Genetics of cystic fibrosis: CFTR mutation classifications toward genotype-based CF therapies.
Pascale Fanen, Adeline Wohlhuter-Haddad, Alexandre Hinzpeter

To cite this version:

HAL Id: inserm-01059046
http://www.hal.inserm.fr/inserm-01059046
Submitted on 29 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Genetics of Cystic Fibrosis:

CFTR Mutation Classifications towards Genotype-Based CF Therapies

Pascale Fanen\textsuperscript{1,2,3*}, Adeline Wohlhuter-Haddad\textsuperscript{2,4}, Alexandre Hinzpeter\textsuperscript{1,2}

\textsuperscript{1}. Université Paris-Est, UMR_S955, UPEC, F-94000, Créteil, France; \textsuperscript{2}. Inserm, U955, Equipe 11, F-94000, Créteil, France; \textsuperscript{3}. AP-HP, Hôpital H. Mondor - A. Chenevier, Département de Génétique, F-94000, Créteil, France; \textsuperscript{4}. McGill University, Faculty of Medicine, Department of Physiology, Montréal, QC, Canada.

*Corresponding author: Pascale Fanen, MD, PhD, INSERM U955, Equipe 11 Hôpital Henri Mondor Créteil, F-94010, France
Tel: +33 1 49 81 28 54/ +33 1 49 81 48 22
Fax: +33 1 48 99 33 45
E-mail: pascale.fanen@inserm.fr

Running title: Genetics of CF

Keywords: CFTR, mutation, CFTR-RD, therapy, Variants of Unknown significance

Abbreviations: CBAVD, congenital bilateral absence of the vas deferens; CF, cystic fibrosis; CFTR, Cystic Fibrosis Transmembrane conductance Regulator; CFTR-RD, CFTR related disease; COPD, chronic obstructive pulmonary disease; EJC, exon junction complex; NBD, nucleotide-binding domains; NGS, next generation sequencing; NPD, nasal potential difference measurements; NMD, Nonsense-Mediated mRNA Decay; PABP, Poly(A)-binding protein; PI, pancreatic insufficiency; PTC, premature termination codon; PS, pancreatic sufficiency; VUs, Variants of Unknown significance.
Abstract

Cystic fibrosis (CF) is an autosomal recessive disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which encodes an epithelial anion channel. Since the identification of the disease in 1938 and up until 2012, CF patients have been treated exclusively with medications aimed at bettering their respiratory, digestive, inflammatory and infectious symptoms. The identification of the CFTR gene in 1989 gave hopes of rapidly finding a cure for the disease, for which over 1,950 mutations have been identified. Since 2012, recent approaches have enabled the identification of small molecules targeting either the CFTR protein directly or its key processing steps, giving rise to novel promising therapeutic tools.

This review presents the current CFTR mutation classifications according to their clinical consequences and to their effect on the structure and function of the CFTR channel. How these classifications are essential in the establishment of mutation-targeted therapeutic strategies is then discussed. The future of CFTR-targeted treatment lies in combinatory therapies that will enable CF patients to receive a customized treatment.
1. Introduction

Mutations disrupting the function of the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene, also referred to as ABCC7, cause Cystic Fibrosis (CF), a severe autosomal recessive disease. CF affects 1 in 2,500 to 4,500 newborns among the Caucasian population. All other ethnic groups are affected to a lesser extent. In 1989, Francis Collins, Lap-Chee Tsui and John R. Riordan (Kerem et al., 1989) identified the CFTR gene and its most frequent mutation, F508del, as being the cause of CF. The CFTR gene is located on the long arm of chromosome 7 (7q31.2) and its structure comprises 27 exons, spanning over 190kb (Fig. 1A). After splicing of the large introns, the CFTR mRNA is 6.5kb and all of its exons are necessary to produce a functional CFTR protein.

Since 1990, the laboratories dedicated to CF genetic diagnosis and research worldwide have been grouped into the international Cystic Fibrosis Genetic Analysis Consortium (CFGAC). This consortium identifies and describes mutations of the CFTR gene. Inputs are then deposited on the CF mutation database “CFTR1” (http://www.genet.sickkids.on.ca/cftr/) for public access. Since the discovery of the CFTR gene, extensive sequencing enabled the identification of over 1,950 different mutations. Recently, a second CF mutation database “CFTR2” for “Clinical and Functional Translation of CFTR”, has been created (http://www.cftr2.org/). It gathers up-to-date information about newly discovered CFTR gene mutations, classifies them as disease-causing, neutral or mutation of varying clinical consequences, and provides clinical information about specific mutation combinations.

