Article

CFTR Lifecycle Map—A Systems Medicine Model of CFTR Maturation to Predict Possible Active Compound Combinations

Liza Vinhoven [1], Frauke Stanke [2,3], Sylvia Hafkemeyer [4] and Manuel Manfred Nietert [1,5,*]

1   Department of Medical Bioinformatics, University Medical Center Göttingen, Goldschmidtstraße 1,
    37077 Göttingen, Germany; liza.vinhoven@med.uni-goettingen.de
2   Clinic for Pediatric Pneumology, Allergology and Neonatology, Hannover Medical School,
    Carl-Neuberg-Strasse 1, 30625 Hannover, Germany; mekus.frauke@mh-hannover.de
3   Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), The German Center
    for Lung Research, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany
4   Mukoviszidose Institut gGmbH, In den Dauen 6, 53117 Bonn, Germany; shafkemeyer@muko.info
5   CIDAS Campus Institute Data Science, Goldschmidtstraße 1, 37077 Göttingen, Germany
*   Correspondence: manuel.nietert@med.uni-goettingen.de; Tel.: +49-551-39-14920

Abstract: Different causative therapeutics for CF patients have been developed. There are still no mutation-specific therapeutics for some patients, especially those with rare CFTR mutations. For this purpose, high-throughput screens have been performed which result in various candidate compounds, with mostly unclear modes of action. In order to elucidate the mechanism of action for promising candidate substances and to be able to predict possible synergistic effects of substance combinations, we used a systems biology approach to create a model of the CFTR maturation pathway in cells in a standardized, human- and machine-readable format. It is composed of a core map, manually curated from small-scale experiments in human cells, and a coarse map including interactors identified in large-scale efforts. The manually curated core map includes 170 different molecular entities and 156 reactions from 221 publications. The coarse map encompasses 1384 unique proteins from four publications. The overlap between the two data sources amounts to 46 proteins. The CFTR Lifecycle Map can be used to support the identification of potential targets inside the cell and elucidate the mode of action for candidate substances. It thereby provides a backbone to structure available data as well as a tool to develop hypotheses regarding novel therapeutics.

Keywords: cystic fibrosis; CFTR; CFTR maturation; systems medicine model; trafficking; CFTR modulators

1. Introduction

Cystic fibrosis (CF) is an inherited disorder prevalent among the white European population, where, with an incidence of approximately 1 in 3000 newborns, it is one of the most common monogenic autosomal recessive diseases [1]. CF is caused by mutations of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene [2], which encodes a membrane protein that serves as a chloride and bicarbonate channel in exocrine epithelia of various organs, and thereby regulates the viscosity of the mucus lining [3]. Defective CFTR, therefore, has severe implications throughout the body, its major hallmarks being recurrent pulmonary infections and pancreatic insufficiency [2,3].

CFTR is an approximately 170 kDa membrane-spanning glycoprotein composed of 1440 amino acid residues and complex glycosylation [4–6]. It belongs to the ATP binding cassette (ABC) Transporter Superfamily [7,8], but, unlike most of them, acts as an ion channel as opposed to an active transporter. As an ion channel, CFTR only requires ATP for opening, whereas other ABC transporters also require ATP for the active transport of substrates across membranes [9]. In accordance with the other ABC transporters, CFTR consists of two transmembrane domains (TM1 and TM2) and two cytosolic nucleotide-binding

domains (NBD1 and NBD2) with an additional highly flexible regulatory region (R-region) at its center [4,9–11]. Several in-depth reviews cover the structure and opening mechanism of CFTR in great detail [9,12–14]. CFTR undergoes an intricate maturation pathway with complex folding and core glycosylation at the endoplasmic reticulum (ER), before being trafficked through the secretory pathway and further glycosylated at the Golgi apparatus. At the ER, CFTR is subject to extensive quality control mechanisms, often resulting in premature degradation through the ER-associated degradation pathway (ERAD), which leads to only 20–40% of nascent peptides of wild-type (wt) CFTR being correctly folded and trafficked to the apical plasma membrane (PM). After being integrated into the membrane, CFTR undergoes continuous endocytosis, recycling and, when misfolded, degradation in the lysosome [15].

To date, more than 2100 mutations of the *CFTR* gene are known, several hundred of which are known to be disease-causing [16–18]. They can cause defects anywhere in CFTR's complicated and sensitive lifecycle, which is why they were traditionally subdivided into six different classes by the effect they have on the CFTR protein: (I) no protein synthesis, (II) CFTR trafficking defect, (III) dysregulation of CFTR, (IV) defective chloride conductance or channel gating, (V) reduced CFTR transcription and synthesis and (VI) less stable CFTR [19–21]. The traditional classification system has been proposed by Welsh and Smith in 1993 [19] and since has been reviewed and adapted in a range of publications [20,22,23]. However, since many mutations exhibit more complex phenotypes, nowadays, a modified and expanded classification system finds use, where all combinations of the six original mutation classes are regarded as possible classes [21]. For example, the most prevalent CF-causing mutation is the deletion of phenylalanine at position 508 (F508del), which accounts for almost 70% of CF chromosomes worldwide [1,24]. F508del causes a folding defect which results in premature degradation at the ER, making it a class II mutation [6,25]. However, even when rescued and trafficked to the membrane, its channel gating is reduced and it is less stable than wt-CFTR, meaning it exhibits mutation class III/IV and VI characteristics as well [26,27]. According to the expanded classification system, it is therefore classified as a class II-III-VI mutation [21]. Several other mutations also display different defects. Therefore, the wide range of *CFTR* gene mutations, resulting in different defects in the CFTR protein, makes causative treatments for CF difficult to find, and the recently available CF modulators do not target all mutations. As a result, there is still no mutation-specific therapeutic for some patients, especially those with rare *CFTR* mutations. The latest research efforts, therefore, focused on developing combination therapies to target multiple defects at once [28,29]. For this purpose, high-throughput (HT) screens have been performed [30–42], where thousands of substances have been tested in different cell models [43]. These result in a plethora of data and various candidate compounds, often with an unclear mode of action. In order to provide an overview of already tested compounds, we previously established the publicly available database CandActCFTR (https://candactcftr.ams.med.uni-goettingen.de/ (accessed on 6 June 2021)), where substances from 90 publications are listed and categorized according to their influence on CFTR function (manuscript submitted for publication).

In order to support the elucidation of the mechanism of action for promising candidate substances and to be able to predict possible synergistic effects of substance combinations, we used a systems biology approach to create a model of the CFTR maturation pathway in cells. Systems biology modeling aims to gather knowledge on biological systems and translate it into a human- and machine-understandable format in order to analyze its behavior and interactions. To make models reproducible and reusable, there are certain standards and formats to adhere to. The most well-established, standardized format in the graphical representation of biological processes is the Systems Biology Graphical Notation (SBGN), consisting of three languages [44]. Of the three, the SBGN Process Description (PD) language allows the most detailed representation of molecular mechanisms using nodes and directed edges. Molecular entities are shown as nodes and have different shapes depending on their molecular species. For example, proteins are represented as rounded

rectangles, RNAs as parallelograms and small molecules as ovals. The reactions (e.g., transcription, translation and state transition) and reaction regulations (e.g., catalysis and inhibition) between the molecular entities are represented by edges, shaped as arrows with differently shaped heads depending on the type of interaction [44]. A glossary of systems biology and bioinformatics terms used here can be found in the Supplementary Materials.

Here, we present a systems biological model of the CFTR lifecycle in a standardized, explorable and tractable format. The model is composed of two datasets, a core map manually curated from small-scale experiments in human cells, and a coarse map including interactors identified in high-throughput (HT) efforts [45–48]. Interactors are here defined as molecular entities that influence CFTR directly or indirectly. Both data layers are divided into submaps focusing on different stages of the CFTR lifecycle and different processes the ion channel is involved in. Overall, the manually curated core map includes 170 different molecular entities and 156 reactions from 221 experimental publications. The high-throughput data layer encompasses 1384 unique interactors from four publications by Wang et al., 2006, Pankow et al., 2015, Santos et al., 2019 and Matos et al., 2019 [45–48]. The CFTR Lifecycle map provides a tool to structure and exploit existing knowledge and data, as well as develop a hypothesis regarding synergistic drug targets and novel therapeutics.

2. Results

2.1. CFTR Map

The CFTR map encompasses information from small-scale experiments as well as high-throughput efforts, leading to differences in the degree of detail and confidence. It was therefore split into different data layers. The first data layer, the core map, was manually curated and only includes high-confidence interactors, confirmed by at least two independent small-scale experiments or acknowledged by two reviews from different research groups. As a result, the number of molecular interactions in the core map is limited but each is described with a high level of detail. The second data layer in the coarse map represents the high-throughput interactome of wt-CFTR (the interactome of F508del was excluded) as published by Wang et al., 2006, Pankow et al., 2015, Santos et al., 2019 and Matos et al., 2019 [45–48] in a structured cell layout. Here, the large-scale experimental method does not allow for conclusions regarding the nature of the interactions. Therefore, a small level of detail, but a high number of interactors, is included in the map. Overall, the manually curated core map comprises 170 different interactors and the coarse map from large-scale efforts contains 1384 interactors; 46 interactors could be found in both (Figure 1). A list of all interactors in both maps, as well as their overlap, can be found in Supplementary Table S1. To prevent redundancies, the overlapping interactors were only kept in the manually curated core map and excluded from the coarse map.
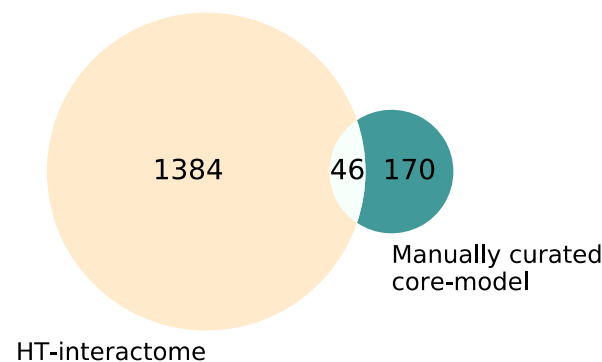

Figure 1. Venn diagram of the interactors in the manually curated core map vs. coarse map. The manually curated list of interactors within the core map comprises 170 interactors, the coarse map derived from high-throughput data contains 1384 interactors, and 46 interactors occur in both lists. The overlap was subtracted from the high-throughput interactome for the visualization to avoid redundancies, resulting in 1338 interactors in the coarse map.

