Molecular basis of Fanconi anemia

MARIA D’APOLITO, LEOPOLDO ZELANTE, ANNA SAVOIA
Servizio di Genetica Medica, IRCCS - Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy

Abstract

Background and Objective. Fanconi anemia (FA) is an autosomal recessive disease characterized by pan- cytopenia, congenital malformation and high predisposition to developing malignancies. The phenotypical heterogeneity is associated with genetic heterogeneity: at least 8 complementation groups (FA-A to FA-H) have been identified, each group presumably corresponding to a separate disease gene (FAA to FAM). The FAA and FAC genes, which account for 75-80% of the patients, have been cloned. Their protein products have no significant homology to any known protein, or to each other, and may therefore represent elements of a new pathway. This review describes some of the recent contributions to the understanding of the molecular basis of FA.

Evidence and Information sources. The authors of the present review have been working in the field of FA for several years. In the present review they have critically examined articles published in journals covered by the Science Citation Index and Medline.

State of Art and Perspective(s). A variety of biochemical and cellular approaches have been used to determine the molecular defect of FA. Evidence for a possible role of the defective proteins in cell cycle regulation, apoptosis, or DNA stability have been reported. Recently, it has been demonstrated that FAA and FAC proteins bind each other and form a complex found in similar abundance in both cytoplasm and nucleus, suggesting a possible function in nucleus activities. Knowledge of the mutation spectrum occurring in the FA genes may contribute significantly to pathogenesis studies in FA and help to design mutation screening strategies. Further functional studies and the cloning of other FA genes will provide insights into the biological basis of FA and information for developing specific therapies for the disease.

©1998, Ferrata Storti Foundation

Key words: Fanconi anemia, complementation groups, FAA gene, FAC gene, mosaicism

In the last few years there have major achievements concerning Fanconi anemia (FA). This review describes some of the recent contributions to the understanding of the molecular basis of FA.

Clinical aspects and cellular phenotype

Clinical heterogeneity

Fanconi anemia is an autosomal recessive disease characterized by a variety of clinical symptoms and cellular features.1 FA is a rare disease with an estimated prevalence of 1-3 per million.2 However, prevalence may vary with geographic region and ethnic structure of the population. In white Africans a birth incidence of at least 1 in 22,000 has been attributed to founder effect.3 A high FA incidence has also been reported in Italy with areas of major density in the South, especially in Campania.4

Clinical manifestations are extremely pleomorphic and include bone marrow failure, congenital malformations, and predisposition to cancers.5 An important clinical hallmark, with a mean age of onset around 5 years, is a progressive pancytopenia initially characterized by thrombocytopenia and subsequently by granulocytopenia and anemia.6,7 Multiple congenital malformations have previously been reviewed.8-10 Clinical data from 370 patients of the International Fanconi Anemia Registry (IFAR) showed that more than half patients exhibit growth retardation, short stature, and abnormal skin pigmentation (cafe-au-lait spots). Skeletal malformations, particularly dysplasia of the radius and the thumb, represent a typical finding of the syndrome and are reported in about 50% of the patients. Other frequent abnormalities include ocular, kidney, and cardiac defects. Microcephaly is present in 36% of the patients while mental retardation accounts for 26% of the subjects. Hypogenitalism affects 20% of the male who show gonadal hypoplasia and defective spermatogenesis. Clinical features can be variably associated: 39% of the patients exhibit the most severe trait with both hematological and congenital defects, whereas only anemia or malformations affect 30% and 24%, respectively. A minor group of patients are clinically asymptomatic.8

The complex clinical trait of FA creates difficulties in differential diagnosis with other syndromes including, VACTERL association with hydrocephalus, Bailer-Gerold, thrombocytopenia with absence of radii, Diamond-Blackfan syndrome, and dyskeratosis congenita.11-15 Misdiagnosis raises the question of whether these syndromes are distinct disorders or phenotypic variations of the same disease.