To better our understanding of CF pathogenesis, as well as to facilitate diagnosis and choice of treatment, two classification systems for CFTR mutations have been created, one according to clinical consequences, the other according to CFTR structure-function assessment. Currently available CF therapies mainly treat the symptoms of the disease, and have increased the mean life expectancy of CF patients from 5 years in the 1970s to about 35 to 40 years of age today. To address the underlying cause of the disease, recent approaches have enabled the identification of small molecules targeting either the CFTR protein directly, or its key processing steps. These novel therapeutic tools correct specific defects responsible for CFTR protein loss-of-function and need to be adapted to each patient’s genotype. This highlights the importance of thoroughly characterizing each patient at the molecular level.
This review first presents the clinical spectrum of Cystic Fibrosis, as well as the two existing classification systems for CFTR mutations. How CFTR mutations are assigned to CF patients is then described. The use of this essential information in the development of genotype-based therapeutic strategies is then discussed.

2. Classification of CFTR mutations

2.1. The clinical spectrum of CF

The first description of the disease was done in 1938 by Dorothy Hansine Andersen, describing an abnormal pancreas, which presented with cysts and fibrosis (Andersen, 1938). The clinical spectrum of CF has since greatly expanded, giving rise to diagnoses of classic and non-classic CF presented in Figure 2 (De Boeck et al., 2006, Farrell et al., 2008).

Although classic CF most often presents with a severe multi-organ phenotype, and non-classic CF with milder single-organ phenotypes, this is not always the case. Whether the phenotype is "severe/mild" or "multi/single-organ" is not intrinsically linked to the classic or non-classic CF diagnosis. These are subjective terms, which should be used for descriptive purposes exclusively.

The great phenotypic variability of CF has been shown to implicate not only the type of CFTR mutations, but also other genetic factors such as modifier genes, (see chapter from Guillot and Corvol in this special issue) and environmental factors (lifestyle, treatment) (De Boeck et al., 2006). In the diagnostic algorithms proposed by the European Consensus Group, patients with one or more phenotypic characteristic suggestive of CF are classified according to their pilocarpine sweat test result (chloride >60 mmol/L, 30–60 mmol/L or <30 mmol/L) (Castellani et al., 2008). To date, no test has proven to be as practical or reliable as the sweat test for clinical diagnostic purposes.

Classic CF – The majority of CF patients suffer from classic CF, yet their organs are affected to varying degrees. Patients are diagnosed with classic CF if they have one or more phenotypic characteristic and a sweat chloride concentration of >60 mmol/L (Fig. 2). They may have exocrine pancreatic insufficiency (PI) or pancreatic sufficiency (PS). The disease can have a severe course with rapid progression of symptoms or a milder course with very little deterioration over time. Classic CF patient mortality is mainly due to progressive respiratory disease. From a genetic standpoint, classical CF is characterized by one established CF-causing mutation on each CFTR allele.

Non-classic CF – Non-classic CF describes individuals with at least one CF phenotypic
characteristic and a normal (<30 mmol/L) or borderline (30-60 mmol/L) sweat chloride level, in whom detection of one mutation on each CFTR allele, or direct quantification of CFTR dysfunction by NPD measurement has been confirmed (Boyle, 2003, Rosenstein, 2003) (Fig. 2). Non-classic CF patients have either multi- or single-organ involvement. Most of them have exocrine PS and milder lung disease. As described by the CF Diagnostic Working Group, “some patients with single-organ involvement resulting from CFTR dysfunction may be more appropriately given an alternative diagnostic label, as recommended in the World Health Organization diagnostic list” (WHO, 2000). These alternative diagnoses for non-classic CF, or CFTR-related disorders, include isolated obstructive azoospermia, chronic pancreatitis, disseminated bronchiectasis, allergic bronchopulmonary aspergillosis, diffuse panbronchiolitis and sclerosing cholangitis, and all have at least one identified associated CFTR mutation, some of which are discussed hereafter.