2.2. Representation of the CFTR Lifecycle in the CFTR Core Map

The CFTR core map represents the molecular mechanisms affecting wt-CFTR during its lifecycle. It is the product of an exhaustive literature curation process and the manual integration of different data sources. As the whole model is represented and written in the standardized SBGN format, it is human understandable as well as computationally tractable. It was created in the editor CellDesigner4.4.2 [49,50], adhering to the Process Description language of the SBGN format [44]. At the moment, it encompasses 262 different molecular entities and 156 reactions in 6 main cellular compartments. The biomolecules are categorized into 149 proteins, 58 complexes, 28 simple molecules, 13 ions, 6 genes, 5 RNAs, and 3 pools of degraded protein, amino acids, or nucleotides. Proteins can be subdivided into generic proteins, truncated proteins, ion channels, and receptor proteins. The color of the molecular entity indicates whether it was identified in at least one polarized cell line (green) or non-polarized cell lines only (yellow). Reactions are specified as state transitions, inhibitions, catalysis, transports, and heterodimer associations and dissociations. Each interaction is supported by at least two independent publications, leading to an overall number of 221 publications. A complete list of all interactors and references can be found in Supplementary Tables S2–S6.

The CFTR core map (Figure 2) has a roughly cell-shaped layout, with CFTR making its way from the nucleus at the bottom all the way up to its site of action at the plasma membrane. It covers the molecular interactions CFTR undergoes on its way from being transcribed in the nucleus to being a functional ion channel at the apical plasma membrane, including its activity and regulation there, as well as endocytosis, recycling and degradation of the mature protein. The map can be subdivided into five submaps (Table 1) to enable the user to either look at its whole or individual processes, depending on their focus or interest. The five submaps are guided by the subcellular location and process they focus on.

1. Transcription—Nucleus: The Nucleus submap covers the transcriptional regulation of the *CFTR* gene into its mRNA.
2. Translation, Folding and ER Quality Control—ER: The ER submap step-by-step describes the translation of the mRNA into the CFTR peptide and its integration into the membrane as well as folding steps modulated by chaperones, core glycosylation, and the calnexin cycle involved in ER quality control. Depending on the folding success, CFTR may progress through the secretory pathway or be degraded through ER-associated degradation.
3. Secretory Pathway—ER, Golgi Apparatus, Plasma Membrane: The Secretory Pathway submap covers COPII vesicle-mediated trafficking between the ER, Golgi, and the plasma membrane and the full glycosylation at the Golgi apparatus. It also describes unconventional trafficking of the protein between the ER and plasma membrane, which has been found to be an alternative route CFTR may take.
4. Activity and Regulation—Plasma Membrane: The Activity submap covers the phosphorylation-dependent activation of CFTR through the cAMP signaling cascade, channel opening, closing, and ion conductance as well as regulatory interactions with other ion channels and stabilization through interactions with the cytoskeleton.
5. Endocytosis, Recycling and Degradation—Plasma Membrane, Endosomes, Lysosomes (Figure 2c): The final submap describes the endocytosis of the mature CFTR protein from the plasma membrane, which can be recycled back to the membrane or degraded in lysosomes.
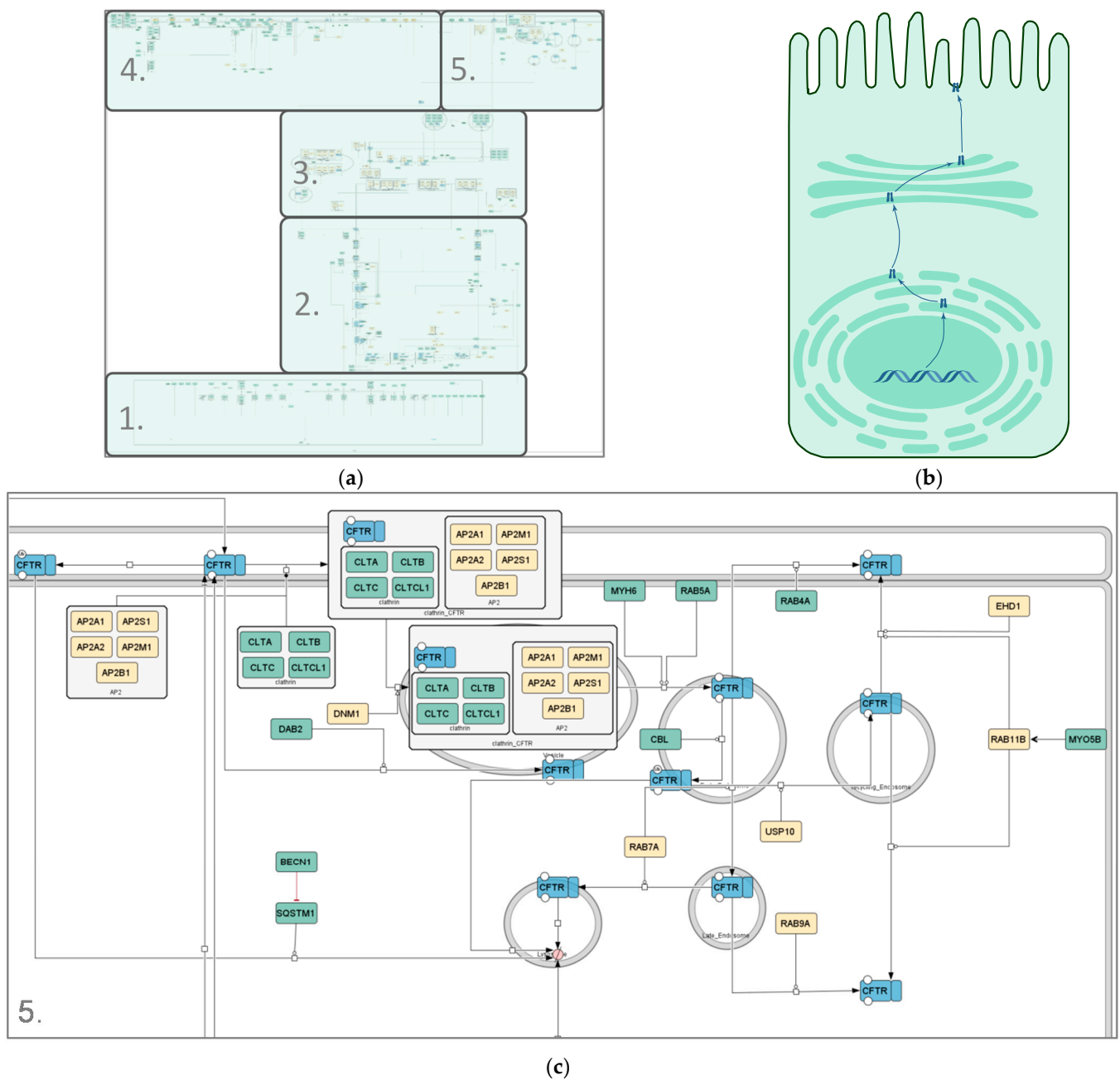
Figure 2. Different representations of the CFTR lifecycle. (**a**) Image of the SBGN-compliant manually curated CFTR lifecycle core map; (**b**) cartoon representation of the CFTR lifecycle in an apical epithelial cell; (**c**) zoomed-in section of the endocytosis pathway in the manually curated CFTR lifecycle core map (submap 5). Compartments are depicted in grey, CFTR in blue and interactors in differing shades of green and yellow. The green color scheme represents interactors identified in at least one polarized cell line; yellow interactors were identified in non-polarized cell lines only. State transitions, catalysis and positive influences are shown in black; negative influences and inhibitions are displayed in red. Different shapes represent different kinds of interactors. Rounded rectangles correspond to proteins, ovals and circles to small molecules and ions, respectively, rectangles correspond to genes, rhomboids to RNA molecules and chevron shapes to receptors. The map was created using CellDesigner4.4.2.

Table 1. Description of the five submaps of the CFTR core map.

| Process	Localization	Molecular Entities Present in the Model	N [1]	Proportion of the Interactors Identified in Polarized Cells
Transcription	Nucleus	Proteins RNAs and gene elements Small molecules and ions	28 16 1	97%
Translation, Folding and ER quality control	ER	Proteins Small molecules and ions	45 13	69%
Secretory pathway	ER, Golgi apparatus, Plasma Membrane	Proteins Small molecules and ions	27 8	52%
Activity and Regulation	Plasma Membrane	Proteins Small molecules and ions	44 20	82%
Endocytosis, Recycling and Degradation	Plasma Membrane, Endosomes, Lysosomes	Proteins	23	74%

[1] Number of different molecular entities present in the respective submap.

As different cell lines can have different effects on the interactome of proteins [43,51], the cell lines used in the small-scale experiments were identified from each reference and categorized into polarized and non-polarized cells. As the question whether cells are to be characterized as polarized or non-polarized can often be a matter of debate when grown under experimental conditions, polarized cells are here defined as cells with the general ability to polarize. For each interactor, it is indicated by color whether or not they were identified in polarized (green color scheme) or non-polarized cell lines (yellow color scheme), whereby the specific cell lines for each interactor can be found in Supplementary Tables S2–S6. The percentage of interactors identified in at least one polarized cell line was calculated for each submap, amounting to 97% in the Transcription submap, 69% and 52% in the ER and Golgi submap, respectively, 82% in Activity and Regulation and 74% in Endocytosis, Recycling and Degradation.

