector hit rate	Modulator hit rate
NBD1:NBD2 interface	205	8	5	3.9%	6.3%
F508del site	101	5	3	5.0%	7.9%
Multi-domain interface	190	2	2	1.1%	2.1%
Total	496	15	10	3.0%	5.0%