2.2. Classification of CFTR mutations based on clinical consequences

In 2007, the Consensus Conference, organized by the European Cystic Fibrosis Society with the partnership of the European Society of Human Genetics and the EuroGentest Network of Excellence, allowed for the establishment of a classification of CFTR mutations, into the following four groups, according to their clinical consequences (Table 1). Group A comprises mutations that cause classic CF (CF-causing), Group B includes mutations that cause non-classic CF (associated to CFTR-related disorders), Group C comprises mutations with no known clinical consequences, and Group D consists of mutations with unknown or uncertain clinical relevance (also referred to as VUs) (Castellani et al., 2008). Only a few of the hundreds of CFTR mutations identified to date have been shown to result in classic CF. Within a given ethnic group, specific mutations may be more frequent than others. Most CF patients are Caucasian, however, other ethnicities are also affected, and p.Phe508del remains the most frequent classic CF mutant allele, accounting for approximately 70% of worldwide chromosomes (Lucotte et al., 1995). As more and more patients present with non-classic phenotypes, DNA screens for common CFTR mutations have become essential for CF diagnosis. To that effect, the American College of Medical Genetics (ACMG) has recommended a core panel of 23 mutations to assess CF risks in diagnosis, carrier screening and sequencing analyses (Table 2). These 23 mutations fall into Group A, as classic CF-causing mutations, with the exception of p.Arg117His, which is discussed later in a paragraph dedicated to complex
alleles. This core panel may be complemented with the screening of additional ethnic-specific mutations found in European and/or other selected population. Some mutations may be associated with different clinical presentations. Groups A and B are therefore not mutually exclusive. For example, the p.Asp1152His mutation may result in isolated congenital bilateral absence of the vas deferens (CBAVD), or in a fully expressed CF lung disease with pancreatic sufficiency. Non-classic CF disorders have a pathophysiology that may be attributed to CFTR defect, as these patients are characterized by a prevalence of CFTR mutation higher than that expected in the general population.

Group D mutations have unknown or uncertain clinical relevance to CF phenotypic characteristics. It is often difficult to assess the association between CFTR mutations and clinical entities that, in some cases, may have other non-CFTR related aetiologies. Other factors, such as the environment, modifier genes and the progression of the disease with age, can also affect the clinical heterogeneity of patients carrying these large spectrum mutations.

It is important to note that CFTR genotype is not a useful predictor of the severity of lung disease in the individual patient, and should not be used as an indicator of prognosis (Castellani et al., 2008).

2.3. Classification of CFTR mutations based on CFTR structure-function

Five major classes of CFTR mutations were created according to their consequences on CFTR function (Welsh and Smith, 1993): mutations interfering with protein synthesis, mutations affecting protein maturation, mutations altering channel regulation, mutations affecting chloride conductance, and mutations reducing the level of normally functioning CFTR at the apical membrane (Welsh and Smith, 1993) (Fig. 3). A sixth class consisting of mutations decreasing the stability of CFTR present at the plasma membrane or affecting the regulation of other channels was proposed, but has been poorly investigated (Haardt et al., 1999). These class VI mutations are now combined with class V, as mutations leading to a reduced amount of functional CFTR protein (Welsh et al., 2001).

Class I: mutations preventing the production of a full-length CFTR protein

Class I mutations result in the total or partial lack of production of a functional CFTR protein (Fig. 3). Such mutations may arise either due to (1) a nucleotide substitution introducing an in-frame premature termination codon (PTC) -UAA, UAG or UGA-, (2) frame-shifting insertions or
deletions, (3) mutations at the invariant dinucleotide splice junctions, or introduction of a PTC, resulting in complete skipping of an out of frame exon (Fig. 1B), (4) a complete or partial deletion of the CFTR gene or (5) a rearrangement in the gene altering the exon sequence. Mutations that generate PTCs can reduce the steady-state level of mRNA via nonsense-mediated decay (NMD), which degrades the abnormal PTC-containing mRNA (Maquat, 1995). Such a phenomena has been shown with p.Gly542*, p.Arg553*, p.Trp1282*, and one frameshift mutation (p.Phe316LeufsX12), but a notable exception is p.Arg1162*, which produces a normal CFTR mRNA level as compared to wild-type CFTR (Will et al., 1995).

The p.Gly542* mutation is the most frequent worldwide mutation of its class, affecting at least one allele of up to 4% of CF patients. Other mutations reach a higher frequency in specific populations, due to a founder effect; for example 48% of all alleles in Ashkenazi Jews in Israel carry the p.Trp1282* mutation and 24% carry the p.Tyr122* mutation in the French Reunion Island (Dugueperoux et al., 2004, Shoshani et al., 1992).