2.3. Protein–Protein Interaction Network and Topological Analysis of the CFTR Core Map

To analyze the manually curated core map with regard to interactions between the proteins included, a protein–protein interaction network was created using the list of genes present in the model (Figure 3a). All proteins (nodes) of the protein–protein interaction network were identified as CFTR interactors through the manual literature curation, whereas all interactions (edges) between them were identified through the BioGrid database. A list of all interactions present in the protein–protein interaction network can be found in Supplementary Table S7. There are 145 nodes and 326 edges present in the network, and the average number of neighbors amounts to approximately 4.5. Another important property to assess when analyzing protein–protein interaction networks is its degree distribution. Here, the degree of a protein (node) is the number of interactions (edges) it shares with another protein. Most biological networks are considered 'scale-free', meaning that the majority of nodes have a low degree with only a few highly interconnected hubs, represented by a large diameter in Figure 3a. In order to analyze whether the CFTR protein–protein interaction network is scale-free, the degree distribution was calculated (Figure 3b). As can be seen, it follows a power law, confirming that it is, indeed, scale-free. Their scale-free character lends biological networks features important in biological systems. On the one hand, they are quite stable, as a random failure is likely to affect a protein with a low degree due to their high prevalence. Therefore, random failures and changes are unlikely to have a large effect on the overall network. On the other hand, targeted interventions at one of the few hubs will have a large effect on the whole network, which can be important when treating diseases and considering side effects. Here, CFTR is, as expected, the node with the highest degree, which can also be seen in the visualization in Figure 3a, where CFTR is the

largest node. Apart from CFTR, the five nodes with the next highest degree are HSP90AA1 (heat shock protein 90), STUB1 (ubiquitin–protein ligase CHIP), UBE2I (SUMO-conjugating enzyme), NR3C1 (Glucorticoid receptor), and VCP (Transitional ER ATPase).
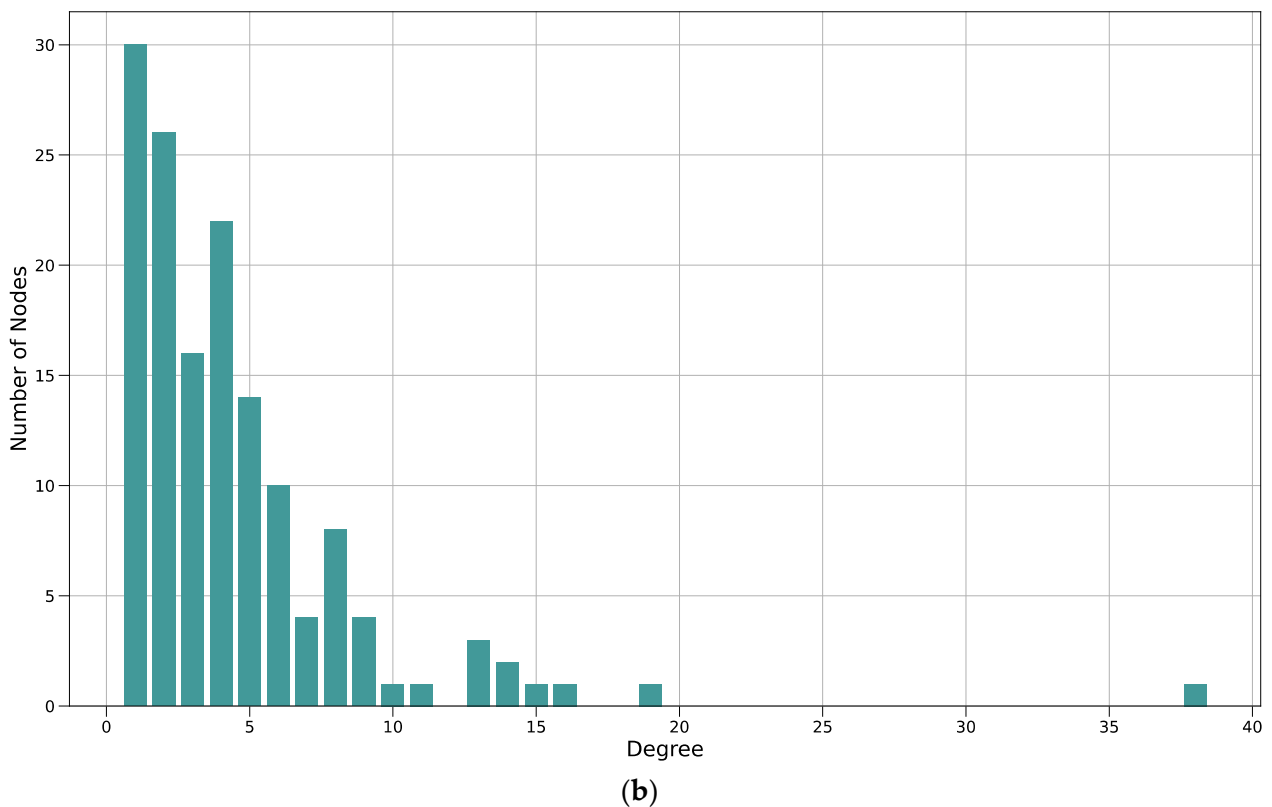

(**a**)

Figure 3. *Cont.*

(b)
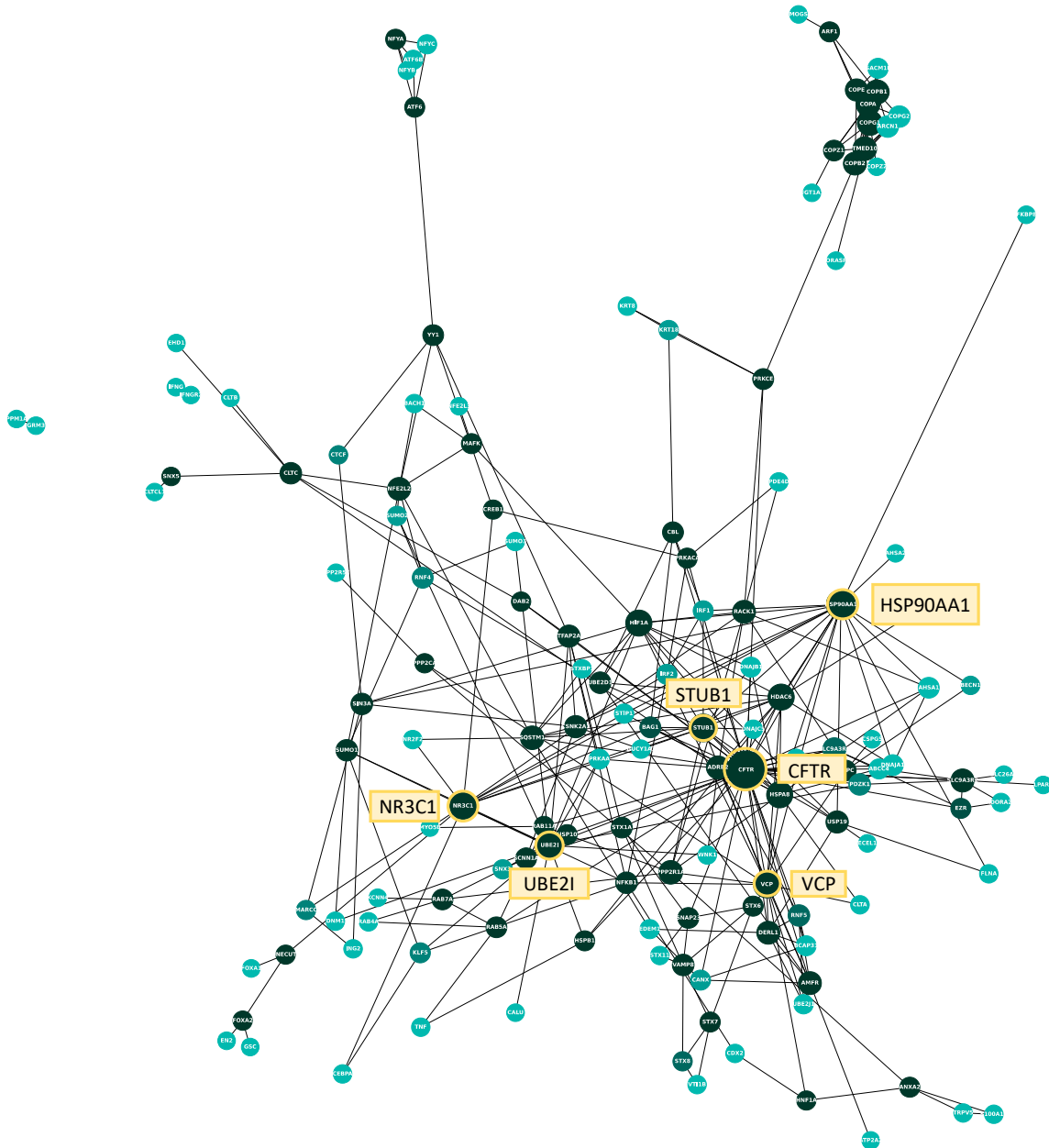
Figure 3. Protein–protein interaction network and degree distribution of the manual CFTR map. (**a**) Each node represents one protein, each edge between them a physical interaction shown in a small-scale study and reported on BioGrid. The larger a node, the higher its degree (i.e., the more interactions it shares with other proteins). CFTR and the five proteins with the next-highest degree are marked in yellow. The color of the protein represents its betweenness centrality, which is a measure of how important the node is to the flow of information through the CFTR Lifecycle Map. The betweenness centrality of a protein is the number of times it lies on the shortest path between two other proteins. The darker the node, the higher its betweenness centrality; (**b**) bar plot of the degree distribution of the protein–protein interaction network in A. The x-axis shows the degree of a protein; the degree is the number of other proteins a protein interacts with. The y-axis shows the number of proteins in the network with a certain degree. For example, the number of proteins that interact with only one other protein in the network (i.e., have a degree of 1) is 30, the number of proteins that interact with six other proteins is ten. The node with the highest degree (38) is CFTR.

2.4. Visualization of the wt-CFTR Interactome as Coarse Model

In addition to the extensive manually curated interaction pathways, which include all high-confidence interactors and detailed interactions, a second data layer was included for interactors with lower confidence and detail. The second data layer represents information from large-scale experiments, namely the wt-CFTR core interactomes published by Wang et al., 2006, Pankow et al., 2015, Santos et al., 2019 and Matos et al., 2019 [45–48]. As the data are stored in long gene lists, in contrast to the detailed, text-based description of interactions in the manual model, different tools were used to represent the high-throughput interactome. The coarse map is also written in the SBGN format, but was created using libsbgnpy [52] and CellDesigner [49,50] and is represented in the SBGN Activity Flow notation, which lacks mechanistic information. The use of the python library libsbgnpy allowed for an automated construction of the maps. In order to structure the information into an intuitive, cell-based layout, the interactors were grouped according to their function and subcellular localization. Again, the model was divided into several submaps, based on the functional categorization (Table 2). The functional category of each interactor was solely based on information from the respective publications and not inferred from other sources. Interactors for which no function was indicated were assigned to the

'other/unknown' category. Figure 4 shows an exemplary image of one of the submaps, which is focused on Endocytosis, Recycling and Degradation. Each submap focuses on one main step or area of function in the CFTR-lifecycle, abstracted into a state-transition reaction. Six of them correspond to those from the core map, with an extra map for mRNA processing between the "Transcription" and "Translation, Folding and ER Quality control". Additional maps focus on interactions with the cytoskeleton, as well as immune-related and other interactors.

1. Transcription—Nucleus: The Transcription submap focuses on the *CFTR* gene and the production of pre-mRNA. All interactors are divided into two functional categories, those that affect the gene directly, e.g., "DNA repair" and "replication", and those that affect the transcription, such as "transcription" and chromatin structure". Apart from the CFTR entities, it includes 17 nodes, seven affecting the gene and ten affecting the state transition.