Class II: mutations altering the cellular processing of the protein

Class II mutations are associated with defective processing due to misfolding of the protein (Fig. 3) and may be found within any domain of the CFTR protein. The misfolded protein is retained in the Endoplasmic Reticulum (ER), retro-translocated to the cytoplasm, where it is degraded by the ubiquitin/proteasome pathway (Jensen et al., 1995, Ward et al., 1995). CFTR maturation through the cellular compartments can be assayed by western blot. The core-glycosylated CFTR (band B) migrates at a lower molecular weight than the fully-glycosylated CFTR protein (band C). Depending on the mutation, one may observe either a partial reduction (p.Leu206Trp, third transmembrane segment) or a complete absence (p.Arg1066Cys, fourth intracellular loop or p.Phe508del, NBD1) of mature CFTR. The most frequent CF mutation, p.Phe508del, has been well studied. It has been shown that p.Phe508del leads to energetic and kinetic instability of the NBD1 domain due to improper local folding (Protasevich et al., 2010). Additionally, CFTR domain assembly is altered, namely the interaction between NBD1 and the fourth cytoplasmic loop (CL4) (Cui et al., 2007, He et al., 2010). Correction of both NBD1 energetics and interface instability is therefore required to restore p.Phe508del processing (Mendoza et al., 2012, Rabeh et al., 2012). Initially, mutations belonging to this class were thought to cause severe CF similarly to p.Phe508del or p.Arg1066Cys, but p.Leu206Trp was shown to be associated with variable phenotype (Clain et al., 2005a).
Class III: mutations disturbing the regulation of the Cl⁻ channel

These mutations are frequently located in the ATP binding domains (NBD1 and NBD2) and are referred to as gating mutations (Fig. 3). They are missense mutations producing a protein efficiently inserted in the membrane at normal levels, but resistant to activation by protein kinase A. The main example is p.Gly551Asp, which abolishes ATP-dependent gating, resulting in an open probability that is ~100-fold lower than that of the wild type channel (Anderson and Welsh, 1992). Other class III mutations such as the frequent mutations p.Arg560Thr or p.Gly970Arg are listed in Figure 3 (Seibert et al., 1996). Noteworthy, this latter mutation is located in the third cytoplasmic loop (CL3) of the protein, highlighting the fact that knowing the localization of the amino acid substitution is insufficient to predict the resulting CFTR structural and functional anomalies.

Class IV: mutations altering the conduction of the Cl⁻ channel

These mutations are mostly located within membrane spanning domains implicated in the constitution of the channel pore (Fig. 3). The missense mutations located in these regions produce a protein efficiently inserted in the membrane, which retains a cAMP-dependent Cl⁻ channel activity, but with a reduced channel conductance. Alleles in this class are typically associated with a milder form of the disease. The p.Arg117His mutation is the best-characterized class IV mutation. While this mutant is correctly processed and generates cyclic AMP-regulated apical Cl⁻ currents, patch-clamp analysis demonstrated a reduction in both channel open probability and conductance (Reddy and Quinton, 2001, Sheppard et al., 1993).

Recently, a subclass IVb (b stand for bicarbonate) has been proposed for mutations affecting CFTR bicarbonate conductance specifically, without reducing its chloride conductance (Schneider et al., 2011). As of today, only a few of such mutants have been identified, namely p.Arg75Gln and p.Ile148Thr (Choi et al., 2001, Schneider et al., 2011). While p.Ile148Thr appears not to be deleterious (Claustres et al., 2004), p.Arg75Gln was found in some studies at a higher frequency in cohorts of patients presenting CFTR-RD (Divac et al., 2004, Schneider et al., 2011). Other mutations affecting the CFTR bicarbonate conductance are bound to be identified, probably in patients presenting with a CFTR-RD.

Class V: mutations reducing the amount of functional CFTR protein

Most of the class V mutations reduce the total amount of CFTR protein by affecting pre-mRNA splicing (Fig. 3). These splice site mutations can induce complete or partial exclusion of
an exon. In the latter case, production of normal mRNA is maintained, but in lower quantity. When the skipped exon is in phase (Fig. 1A), the transcript will produce an incomplete non-functional channel. On the other hand, when the skipped exon is out of phase, a PTC will rapidly appear and the transcript will be degraded by the NMD pathway. Classically, splicing mutations are located in the introns close to the splice sites, where they alter key splicing signals implicated in proper exon recognition (acceptor site, donor site, branch point or polypyrrimidine tract). Yet, a growing number of them are being identified within exons. These exonic nucleotide substitutions can induce exon skipping by either disrupting Exonic Splicing Enhancer (ESE) or creating Exon Splicing Silencer (ESS) motifs (Pagani and Baralle, 2004). Since direct RNA analysis is not routinely performed, the number of mutations causing splicing defects is underestimated.

The most frequent and well-studied is the skipping of exon 10 (formerly named exon 9). The presence or absence of this exon is correlated with a polymorphism within the polypyrrimidine tract located upstream of the acceptor splice site (Poly-T tract). While a Poly-T tract of 7T or 9T ensures up to 90% of normal splicing, the shorter 5T Poly-T tract allows no more than 10 to 40% of normal mRNA to be produced (Chu et al., 1993).

Complex alleles and multi-class mutations

Some CFTR alleles have the particularity of harbouring two distinct mutations. They are referred to as complex alleles. The two mutations can induce additive defects, leading to a more severe phenotype than would each mutation separately (Clain et al., 2001, Clain et al., 2005b). This can be illustrated by the class IV R117H mutation (c.350G>A, p.Arg117His) whose severity is modulated in cis by the 5/7/9T polypyrromidine tract (c.1210-12T(5_9)in intron 9. While the R117H-T7 genotype is associated with milder forms of CF such as CBAVD, and most of the time even absence of symptoms, the R117H-T5 can be identified in patients having elevated sweat chloride and clinical cystic fibrosis, which in some cases is severe (Thauvin-Robinet et al., 2009).

Finally, some point mutations have been shown to alter multiple distinct processes, such as splicing and routing or maturation and channel conductance. This is for example the case of the p.Phe508del mutation, which presents, in addition to a processing defect, a gating defect characterized by a lower open probability of the channel and a reduced stability of the protein at the cell surface due to an increased recycling rate (Dalemans et al., 1991).

2.4. Assigning CFTR mutations to CF patients: a methodological perspective
As of today, only a minority of the 1950 known mutations have been analyzed at the protein level (Ferec and Cutting, 2012) and even fewer at the mRNA level. Studying the DNA/RNA obtained from clinically characterized CF patients has led to the identification of their mutations. With the upcoming introduction of NGS, the number of identified VUs will be greatly increased (Abou Tayoun et al., 2013, Trujillano et al., 2013). The labelling of these VUs, as disease-causing or as neutral (Group C) mutations, will be challenging. Presented hereafter are some guidelines and available tools enabling such a distinction.

One such tool is the CFTR2 database, initiated to establish disease liability of individual CFTR mutations, as well as to determine the relationships between specific mutations and symptoms of CF. This project (http://www.cftr2.org/) has compiled clinical data from nearly 40,000 CF patients and relates the clinical phenotype (pancreatic status, lung function, pulmonary infection) to the electrophysiological phenotype (sweat chloride and ion transport in epithelial cells transfected with the corresponding mutated CFTR). Among the 1221 mutations reported so far on the CFTR2 database, 190 mutations have already been characterized. For the remaining mutations, indirect evidence of pathogenic potential may be derived from the following (adapted from (Castellani et al., 2008)):

1) *in vivo* assessment of abnormal CFTR-mediated chloride transport, as assessed by sweat test or NPD and *ex vivo* assessment by short circuit measurements using Ussing chambers (nasal brushing or rectal biopsy),

2) molecular assessment of sequence variations

*CFTR* mutations with high confidence for pathogenic potential:
- change in an amino acid residue predicted to affect CFTR synthesis and/or function
- premature termination signal (insertion, deletion or nonsense mutation)
- alteration of the invariant dinucleotide of the donor or acceptor splice sites.

*CFTR* mutations with lower confidence for pathogenic potential:
- sequence variation detected in several unrelated individuals with CF
- change in an amino acid residue absent in at least 100 non-CF chromosomes of healthy carriers from the same ethnic group
- change in a highly evolutionarily conserved amino acid residue
- sequence variation creating/unmasking a cryptic splice site
- exonic or intronic sequence variation modifying splicing.
When the observed criteria point to a mutation with lower confidence, additional tools and approaches are necessary to confirm or invalidate its pathogenic potential. The first approach is phenotype-driven, based on the analysis of well-phenotyped subjects with classic CF, thus helping to interpret the clinically relevant mutations (Sosnay et al., 2011). The second are bioinformatic tools, which are now extensively used in clinical diagnosis. However, the computational methods mainly based on evolutionary conservation and protein structure, have limitations (Dorfman et al., 2010). Notably, for the prediction of splicing defects, the use of a combination of tools is more powerful than individual ones (Aissat et al., 2012, Raynal et al., 2013). In silico predictions need to be validated by either in vivo measures (when possible) or in vitro assays. Many in vitro assays are available today, exploring the effect of mutations on specific cellular processes. Indeed, splicing defects are reproduced using minigenes, and defects in transcriptional level estimated by reporter gene assays. CFTR processing and stability are monitored by biochemical assays such as western blot and surface biotinylation. Protein localization can be visualized by immuno-cytochemistry and channel function measured by electrophysiological techniques (Ussing chamber, halide-sensitive assays or patch-clamp). While these tools enable to correctly identify major defects, subtle changes in the individual cellular processes tested are more difficult to interpret.