2. RNA processing: The additional RNA processing map describes the conversion of pre-mRNA to mature mRNA. It includes interactors with functional categorizations, such as nuclear export and RNA splicing, but also RNA degradation, and contains 36 nodes apart from CFTR.

3. Translation, Folding and ER Quality Control—ER: The third submap summarizes the processes taking place in the ER in two state transitions. One is the processing from mature mRNA to folded, core-glycosylated CFTR peptide and degradation at any stage during ER quality control, resulting in an overall number of 45 interactors. The interactors are color-coded depending on whether they affect folding (57 interactors, green), degradation (one interactor, red), both (three interactors, red), or the interaction is unspecified (653 interactors, yellow). The 653 unspecified interactors are mainly from the data published by Santos et al. [47], where the authors characterize the interactome of CFTR prior to its exit from the ER.

4. Secretory Pathway: In accordance with the core map, the Secretory Pathway submap shows the trafficking of the CFTR peptide between the ER, Golgi and PM after folding and core glycosylation. For reasons of simplicity and a lack of information, all 22 interactors were depicted as influencing CFTR trafficking between the ER and Golgi, even though they might be affecting different steps.

5. Activity: Here, all reactions involved in the activity and regulation of and by mature CFTR within the plasma membrane PM are summarized as channel opening, influenced by 38 different entities. It also includes 145 unspecified interactors, that were reported to interact with CFTR at the PM [48], but for which the nature of the interaction is unclear.

6. Recycling and Degradation (Figure 4b): This submap is split into the recycling and degradation of mature CFTR. Endocytosis-regulating interactors are included in the recycling category, resulting in 12 interactors affecting recycling and 32 influencing degradation.

7. Cytoskeleton: An additional submap is designated for interactors with an influence on the anchoring of CFTR in the cell, including 62 entities apart from CFTR.

8. Immunity: A separate submap shows nine interactors playing a role in immunity (10 interactors).

9. Other Functions: In order to represent the whole datasets published, another submap includes all interactors that fall into none of the categories above. These include, for example, proteins involved in metabolism and those for which no function regarding CFTR could be specified (250 interactors).

Table 2. Description of the five submaps of the CFTR high-throughput model.

Process	Localization	Functional Category	N [1]
Transcription	Nucleus	DNA Replication	7
		Transcription	10
RNA Processing	Nucleus–Cytoplasm		36
Translation, Folding and ER quality control	ER	Folding	57
		ER-associated degradation	1
		both	3
		unspecified	653
Secretory pathway	ER, Golgi apparatus, Plasma Membrane		22
Activity and Regulation	Plasma Membrane	Activity	38
		unspecified	145
Endocytosis, Recycling and Degradation	Plasma Membrane, Endosomes, Lysosomes	Recycling	12
		Degradation	32
Cytoskeleton	Cytoplasm, Plasma Membrane		62
Immunity			10
Other/Unknown			250

[1] Number of interactors present in the respective submap.

(**a**)   (**b**)
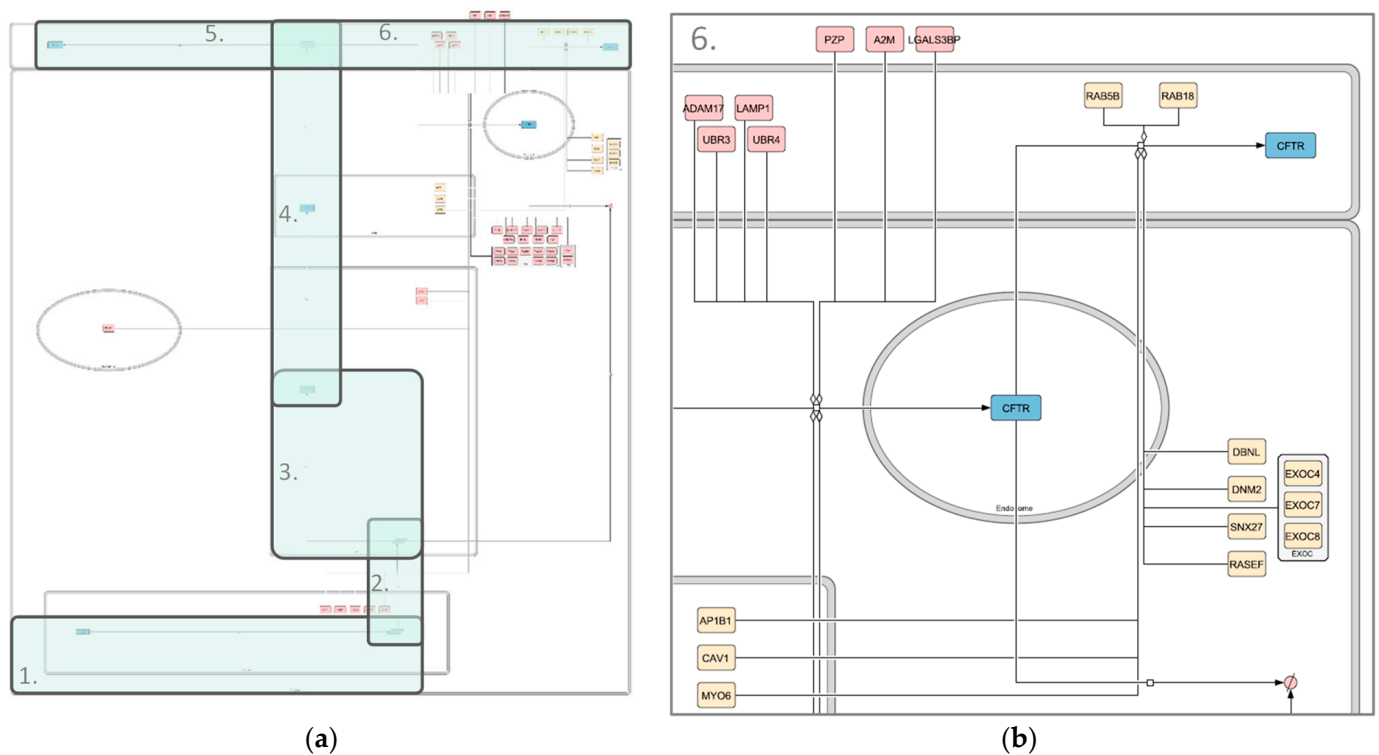
Figure 4. Image of the high-throughput CFTR Endocytosis, Recycling and Degradation map. (**a**) Whole map; (**b**) zoomed-in section of the map. Compartments are depicted in grey, CFTR in blue and interactors in shades of red (degradation associated) and yellow (recycling associated). State transitions and modulations are shown in black. Different shapes represent different kinds of interactors. Rounded rectangles correspond to proteins, rectangles correspond to genes and rhomboids to RNA molecules. The map was created using libsbgnpy 0.2.2 and CellDesigner4.4.2.

2.5. Systemic Interpretation and Comparison of Manually Curated Model and Large-Scale Interactome

The lists of interactors from the core map and coarse map were compared for overlaps. Interestingly, of the 170 manually curated interactors and the 1384 interactors from the high-throughput screens, only 46 interactors could be found in both datasets (Figure 1).We next wanted to see whether the interactors in both datasets belong to similar functional categories, as this would indicate that similar pathways of importance for CFTR folding and maturation have been identified by the targeted analysis of interactors (core map) or by hypothesis-free high-throughput analysis (coarse map).

Datasets of the core and the coarse map were analyzed using the BioInfoMiner web application [53], which performs a biological interpretation based on a list of genes, resulting in prioritized lists of systemic processes and genes, similar to a gene enrichment analysis. Here, the gene ontology (GO) [54,55] terms and Reactome Pathway Database [56] terms were assigned to all interactors. Reactome is a manually curated and peer-reviewed database for cellular pathways on a molecular level, whereas the gene ontology knowledgebase provides a model of biological systems to represent the current knowledge on the function of genes from the molecular to organism level. Therefore, while Reactome and Gene Ontology share a certain overlap, they mainly complement each other, as Reactome associates genes with specific molecular processes, whereas Gene Ontology also takes broader biological processes into account.

Of the top 20 prioritized Gene Ontology terms of the core model-derived gene list and of the coarse model-derived high-throughput gene list, 12 were shared between core and coarse model, including *cellular localization*, *macromolecule localization*, *protein localization* and *vesicle-mediated transport* (Figure 5). Furthermore, the heat shock protein *HSP90AA1*, as well as the ER ATPase *VCP* were found in the top 45 prioritized genes of both gene sets.

In addition to the analysis using Gene Ontology terms, the BioInfoMiner analysis was repeated with terms from the Reactome Pathway Database [56]. Here, however, there was no overlap between Reactome terms assigned to the two gene lists derived from the core model and the coarse model, respectively. While the top-ranked terms for the manually curated gene list derived from Reactome were mostly intracellular transport-related, the top-ranked terms for the high-throughput gene list mostly related to the cellular response to stress and infection. The complete results of the analysis can be found in Supplementary Tables S8–S12.

The diverging results between the Gene Ontology-based analysis and Reactome-based analysis most likely stem from the different levels of biological activity that are defined by these two databases. While the analyzed Gene Ontology terms mainly include the broad biological process terms, such as *intracellular transport*, the Reactome pathway terms are on a smaller, more detailed level, such as *SRP-dependent cotranslational protein targeting to membrane*. Consequently, the same major biological processes appear to be overrepresented in the manually curated core model and the high-throughput interactome represented in the coarse model, while different specific molecular pathways from Reactome are recognized in both layers of the CFTR Lifecycle Map.
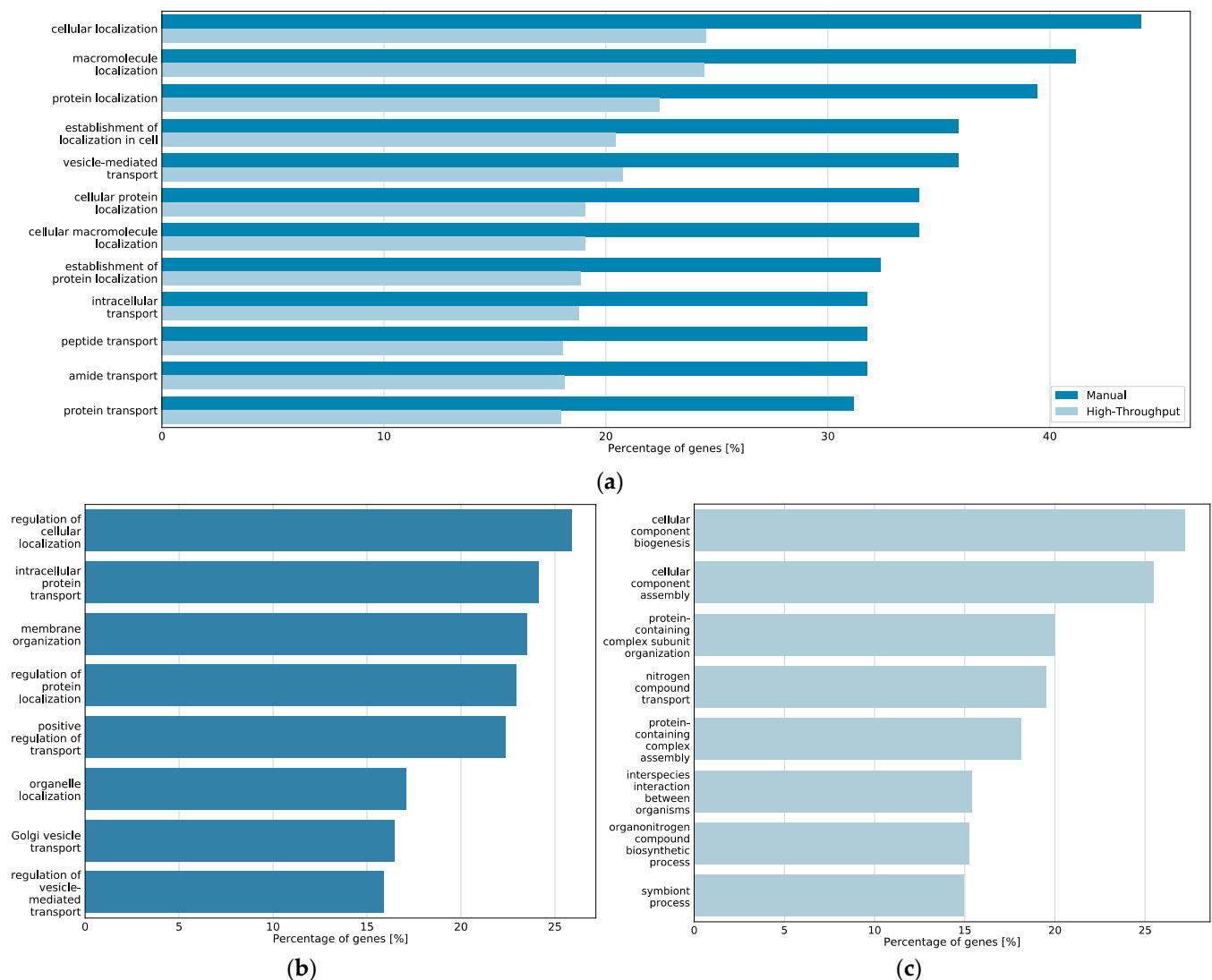
Figure 5. Top 20 prioritized Gene Ontology processes from BioInfoMiner analysis of the CFTR Lifecycle Map. (**a**) Bar plot of the processes prioritized among the top 20 in both the core map and the coarse map datasets and percentage of genes from the respective map associated with the processes; (**b**) bar plot of the processes prioritized among the top 20 in only the manually curated gene list of the core map and percentage of genes from the core map associated with the processes; (**c**) bar plot of the processes prioritized among the top 20 in only the high-throughput interactome of the coarse map and percentage of genes from the coarse map associated with the processes.

3. Discussion

Extensive research has been dedicated towards the elucidation of the processes CFTR undergoes from its transcription on the way to becoming a folded, complex glycosylated and fully functional ion channel at the plasma membrane of apical epithelial cells. In the CFTR Lifecycle Map, we collected the knowledge accumulated by researchers over the course of three decades and represented it in a human-and machine-readable way. In doing so, we adhered to the standards established by the systems biology community. We applied common literature curation criteria and followed the well-established SBGN format, as well as MIRIAM guidelines to create and annotate a core map of the CFTR lifecycle. Additionally, we included data from large-scale efforts to identify the CFTR interactome in a coarse map as a second data layer.

When comparing the manually curated data from small-scale experiments within the core model with the data from high-throughput screens by Wang et al., 2006, Pankow et al., 2015, Santos et al., 2019 and Matos et al., 2019 [45–48] embedded into the coarse

model, it becomes evident that the overlap between them is surprisingly small. To detect whether genes from the core and the coarse model belong to shared functional categories, BioInfoMiner was also used to prioritize genes in the core model and in the high-throughput generated interactome of the coarse model. BioInfoMiner is a tool used for the analysis of semantic networks and prioritizes key systemic processes and related genes present in a set of genes based on different databases [53]. Here, functionalities were assigned to the CFTR interaction partners based on Gene Ontology biological process terms [54,55] and the terms from the Reactome Pathway Database [56]. Gene Ontology provides information on the function of genes and gene products, which is derived from the scientific literature by the Gene Ontology Consortium. The Reactome Pathway Database is a manually curated and peer-reviewed database for molecular pathways by an international multidisciplinary team. The overlap between core model and coarse model genes detected by the Gene Ontology categories indicate that similar pathways of importance for CFTR folding and maturation have been identified by the targeted analysis of interactors (core map) or by hypothesis-free high-throughput analysis (coarse map). Using the Gene Ontology term analysis, two genes were prioritized among the top ten for both datasets. Both of them, HSP90AA1, better known as heat shock protein 90, and the ER ATPase VCP were also found amongst the hubs in the protein–protein interaction network of the manually curated interactors. Both VCP and HSP90AA1 are involved in a multitude of different cellular processes. It is therefore not surprising that they are highly connected in the protein–protein interaction network and are ranked as important genes in the Gene Ontology term analysis. They most likely also play an important role in CFTR maturation; however, they are very unspecific and likely influence the folding and maturation of proteins other than CFTR as well.

In contrast to the overlap in Gene Ontology categories, where entries of the core and the cores model were shared, the pathway terms from the Reactome Pathway Database differ substantially between the manually curated core model and the high-throughput-derived coarse model gene lists, which reinforces the poor overlap of only 46 interactors present in both layers of the CFTR Lifecycle Map. These differences on the small-scale level may stem from the different experimental approaches used to identify the interactors. While the manual curation from small-scale experiments includes a lot of information on the immature CFTR, the high-throughput screens use large-scale co-immunoprecipitation-based approaches to identify the interactome of the mature, or at least folded, wt-CFTR. These results, however, also indicate that, although a lot is known about the maturation and processing of CFTR, there are still substantial knowledge gaps. This is especially true for the interaction with other ion channels at the plasma membrane. In the last decade, other ion channels that may be in regulatory interaction with CFTR, such as the calcium-activated chloride channel Anoctamin-1 (also known as ANO1, TMEM16A and ORAOV2), or other chloride channels which are being discussed as CFTR alternatives, have been brought to the center of attention [57–59]. However, there is still work to be carried out in the elucidation of the CFTR regulatory network and its role in ion homeostasis at the plasma membrane, especially considering that potentially interacting ion channels, such as CLCA2 (also known as CACC3) and KCNJ1 (also known as ROMK) had to be rejected during the construction of the CFTR map due to a lack of experimental evidence. Knowledge gaps also seem to exist with regard to the intracellular trafficking pathway. It is known that CFTR is usually trafficked through conventional Golgi-mediated exocytosis, but may also circumvent the Golgi via an unconventional route [60–62]. This may be relevant for the rescue of misfolded variants such as F508del-CFTR [62–65]. Differences in the observed pathways and interactions may also arise from the variety of cell models used, which we addressed by color-coding interactors from polarized and non-polarized cells in our model. However, it becomes clear that the molecular mechanisms and interactors of both pathways are not fully elucidated yet, as can be seen in the lack of detail in the CFTR core map.

The CFTR Lifecycle Map is part of the CandActCFTR project, which established a publicly available database of candidate cystic fibrosis therapeutics, combining data from different sources, such as high-throughput- and small scale screens, data from relevant

databases and unpublished primary data (candactcftr.ams.med.uni-goettingen.de/). The CFTR Lifecycle Map, as a second part of the project, aims to provide the means to identify promising drug targets and elucidate the mode of action for candidate substances. Furthermore, it will be used to predict possible additive effects of different substance combinations. It thereby simultaneously provides a backbone to structure available data as well as a tool to develop hypotheses regarding novel therapeutics.

4. Materials and Methods

4.1. Creation of the Core Map

The map was drawn using the Process Description (PD) language of the Systems Biology Graphical Notation (SBGN) syntax [44] and the diagram editor CellDesigner4.4.2 [49,50]. The PD language of SBGN allows the representation of molecular biological models in a standardized manner using nodes and edges. The nodes, i.e., the molecular entities, of the network are represented by different symbols according to their molecular species. This map includes entities specified as genes, mRNA, proteins, truncated proteins, protein complexes, receptors, ion channels, simple molecules, and ions. The edges between molecular entities serve to specify reactions and reaction regulations. Reaction types present in the map include transcription, translation, state transition, complex formation and dissociation and transport, and reaction regulations are of the type catalysis, inhibition, physical stimulation, or modulation, where the exact nature of the interaction is unknown. Each molecular entity is named according to its HUGO-approved gene symbol [66] (https://www.genenames.org/ (accessed on 26 January 2021)) for genes and gene products, and ChEBI name [67] (www.ebi.ac.uk/chebi/ (accessed on 26 January 2021)) for simple molecules and ions. The model is split into six main cellular compartments (cytoplasm, plasma membrane, extracellular space, nucleus, endoplasmic reticulum (ER), and Golgi apparatus) and smaller compartments (vesicles and endosomes). Following the Minimal Information Required for the Annotation of Models (MIRIAM) Guidelines [68], an established standard for annotating systems biology models, all components are annotated by unique identifiers. The HUGO ID [66] (https://www.genenames.org/ (accessed on 26 June 2021)) is used for genes, and gene products, the UniProt ID [69] (www.uniprot.org (accessed on 26 January 2021)) for proteins, the PubChem CID [70] (pubchem.ncbi.nlm.nih.gov/ (accessed on 26 January 2021)) and ChEBI ID [67] (www.ebi.ac.uk/chebi/ (accessed on 26 January 2021)) for simple molecules. Furthermore, all interactors are annotated by the PMID (pubmed.ncbi.nlm.nih.gov (accessed on 6 June 2021)) of the references they are derived from. We define interactors as molecular entities that influence CFTR either through direct physical interactions or indirect regulatory interactions.