Indeed, it must be pointed out that 1-1.5 % of CF alleles remain undetermined in patients with classic CF (Castellani et al., 2008). The classification and assignment of CFTR mutations to CF patients is the basis for current and future genotype-based therapeutic strategies.

3. Current mutation-targeted therapeutic strategies

As our understanding of the genetics of CF grows, and as technological advances render molecular diagnostic tools more powerful, many CF researchers worldwide now focus on the development of drugs that aim to treat CF by targeting the underlying cause of the disease. These drugs include molecules inducing readthrough of premature termination codons, potentiator molecules, as well as corrector molecules, presented hereafter.

3.1. Molecules inducing readthrough of PTC

Premature termination codons usually induce the destruction of the PTC-harbouring transcript through the NMD pathway; however, some PTCs can lead to the production of a truncated protein. In the nucleus, the removal of the introns leaves an exon junction complex (EJC) located
at each exon-exon boundary. The removal of these complexes by the ribosome during the first round of translation serves to monitor the pre-mRNA. When a transcript harbours a PTC, the ribosome will stall at this particular position and leave downstream EJCs. The EJC will then recruit UPF1, via its interaction with other UPF proteins (UPF2 and UPF3/3X (or UPF3a or UPF3b)). This in turn will lead to the degradation of the PTC containing transcript by the NMD pathway (Amrani et al., 2006, Bidou et al., 2012, Maquat, 2004). This model predicts the recognition of PTCs located in all exons, with the exception of the last exon. Additionally, PTCs located in the last 50 nucleotides of the penultimate exon were shown to bypass NMD recognition, probably due to steric hindrances with the EJC. Hence, PTCs have been classified as either NMD-sensitive or NMD-irrelevant depending on their position. More recent data challenged this model, as some NMD-irrelevant transcripts were found to be degraded, whereas some NMD-sensitive transcripts were not. This led to the Faux-3’ UTR model, where mRNA degradation by NMD is dependent on the distance between the PTC and the Poly(A)-binding proteins (PABPs) (Eberle et al., 2008, Silva et al., 2008). According to this model, the latter distance will influence the kinetics of dissociation of the ribosome/polypeptide complex, affecting the recruitment of UPF1 (Muhlemann, 2008). In any case, most transcripts harbouring PTCs will be degraded, preventing the production of deleterious truncated proteins.

Small molecules have been shown to induce readthrough of PTCs. Readthrough therapy targets a molecular defect common to all genetic diseases. It could therefore treat a substantial proportion of patients as in-frame nonsense mutations account for 12% of all hereditary disease-causing mutations (Kellermayer, 2006). Studies have been performed in both in vitro models (over 45 different genes in more than 80 studies) and in clinical trials (8 different genes in 16 studies) (reviewed in (Lee and Dougherty, 2012)).

Compounds inducing readthrough favor an improper recognition of the PTC and the recruitment of a near cognate tRNA in the A site of the ribosome in place of the termination complex. If enough PTCs are recoded into sense-codons, enough full-length, functional protein may then be restored to provide a therapeutic benefit.

The best characterized drugs active against PTCs are aminoglycosides. Studies demonstrated the restoration of CFTR-dependent Cl− secretion and protein expression with improved clinical endpoints (Sermet-Gaudelus et al., 2007, Wilschanski et al., 2003). PTC124 (trade name Ataluren) was more recently identified in a high throughput screen (Welch et al., 2007). This
molecule was found to suppress nonsense mutations more efficiently than aminoglycosides without nephro- or oto-toxic side effects. A double-blind, placebo-controlled Phase III study is now under evaluation. Novel readthrough molecules are being identified, such as RTC13 and RTC14 (Du et al., 2009) or NB124 (Rowe et al., 2011). Interestingly, an NMD inhibitor, amlexanox, has recently been shown to induce readthrough (Gonzalez et al., 2012). Such drugs presenting dual activities could increase readthrough efficiency by increasing the amount of target transcripts.

Before these molecules can be used in routine therapy, certain aspects of their mechanisms of action still need to be understood. Among these, what will be their consequence on normal stop codons? It appears that the rapid dissociation of the terminating ribosome at the normal stop codon does not enable efficient readthrough. This is comforted by the absence of elongated proteins after PTC124 treatment (Welch et al., 2007). Additionally, if the ribosome decodes the 3’ UTR sequence, it will either rapidly encounter an in-frame stop codon or will decode the polyA-tail, leading to the synthesis of a poly-Lysine tail at the end of the nascent protein. This, in turn, initiates a quality control mechanism named non-stop decay, leading in fine to the degradation of both the polypeptide and the transcript (Graille and Seraphin, 2012).