In order to provide additional information, all interactors in the map are color-coded. Blue entities represent CFTR or versions thereof (gene, mRNA, protein). All interactors that could be confirmed in at least one polarized cell line are depicted using a green color scheme, and interactors only identified in non-polarized cell lines are shown in yellow.

4.2. Literature Curation for the Core Map

In order to assess the current state of knowledge for CFTR maturation and activity, a literature overview was created, starting from 23 relevant main reviews of the last two decades, which can be found in Supplementary Table S13. Highly cited examples of these reviews are Riordan, 2008 [71], Lukacs and Verkman, 2012 [25] and Farinha and Canato, 2017 [72], which cover the expression, folding, maturation and function of CFTR in great detail and were published over the course of nearly one decade, thereby providing extensive descriptions of recent as well as earlier findings. From there, a preliminary consensus network of CFTR-relevant pathways and a list of resulting interaction partners was created. To validate the list of interactors, the literature database PubMed [73] was searched for experimental publications that confirmed the interaction through small-scale experiments using methods such as co-immunoprecipitation, pull-down assays, two-hybrid assays, NMR, X-ray crystallography, surface plasmon resonance, and functional

(mutational) assays. A complete list of the 221 references published between 1991 and 2020, which were used for validation, can be found in Supplementary Table S14. An interactor was accepted for the model when it could be identified in at least two small-scale experiments conducted in human cells from independent references or was considered as acknowledged by the research community when described in at least two reviews from different research groups. For example, the interaction of CFTR with the gene product of *SLC9A3R1* (also known as NHRF1 or EBP50) at the plasma membrane is well documented in experimental publications as well as reviews [72,74–76], and it was therefore accepted for the CFTR core map. On the other hand, the serum response factor (SRF) has been reported to act as a transcription factor for CFTR by René et al. [77] but, to our knowledge, has not yet been confirmed by other research groups. Consequently, it was not included in the CFTR core map for now. This does not mean that SRF is not considered a CFTR interactor; it is merely a quality control check to ensure a high data quality and high evidence of the interactions compiled in the core map.

4.3. Integration of Protein–Protein Interaction Databases

To ensure that no high-confidence interactors were missed, the manually curated list of interactors was then complemented with data from small-scale experiments from existing protein–protein interaction databases (SIGNOR2.0 [78], BioGRID [79], the Human Protein Reference Database [80], String DB [81], MINT [82], InnateDB [83], APID [84] and IntAct [85]). The same literature criteria were applied. References to large-scale experiments were excluded from the database searches for the core map and later added as a second data layer in the coarse map (see Section 4.5).

4.4. Consideration of Cell Polarity

The proper maturation and integration of functional CFTR into the plasma membrane is highly dependent on the polarization of the cell [51]. In order to take this into account in our model, the experimental method was specified and the cell lines used were listed for each interaction. Cell lines were divided into non-polarized and polarized cells and it was specified for each interaction whether or not it was shown in at least one polarized cell line. Under experimental conditions, the question whether cells are to be characterized as polarized or non-polarized can often be a matter of debate, which is beyond the scope of this study. Therefore, polarized cells are here defined as cells with the general ability to polarize. A complete list of interactions and the respective publications with the experimental method and cell line used can be found in Supplementary Tables S2–S6. During the literature curation, interactors were successively added to the list and model, which was visualized using the diagram editor CellDesigner4.4.2 [49,50] and the cell-polarity was indicated by color. Therefore, our model rudimentarily takes into account that a range of different model systems was used to study the CFTR lifecycle.

4.5. Visualization of the High-Throughput Interactome as Coarse Map

In the last 15 years, many potential interactors of CFTR have been identified through high-throughput methods [45–48]. While these methods are able to generate high amounts of data, they provide less information on the nature of the interaction and the confidence of the interaction is lower than through the identification in several small-scale studies, as more false-positives may occur. In order to address this discrepancy in data quality and still provide an extensive interaction map, the wt-CFTR interactomes published by Wang et al., 2006, Pankow et al., 2015, Santos et al., 2019 and Matos et al., 2019 [45–48] were added as separate data layer from large-scale experiments. Only the wt-specific interactome was considered, and interactors unique to the F508del-variant were excluded. Missing information on the subcellular localization of the interactors reported was gathered from UniProt and the Human Protein Atlas, using the respective application programming interface, to update to the knowledge on subcellular localization in 2021. The subcellular localizations specified in the published data were combined into more general main

localizations (nucleus, endoplasmic reticulum (ER), Golgi apparatus, endosome, plasma membrane, mitochondrion, cytoplasm, and extracellular space). Based on the functional categorization specified in the published data, the interactors were grouped into general functional categories (DNA replication, transcription, RNA processing, translation and folding, ER-associated degradation (ERAD), trafficking pathway, cytoskeleton and stabilization, activity and regulation, recycling, degradation, immune response, and other). The functional categorization was solely based on the information provided in the publications and not inferred from other sources. Interactors already present in the core map (46 interactors) were excluded in the coarse map to avoid redundancies between the data layers. A complete list of the interactors from the large-scale experiments, together with their subcellular localization and functional categorization, can be found in Supplementary Table S15. The list of interactors was split according to their functional categorizations and visualized in individual maps using the SBGN Activity Flow notation through the Python library libsbgnpy 0.2.2 [52] and CellDesigner4.4.2 [49,50].

Through this procedure, the findings from high-throughput efforts are available in the model, while the different data-quality, compared to the manually validated interaction partners, becomes visible for the user. While the high-throughput data have a lower confidence, they are also free from prior assumptions, whereas the small-scale experiments offer a high confidence but stem from restricting hypothesis-driven approaches.

4.6. Analysis of the Protein–Protein Interaction Network within the CFTR Core Map

To analyze the manually curated core map with regard to cross-interactions between the proteins included, a protein–protein interaction network was created using the list of genes present in the core map. Physical interactions between the manually curated interactors of CFTR were identified in GeneMania [86] using only physical interactions reported by BioGrid-small scale studies [79]. The complete list of interactions can be found in Supplementary Table S7. The network was visualized using the Python plotting library Matplotlib [87] and analyzed as a weighted, undirected graph using the Python packages NetworkX [88] and python-louvain [89].

4.7. Comparison of Content Provided to the Model by Small-Scale and Large-Scale Experiments

BioInfoMiner [53] was used to analyze and compare the gene lists derived from the manually curated core map and from the coarse map containing the interactomes published by Wang et al. [45], Pankow et al. [46], Santos et al. [47] and Matos et al. [48]. The BioInfoMiner tool provides a topological analysis of semantic networks and prioritizes key systemic processes and related genes present in a set of genes. Here, analysis based on Gene Ontology [54,55] and the Reactome Pathway Database [56] were conducted. The Reactome Pathway Database represents molecular pathways that are part of human biological processes, while Gene Ontology also assigns broader biological process terms to genes and gene products. Hence, the Reactome Pathway Database mainly associates genes and gene products to rather specific terms (e.g., *N-glycan trimming in the ER and Calreticulin cycle*), whereas they might be assigned to more general terms in Gene Ontology (e.g., *Protein Folding*).

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22147590/s1, Glossary, Table S1: gene lists of core map and coarse map, Tables, S2–S6: lists of interactors from core map, split by submaps, Table S7 protein–protein interactions between proteins in the core map, Tables S8–S12: comparison and systemic analysis of core map and coarse map (Sheet 1: overrepresented systemic processes from Gene Ontology and Reactome Pathway Database, Sheet 2: overrepresented genes from core map from Gene Ontology-based analysis, Sheet 3: overrepresented genes from core map from Reactome Pathway Database-based analysis, Sheet 4: overrepresented genes from coarse map from Gene Ontology-based analysis, Sheet 5: overrepresented genes from coarse map from Reactome Pathway Database-based analysis), Tables S13 and S14: lists of references for primary literature overview (Table S13) and validation of interactors (Table S14), Table S15: list