Efficiency of readthrough therapy is linked to various parameters, some of which are predictable while others are measurable. Predictions can be made knowing the identity of the stop codon (UAA<UAG<UGA) and the local nucleotide sequence, where nucleotides in position -1 and +4 appear to have the greatest effect (Floquet et al., 2012). The amount of target transcripts, dependent on the level of NMD, also directly influences readthrough efficiency; hence a positive correlation has been obtained between Cl- transport modification and the level of increase in CFTR transcript after gentamicin nasal application (Linde et al., 2007). Skipping of the PTC harboring exon should also be taken into account as it will also reduce the amount of target transcripts (Hinzpeter et al., 2013). Indeed, taking into account these two parameters, it appeared that 8% of the CFTR transcripts would be targeted by readthrough therapy in a patient homozygous for the p.Glu831X mutation, which would be predictive of poor treatment efficiency (Hinzpeter et al., 2010).

Finally, readthrough of a PTC leads to the incorporation of a random amino acid at the considered position, leading to the production of a pool of full-length proteins bearing a maximum of one single amino acid substitutions. These substitutions will lead to either
functional or non-functional proteins. Therefore, PTCs located in key regions implicated in protein processing or function would be less efficiently corrected, even if high levels of expression can be achieved.

3.2. Potentiator therapies or activation of CFTR channels

Molecules enhancing channel activity, such as molecules belonging to the phytoestrogen family (genistein and isoflavone), have been identified. A high-throughput screen enabled the identification of VX-770, Ivacaftor (trade name Kalydeco, Vertex Pharmaceuticals). This molecule specifically targets the p.Gly551Asp gating mutation by increasing the channel open probability and consequently the flow of ions transported through the channel (Ramsey et al., 2011), (Eckford et al., 2012). It also has the ability to increase the open probability of the wild-type channel and other mutants, namely p.Phe508del (Van Goor et al., 2009) and p.Arg117His (Van Goor et al., 2012). This molecule will thus also benefit patients retaining low level wild-type CFTR expression, e.g. class V mutations. Finally, it has recently been proposed that a reduction of CFTR activity could lead to a CF-like acquired disease. Cigarette smokers presenting with or without Chronic Obstructive Pulmonary Disease (COPD) showed an increased sweat chloride concentration and reduced intestinal CFTR currents (Raju et al., 2013). This can be explained by a reduced transcriptional expression of CFTR (Cantin et al., 2006) and a reduced stability of the protein (Clunes et al., 2012). As CFTR potentiators have been shown to increase wild-type CFTR activity, it would be tempting to increase CFTR activity under such conditions.

3.3. Corrector therapies or routing CFTR to the plasma membrane

Bringing misfolded p.Phe508del protein to the cell surface is a major goal for the treatment of CF. Some chemical compounds have been reported to allow the protein to evade endoplasmic reticulum quality control, enabling surface expression. These molecules named correctors, have been identified by high-throughput screening, and include molecules such as benzo[c]quinolizinium (MPB) compounds, Corr4a or VX-809 (Grove et al., 2009, Pedemonte et al., 2005, Pedemonte et al., 2011, Ye et al., 2010). The most promising, VX-809 can restore up to 15% of non-CF channel activity in primary respiratory epithelia (Van Goor et al., 2011, Van Goor et al., 2006). Recently, VX-809 has been shown to correct processing mutations located within the N-terminus of the protein more efficiently. Indeed, some of these processing mutants could be fully corrected under the tested conditions as compared to a maximum 15% correction.
obtained for p.Phe508del (Ren et al., 2013). Certain correctors could therefore be best suited for specific genotypes.

It must be pointed out that all these correctors have a ceiling efficacy in restoring chloride conductance of approximately 10-15 % (Pedemonte et al., 2005, Van Goor et al., 2006). Efficiency could be increased by either combining molecules, by targeting distinct folding steps or by identifying new more potent correctors. Interestingly, some correctors have been shown to be active on other misfolded ABC proteins, namely on the p.Gln141Lys gout-causing mutation in the ABCG2 gene (Woodward et al., 2013). This knowledge opens the application of these drugs to other genetic syndromes.

3.4. Combinatory therapies to optimize efficiency

As mutations may induce several CFTR dysfunctions, distinct drugs could correct individual defects. For example, p.Phe508del CFTR, if located at the plasma membrane, also presents a gating defect (Dalemans et al., 1991). Therefore, potentiators aimed at normalizing defective p.Phe508del CFTR Cl⁻ channel gating may increase the effect of correctors, which rescue the traffic of the protein to the membrane. More generally, potentiators would be beneficial for many classes of mutation. Indeed, they could activate channels produced after suppression therapies for class I mutations or residual wild-type channels expressed in class V mutations. Splicing modulators, such as antisense oligonucleotides, could increase the amount of these latter channels. Combinatory therapies may also be considered for compound heterozygotes.

4. Conclusion

CF is a complex disease with a broader clinical and genetic spectrum than previously thought. This is making CF diagnosis (classic vs. non-classic CF), carrier screening and prenatal screening decisions more and more difficult.

New technologies and research approaches currently enable CF researchers to work at developing drugs that target the dysfunction of the CFTR protein directly. Our knowledge of the existing CFTR mutations and their consequences, both at the structure-function level (how and why the CFTR protein is non-functional) and at the clinical level (what organ(s) are most affected by the mutation), will be essential in order to perfect the design of the new molecules that have mainly been identified by high-throughput screening to date. This information is needed not only for the development of the molecules themselves, but also for the optimization of their delivery to the patient (pill vs. nebulizer etc.).
With our current knowledge, it is still very difficult to associate a specific CFTR mutation to a specific dysregulation of the structure and/or function of the CFTR protein. It is very unlikely that most CF patients may be treated with a single drug. As advances in CF research and development will provide solutions to these challenges, the future gold standard for CF treatment strategy will be a customized combinatory one, targeting both the clinical symptoms and the underlying genetic cause for each CF patient.

Acknowledgments

The authors thank Dr. Fabrice Lejeune for the critical reading of this manuscript and helpful discussions. The authors thank the French association Vaincre la Mucoviscidose for its continuous support.

References


Table 1. Classification of CFTR mutations according to their clinical consequences.
Examples of each of the four groups of mutations are given.

<table>
<thead>
<tr>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic-CF CF-causing mutations</td>
<td>Non-Classic CF CFTR-related disorder associated mutations</td>
<td>No Clinical Consequence</td>
<td>Unknown Clinical Relevance</td>
</tr>
</tbody>
</table>

* Mutations that may belong either to Group A or to Group B.
** Mutations that may belong either to Group B or to Group C.

Table 2. ACMG recommended panel of 23 classic CF-causing mutations. These mutations include missense, stop, splicing and frameshift mutations.

<table>
<thead>
<tr>
<th>23 ACMG recommended panel of classic CF-causing mutations</th>
</tr>
</thead>
</table>

Additional or alternative mutations present at significant frequencies in an ethnic population served by a newborn screening program may be assessed.
**Figure 1**

**A**

Rescaled exon scheme with minimized introns, which are not drawn to scale. At the bottom in normal font, the exon labeling numbers, 1 through 24, according to the old nomenclature. At the top in bold font, the exon labeling numbers, 1 through 27, according to the new nomenclature.

**B**

Correspondence numbering and reading frame for exons and nucleotides. In bold font, common numbering of exons, 1 through 27, according to the new nomenclature. First (start) and last (end) nucleotides are indicated for each exon. The major transcription start site (TSS) is indicated in grey at -133 base pairs from ATG (+1). The flat edges indicate the exons with start and end nucleotides aligned with the reading frame; the arrow tips indicate exon starts and ends with frame shifted nucleotides. cDNA RefSeq NM_000492.

**Fig. 1. Scheme of the CFTR gene.** (A) Rescaled exon scheme with minimized introns, which are not drawn to scale. At the bottom in normal font, the exon labeling numbers, 1 through 24, according to the old nomenclature. At the top in bold font, the exon labeling numbers, 1 through 27, according to the new nomenclature. (B) Correspondence numbering and reading frame for exons and nucleotides. In bold font, common numbering of exons, 1 through 27, according to the new nomenclature. First (start) and last (end) nucleotides are indicated for each exon. The major transcription start site (TSS) is indicated in grey at -133 base pairs from ATG (+1). The flat edges indicate the exons with start and end nucleotides aligned with the reading frame; the arrow tips indicate exon starts and ends with frame shifted nucleotides. cDNA RefSeq NM_000492.
Fig. 2. The Clinical Spectrum of Cystic Fibrosis. This diagram depicts the phenotypic characteristics of CF and distinguishes between classic and non-classic CF clinical descriptions.
Fig.3. Classification of CF mutations according to CFTR structure and function assessment.

Each of the five classes of mutations are described. A cell illustration representing the localization of mutated CFTR relative to normal / wild-type CFTR (cell at the very left) is shown for each mutation class, along with a description of the nature of the defect, potential therapies and frequent mutations for this particular class. In bold, the most prevalent mutation of each class.