of interactors from coarse map, CoreMaps.zip: PDF and .xml files of whole cell core maps and individual submaps, CoarseMaps.zip: PDF and .xml files of coarse map submaps.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Supplementary Material. Future versions of the CFTR Lifecycle map will be made available here https://candactcftr. ams.med.uni-goettingen.de/SystemsBiology/ (accessed on 26 January 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bobadilla, J.L.; Macek, M.; Fine, J.P.; Farrell, P.M. Cystic fibrosis: A worldwide analysis of CFTR mutations—Correlation with incidence data and application to screening. *Hum. Mutat.* **2002**, *19*, 575–606. [CrossRef]
2. O'Sullivan, B.P.; Freedman, S.D. Cystic fibrosis. *Lancet* **2009**, *373*, 1891–1904. [CrossRef]
3. Elborn, J.S. Cystic fibrosis. *Lancet* **2016**, *388*, 2519–2531. [CrossRef]
4. Riordan, J.R.; Rommens, J.M.; Kerem, B.S.; Alon, N.O.A.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.I.; Plavsic, N.; Chou, J.L.; et al. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* **1989**, *245*, 1066–1073. [CrossRef]
5. O'Riordan, C.R.; Lachapelle, A.L.; Marshall, J.; Higgins, E.A.; Cheng, S.H. Characterization of the oligosaccharide structures associated with the cystic fibrosis transmembrane conductance regulator. *Glycobiology* **2000**, *10*, 1225–1233. [CrossRef]
6. Cheng, S.H.; Gregory, R.J.; Marshall, J.; Paul, S.; Souza, D.W.; White, G.A.; O'Riordan, C.R.; Smith, A.E. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **1990**, *63*, 827–834. [CrossRef]
7. Higgins, C.F.; Gallagher, M.P.; Mimmack, M.L.; Pearce, S.R. A family of closely related ATP-binding subunits from prokaryotic and eukaryotic cells. *BioEssays* **1988**, *8*, 111–116. [CrossRef] [PubMed]
8. Higgins, C. Export-import family expands. *Nature* **1989**, *340*, 342. [CrossRef] [PubMed]
9. Csanády, L.; Vergani, P.; Gadsby, D.C. Structure, gating, and regulation of the CFTR anion channel. *Physiol. Rev.* **2019**, *99*, 707–738. [CrossRef]
10. Zhang, Z.; Liu, F.; Chen, J. Molecular structure of the ATP-bound, phosphorylated human CFTR. *Proc. Natl. Acad. Sci. USA* **2018**. [CrossRef] [PubMed]
11. Meng, X.; Clews, J.; Ciuta, A.D.; Martin, E.R.; Ford, R.C. CFTR structure, stability, function and regulation. *Biol. Chem.* **2019**, *400*, 1359–1370. [CrossRef]
12. Aleksandrov, A.A.; Aleksandrov, L.A.; Riordan, J.R. CFTR (ABCC7) is a hydrolyzable-ligand-gated channel. *Pflugers Arch. Eur. J. Physiol.* **2007**, *453*, 693–702. [CrossRef]
13. Moran, O. The gating of the CFTR channel. *Cell. Mol. Life Sci.* **2017**, *74*, 85–92. [CrossRef]
14. Callebaut, I.; Chong, P.A.; Forman-Kay, J.D. CFTR structure. *J. Cyst. Fibros.* **2018**, *17*, S5–S8. [CrossRef]
15. Pranke, I.M.; Sermet-Gaudelus, I. Biosynthesis of cystic fibrosis transmembrane conductance regulator. *Int. J. Biochem. Cell Biol.* **2014**, *52*, 26–38. [CrossRef]
16. Cystic Fibrosis Mutation Database. Available online: http://www.genet.sickkids.on.ca/ (accessed on 26 January 2021).
17. Welcome to CFTR2 | CFTR2. Available online: https://www.cftr2.org/ (accessed on 26 January 2021).
18. Sosnay, P.R.; Siklosi, K.R.; Van Goor, F.; Kaniecki, K.; Yu, H.; Sharma, N.; Ramalho, A.S.; Amaral, M.D.; Dorfman, R.; Zielenski, J.; et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. *Nat. Genet.* **2013**, *45*, 1160–1167. [CrossRef] [PubMed]
19. Welsh, M.J.; Smith, A.E. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* **1993**, *73*, 1251–1254. [CrossRef]
20. Rowe, S.M.; Miller, S.; Sorscher, E.J. Cystic Fibrosis. *N. Engl. J. Med.* **2005**, *352*, 1992–2001. [CrossRef] [PubMed]
21. Veit, G.; Avramescu, R.G.; Chiang, A.N.; Houck, S.A.; Cai, Z.; Peters, K.W.; Hong, J.S.; Pollard, H.B.; Guggino, W.B.; Balch, W.E.; et al. From CFTR biology toward combinatorial pharmacotherapy: Expanded classification of cystic fibrosis mutations. *Mol. Biol. Cell* **2016**, *27*, 424–433. [CrossRef]
22. Zielenski, J.; Tsui, L.C. Cystic fibrosis: Genotypic and phenotypic variations. *Annu. Rev. Genet.* **1995**, *29*, 777–807. [CrossRef]
23. Zielenski, J. Genotype and Phenotype in Cystic Fibrosis. *Respiration* **2000**, *67*, 117–133. [CrossRef]

24. Kerem, B.S.; Rommens, J.M.; Buchanan, J.A.; Markiewicz, D.; Cox, T.K.; Chakravarti, A.; Buchwald, M.; Tsui, L.C. Identification of the cystic fibrosis gene: Genetic analysis. *Science* **1989**, *245*, 1073–1080. [CrossRef]

25. Lukacs, G.L.; Verkman, A.S. CFTR: Folding, misfolding and correcting the ΔF508 conformational defect. *Trends Mol. Med.* **2012**, *18*, 81–91. [CrossRef]

26. Okiyoneda, T.; Barrière, H.; Bagdány, M.; Rabeh, W.M.; Du, K.; Höhfeld, J.; Young, J.C.; Lukacs, G.L. Peripheral protein quality control removes unfolded CFTR from the plasma membrane. *Science* **2010**. [CrossRef]

27. Wang, W.; Okeyo, G.O.; Tao, B.; Hong, J.S.; Kirk, K.L. Thermally unstable gating of the most common cystic fibrosis mutant channel (ΔF508): "Rescue" by suppressor mutations in nucleotide binding domain 1 and by constitutive mutations in the cytosolic loops. *J. Biol. Chem.* **2011**, *286*, 41937–41948. [CrossRef]

28. Martiniano, S.L.; Sagel, S.D.; Zemanick, E.T. Cystic fibrosis: A model system for precision medicine. *Curr. Opin. Pediatr.* **2016**, *28*, 312–317. [CrossRef]

29. Southern, K.W.; Patel, S.; Sinha, I.P.; Nevitt, S.J. Correctors (specific therapies for class II CFTR mutations) for cystic fibrosis. *Cochrane Database Syst. Rev.* **2018**. [CrossRef]

30. Pedemonte, N.; Lukacs, G.L.; Du, K.; Caci, E.; Zegarra-Moran, O.; Galietta, L.J.V.; Verkman, A.S. Small-molecule correctors of defective ΔF508-CFTR cellular processing identified by high-throughput screening. *J. Clin. Investig.* **2005**. [CrossRef]

31. Van Goor, F.; Hadida, S.; Grootenhuis, P.D.J.; Burton, B.; Cao, D.; Neuberger, T.; Turnbull, A.; Singh, A.; Joubran, J.; Hazlewood, A.; et al. Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proc. Natl. Acad. Sci. USA* **2009**. [CrossRef]

32. Berg, A.; Hallowell, S.; Tibbetts, M.; Beasley, C.; Brown-Phillips, T.; Healy, A.; Pustilnik, L.; Doyonnas, R.; Pregel, M. High-Throughput Surface Liquid Absorption and Secretion Assays to Identify F508del CFTR Correctors Using Patient Primary Airway Epithelial Cultures. *SLAS Discov.* **2019**. [CrossRef]

33. De Wilde, G.; Gees, M.; Musch, S.; Verdonck, K.; Jans, M.; Wesse, A.S.; Singh, A.K.; Hwang, T.C.; Christophe, T.; Pizzonero, M.; et al. Identification of GLPG/ABBV-2737, a novel class of corrector, which exerts functional synergy with other CFTR modulators. *Front. Pharmacol.* **2019**, *10*. [CrossRef]

34. Merkert, S.; Schubert, M.; Olmer, R.; Engels, L.; Radetzki, S.; Veltman, M.; Scholte, B.J.; Zöllner, J.; Pedemonte, N.; Galietta, L.J.V.; et al. High-Throughput Screening for Modulators of CFTR Activity Based on Genetically Engineered Cystic Fibrosis Disease-Specific iPSCs. *Stem Cell Reports* **2019**. [CrossRef]

35. Van Goor, F.; Hadida, S.; Grootenhuis, P.D.J.; Burton, B.; Stack, J.H.; Straley, K.S.; Decker, C.J.; Miller, M.; McCartney, J.; Olson, E.R.; et al. Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809. *Proc. Natl. Acad. Sci. USA* **2011**. [CrossRef]

36. Phuan, P.W.; Veit, G.; Tan, J.A.; Finkbeiner, W.E.; Lukacs, G.L.; Verkman, A.S. Potentiators of defective DF508-CFTR gating that do not interfere with corrector action. *Mol. Pharmacol.* **2015**. [CrossRef]

37. Carlile, G.W.; Robert, R.; Goepp, J.; Matthes, E.; Liao, J.; Kus, B.; Macknight, S.D.; Rotin, D.; Hanrahan, J.W.; Thomas, D.Y. Ibuprofen rescues mutant cystic fibrosis transmembrane conductance regulator trafficking. *J. Cyst. Fibros.* **2015**. [CrossRef]

38. Liang, F.; Shang, H.; Jordan, N.J.; Wong, E.; Mercadante, D.; Saltz, J.; Mahiou, J.; Bihler, H.J.; Mense, M. High-Throughput Screening for Readthrough Modulators of CFTR PTC Mutations. *SLAS Technol.* **2017**. [CrossRef]

39. Giuliano, K.A.; Wachi, S.; Drew, L.; Dukovski, D.; Green, O.; Bastos, C.; Cullen, M.D.; Hauck, S.; Tait, B.D.; Munoz, B.; et al. Use of a High-Throughput Phenotypic Screening Strategy to Identify Amplifiers, a Novel Pharmacological Class of Small Molecules That Exhibit Functional Synergy with Potentiators and Correctors. *SLAS Discov.* **2018**. [CrossRef]

40. Van Der Plas, S.E.; Kelgtermans, H.; De Munck, T.; Martina, S.L.X.; Dropsit, S.; Quinton, E.; De Blieck, A.; Joannesse, C.; Tomaskovic, L.; Jans, M.; et al. Discovery of N-(3-Carbamoyl-5,5,7,7-tetramethyl-5,7-dihydro-4H-thieno[2,3-c]pyran-2-yl)-lH-pyrazole-5-carboxamide (GLPG1837), a Novel Potentiator Which Can Open Class III Mutant Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Channels to a High Extent. *J. Med. Chem.* **2018**, *61*, 1425–1435. [CrossRef] [PubMed]

41. Veit, G.; Xu, H.; Dreano, E.; Avramescu, R.G.; Bagdany, M.; Beitel, L.K.; Roldan, A.; Hancock, M.A.; Lay, C.; Li, W.; et al. Structure-guided combination therapy to potently improve the function of mutant CFTRs. *Nat. Med.* **2018**. [CrossRef] [PubMed]

42. Wang, X.; Liu, B.; Searle, X.; Yeung, C.; Bogdan, A.; Greszler, S.; Singh, A.; Fan, Y.; Swensen, A.M.; Vortherms, T.; et al. Discovery of 4-[(2R,4R)-4-({[1-(2,2-Difluoro-1,3-benzodioxol-5-yl)cyclopropyl]carbonyl}amino)-7-(difluoromethoxy)-3,4-dihydro-2H-chromen-2-yl]benzoic Acid (ABBV/GLPG-2222), a Potent Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Corrector for the Treatment of Cystic Fibrosis. *J. Med. Chem.* **2018**, *61*, 1436–1449. [CrossRef] [PubMed]

43. Pedemonte, N.; Tomati, V.; Sondo, E.; Galietta, L.J.V. Influence of cell background on pharmacological rescue of mutant CFTR. *Am. J. Physiol. Cell Physiol.* **2010**, *298*. [CrossRef]

44. Novère, N.L.; Hucka, M.; Mi, H.; Moodie, S.; Schreiber, F.; Sorokin, A.; Demir, E.; Wegner, K.; Aladjem, M.I.; Wimalaratne, S.M.; et al. The Systems Biology Graphical Notation. *Nat. Biotechnol.* **2009**, *27*, 735–741. [CrossRef]

45. Wang, X.; Venable, J.; LaPointe, P.; Hutt, D.M.; Koulov, A.V.; Coppinger, J.; Gurkan, C.; Kellner, W.; Matteson, J.; Plutner, H.; et al. Hsp90 Cochaperone Aha1 Downregulation Rescues Misfolding of CFTR in Cystic Fibrosis. *Cell* **2006**. [CrossRef]

46. Pankow, S.; Bamberger, C.; Calzolari, D.; Martínez-Bartolomé, S.; Lavallée-Adam, M.; Balch, W.E.; Yates, J.R. Δf508 CFTR interactome remodelling promotes rescue of cystic fibrosis. *Nature* **2015**, *528*, 510–516. [CrossRef]

47. Santos, J.D.; Canato, S.; Carvalho, A.S.; Botelho, H.M.; Aloria, K.; Amaral, M.D.; Matthiesen, R.; Falcao, A.O.; Farinha, C.M. Folding Status Is Determinant over Traffic-Competence in Defining CFTR Interactors in the Endoplasmic Reticulum. *Cells* **2019**, *8*, 353. [CrossRef]

48. Matos, A.M.; Pinto, F.R.; Barros, P.; Amaral, M.D.; Pepperkok, R.; Matos, P. Inhibition of calpain 1 restores plasma membrane stability to pharmacologically rescued Phe508del-CFTR variant. *J. Biol. Chem.* **2019**, *294*, 13396–13410. [CrossRef]

49. Funahashi, A.; Morohashi, M.; Kitano, H.; Tanimura, N. CellDesigner: A process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO* **2003**. [CrossRef]

50. Funahashi, A.; Matsuoka, Y.; Jouraku, A.; Morohashi, M.; Kikuchi, N.; Kitano, H. CellDesigner 3.5: A versatile modeling tool for biochemical networks. *Proc. IEEE* **2008**. [CrossRef]

51. Hollande, E.; Fanjul, M.; Chemin-Thomas, C.; Devaux, C.; Demolombe, S.; Van Rietschoten, J.; Guy-Crotte, O.; Figarella, C. Targeting of CFTR protein is linked to the polarization of human pancreatic duct cells in culture. *Eur. J. Cell Biol.* **1998**. [CrossRef]

52. König, M. matthiaskoenig/libsbgn-python: 0.2.2. *Zenodo* **2020**. [CrossRef]

53. Pilalis, E.; Valavanis, I.; Chatziioannou, A. e-NIOS BioInfoMiner. Available online: https://bioinfominer.com/login (accessed on 26 January 2021).

54. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef]

55. The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **2021**. [CrossRef]

56. Jassal, B.; Matthews, L.; Viteri, G.; Gong, C.; Lorente, P.; Fabregat, A.; Sidiropoulos, K.; Cook, J.; Gillespie, M.; Haw, R.; et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **2020**. [CrossRef]

57. Mall, M.A.; Galietta, L.J.V. Targeting ion channels in cystic fibrosis. *J. Cyst. Fibros.* **2015**, *14*, 561–570. [CrossRef] [PubMed]

58. Mall, M.A.; Danahay, H.; Boucher, R.C. Emerging concepts and therapies for mucoobstructive lung disease. *Ann. Am. Thorac. Soc.* **2018**, *15*, S216–S226. [CrossRef] [PubMed]

59. Danahay, H.L.; Lilley, S.; Fox, R.; Charlton, H.; Sabater, J.; Button, B.; McCarthy, C.; Collingwood, S.P.; Gosling, M. TMEM16A Potentiation: A Novel Therapeutic Approach for the Treatment of Cystic Fibrosis. *Am. J. Respir. Crit. Care Med.* **2020**. [CrossRef]

60. Rennolds, J.; Boyaka, P.N.; Bellis, S.L.; Cormet-Boyaka, E. Low temperature induces the delivery of mature and immature CFTR to the plasma membrane. *Biochem. Biophys. Res. Commun.* **2008**. [CrossRef]

61. Luo, Y.; McDonald, K.; Hanrahan, J.W. Trafficking of immature ΔF508-CFTR to the plasma membrane and its detection by biotinylation. *Biochem. J.* **2009**. [CrossRef]

62. Gee, H.Y.; Noh, S.H.; Tang, B.L.; Kim, K.H.; Lee, M.G. Rescue of Δf508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell* **2011**. [CrossRef]

63. Yoo, J.S.; Moyer, B.D.; Bannykh, S.; Yoo, H.M.; Riordan, J.R.; Balch, W.E. Non-conventional trafficking of the cystic fibrosis transmembrane conductance regulator through the early secretory pathway. *J. Biol. Chem.* **2002**. [CrossRef]

64. Gee, H.Y.; Kim, J.Y.; Lee, M.G. Analysis of conventional and unconventional trafficking of CFTR and other membrane proteins. *Methods Mol. Biol.* **2015**. [CrossRef]

65. Piao, H.; Kim, J.; Noh, S.H.; Kweon, H.S.; Kim, J.Y.; Lee, M.G. Sec16A is critical for both conventional and unconventional secretion of CFTR. *Sci. Rep.* **2017**. [CrossRef] [PubMed]

66. Braschi, B.; Denny, P.; Gray, K.; Jones, T.; Seal, R.; Tweedie, S.; Yates, B.; Bruford, E. Genenames.org: The HGNC and VGNC resources in 2019. *Nucleic Acids Res.* **2019**, *47*, D786–D792. [CrossRef]

67. Hastings, J.; Owen, G.; Dekker, A.; Ennis, M.; Kale, N.; Muthukrishnan, V.; Turner, S.; Swainston, N.; Mendes, P.; Steinbeck, C. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.* **2016**. [CrossRef] [PubMed]

68. Le Novère, N.; Finney, A.; Hucka, M.; Bhalla, U.S.; Campagne, F.; Collado-Vides, J.; Crampin, E.J.; Halstead, M.; Klipp, E.; Mendes, P.; et al. Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat. Biotechnol.* **2005**, *23*, 1509–1515. [CrossRef]

69. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515. [CrossRef]

70. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **2021**, *49*, D1388–D1395. [CrossRef]

71. Riordan, J.R. CFTR Function and Prospects for Therapy. *Annu. Rev. Biochem.* **2008**. [CrossRef]

72. Farinha, C.M.; Canato, S. From the endoplasmic reticulum to the plasma membrane: Mechanisms of CFTR folding and trafficking. *Cell. Mol. Life Sci.* **2017**, *74*, 39–55. [CrossRef]

73. PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/ (accessed on 26 January 2021).

74. Naren, A.P.; Cobb, B.; Li, C.; Roy, K.; Nelson, D.; Heda, G.D.; Liao, J.; Kirk, K.L.; Sorscher, E.J.; Hanrahan, J.; et al. A macromolecular complex of beta 2 adrenergic receptor, CFTR, and ezrin/radixin/moesin-binding phosphoprotein 50 is regulated by PKA. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 342–346. [CrossRef]

75. Li, J.; Dai, Z.; Jana, D.; Callaway, D.J.E.; Bu, Z. Ezrin Controls the Macromolecular Complexes Formed between an Adapter Protein Na+/H+ Exchanger Regulatory Factor and the Cystic Fibrosis Transmembrane Conductance Regulator. *J. Biol. Chem.* **2005**, *280*, 37634–37643. [CrossRef]

76. Loureiro, C.A.; Matos, A.M.; Dias-Alves, Â.; Pereira, J.F.; Uliyakina, I.; Barros, P.; Amaral, M.D.; Matos, P. A molecular switch in the scaffold NHERF1 enables misfolded CFTR to evade the peripheral quality control checkpoint. *Sci. Signal.* **2015**, *8*, ra48. [CrossRef]

77. René, C.; Taulan, M.; Iral, F.; Doudement, J.; L'Honoré, A.L.; Gerbon, C.; Demaille, J.; Claustres, M.; Romey, M.C. Binding of serum response factor to cystic fibrosis transmembrane conductance regulator CArG-like elements, as a new potential CFTR transcriptional regulation pathway. *Nucleic Acids Res.* **2005**. [CrossRef] [PubMed]
78. Licata, L.; Lo Surdo, P.; Iannuccelli, M.; Palma, A.; Micarelli, E.; Perfetto, L.; Peluso, D.; Calderone, A.; Castagnoli, L.; Cesareni, G. SIGNOR 2.0, the SIGnaling Network Open Resource 2.0: 2019 update. *Nucleic Acids Res.* **2020**, *48*, D504–D510. [CrossRef] [PubMed]
79. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **2021**, *30*, 187–200. [CrossRef] [PubMed]
80. Keshava Prasad, T.S.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; et al. Human Protein Reference Database—2009 update. *Nucleic Acids Res.* **2009**, *37*. [CrossRef] [PubMed]
81. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**. [CrossRef]
82. Licata, L.; Briganti, L.; Peluso, D.; Perfetto, L.; Iannuccelli, M.; Galeota, E.; Sacco, F.; Palma, A.; Nardozza, A.P.; Santonico, E.; et al. MINT, the molecular interaction database: 2012 Update. *Nucleic Acids Res.* **2012**. [CrossRef] [PubMed]
83. Breuer, K.; Foroushani, A.K.; Laird, M.R.; Chen, C.; Sribnaia, A.; Lo, R.; Winsor, G.L.; Hancock, R.E.W.; Brinkman, F.S.L.; Lynn, D.J. InnateDB: Systems biology of innate immunity and beyond—Recent updates and continuing curation. *Nucleic Acids Res.* **2013**. [CrossRef] [PubMed]
84. Alonso-López, D.; Campos-Laborie, F.J.; Gutiérrez, M.A.; Lambourne, L.; Calderwood, M.A.; Vidal, M.; De Las Rivas, J. APID database: Redefining protein-protein interaction experimental evidences and binary interactomes. *Database* **2019**. [CrossRef]
85. Orchard, S.; Ammari, M.; Aranda, B.; Breuza, L.; Briganti, L.; Broackes-Carter, F.; Campbell, N.H.; Chavali, G.; Chen, C.; Del-Toro, N.; et al. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **2014**. [CrossRef]
86. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*. [CrossRef] [PubMed]
87. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
88. Hagberg, A.A.; Schult, D.A.; Swart, P.J. Exploring Network Structure, Dynamics, and Function using NetworkX. In Proceedings of the 7th Python in Science Conference, Pasadena, CA USA, 19–24 August 2008; Varoquaux, G., Vaught, T., Millman, J., Eds.; Los Alamos National Lab.: Walnut Creek, CA, USA, 2008; pp. 11–15.
89. Aynaud, T. Python-Louvain x.y: Louvain Algorithm for Community Detection 2020. Available online: https://github.com/taynaud/python-louvain (accessed on 26 January 2021).