Myelodysplastic syndromes (MDS) and acute
myeloblastic leukemia (AML) are the most severe complications in FA.7,16 Actuarial risk of MDS and AML was 52% by 40 years of age. However, a variety of solid tumors, such as basal cell carcinomas, gastrointestinal and gynecological cancers, are also observed.17,18 The average age for developing cancers is 15 years for leukemia, 16 years for liver tumor, and 23 years for other tumors.19

**Cellular phenotype**

Similar to clinical manifestations, cellular phenotype shows a wide spectrum of features. FA cells exhibit elevated levels of spontaneous chromosomal aberrations with damage mainly of chromatids, including gaps, breaks, and interchanges.20 Because of chromosomal instability, FA has been classified as a DNA repair disease and is often compared with other syndromes, such as ataxia telangiectasia, xeroderma pigmentosum, trichothiodystrophy, Cockayne syndrome, and Bloom syndrome.21-27 The aspect that differentiates FA among these disorders is the hypersensitivity of cells to cross-linking agents, such as diepoxybutane (DEB) and mitomycin (MMC).28,29 In the presence of these drugs, FA cells show an increased sensitivity to chromosomal damage. Clonal cytogenetic abnormalities, more frequently involving chromosomes 1, 7, and 11 are often associated with myelodysplastic syndrome or acute myeloblastic leukemia.7,30-32 FA cells show also an abnormal response to oxygen with reduced growth at standard oxygen levels, suggesting a defect in response to oxygen toxicity.33-36 Prolongation or complete arrest of G2 phase alters cell cycle progression in FA cells.37-38 Since affected cells exhibit a normal G1/S transition, the FA proteins are likely to regulate the transition from G2 to M phase.

**Diagnosis**

Although the diagnosis of FA is complicated by a considerable spectrum of clinical manifestations, the hypersensitivity of FA cells to cross-linking agents can be used as a cellular marker. A highly sensitive and specific test, named DEB test, has to be performed whenever FA diagnosis is suspected.39 The 7% of clinically asymptomatic individuals are diagnosed only on the basis of a positive DEB test, which is therefore pathognomonic for the disease. However, the test is not sufficiently sensitive for FA carrier detection. Alternatively, diagnosis can be performed by delayed passage through the G2 phase of the cell cycle.40

**Therapy**

Without curative support, spontaneous survival is poor, with death occurring during the second decade of life from aplastic anemia, leukemia or other cancers. Treatment of FA is similar to that of acquired aplastic anemia. FA patients respond transiently to therapy with androgens and cytokines, granulocyte and granulocyte-macrophage colony-stimulating fac-

**Genetics**

**Complementation groups**

The most peculiar characteristic of the cellular phenotype of FA cells is a high sensitivity to cross-linking agents. When lymphoblastoid cell lines derived from unrelated patients are fused, the hybrids can be studied for cross linking sensitivity. They may be either resistant and therefore have a normal phenotype (complementation) or sensitive and be phenotypically FA cells (no complementation). Patients whose cell lines do not complement each other in a fusion hybrid, belong to the same complementation group. To determine the extent of genetic heterogeneity in FA, somatic cell hybrids have been analyzed. The first two complementation groups were designated FA-A and non FA-A.41 As a continuation of this analysis, a further genetic heterogeneity was described in non FA-A with evidence for at least another three new subtypes, FA-B, FA-C, and FA-D.42 This study was based on the analysis of only seven cell lines so that, when a European collaborative study (EUFAR, European Fanconi anemia research) was established with the aim to characterize a hundred FA patients, evidence for a fifth (FA-E) complementation group was reported.43 To assess a further variability within subtype E, in which cell lines classified as non-ABCD were included, some fusion hybrids among FA-E cells indicated three new groups (FA-F, FA-G, and FA-H), providing evidence for a minimum of eight distinct complementation groups (Figure 1).45

The relative prevalence of FA-A and FA-C has been estimated at an average of about 65-69% and 12-18%, respectively but may widely vary according to ethnic background.46 In Italy, 11 out of 12 patients analyzed by cell fusion studies were assigned to group FA-A, suggesting an unusual high relative prevalence of this subtype.51 In the Netherlands, the majority (67%) of patients were found to belong to subtype FA-C.52 Only a few cell lines have been assigned to FA-B and FA-D whereas only a single patient is known to belong to each of the other groups FA-E to FA-H.
FA genes

Each of the eight complementation groups presumably corresponds to a distinct disease gene (FAA to FAH). This assumption is true at least for FAA and FAC which have been cloned and for FAD which maps on a different chromosome, 3p22-26. FAC was the first gene to be identified by functional complementation cloning, which consisted of transfecting recipient FA-C lymphoblastoid cell lines with a cDNA library and successive selection in medium supplemented with DEB and MMC. Since only cells that received the wild-type FAC gene were able to survive in these restricted conditions, the FAC gene was rescued from selected clones. The FAC gene is 4,566 bp and encodes a 63-kD protein of 558 amino acids. Differentially processed forms of 2.3, 3.2, and 4.6 kb, which differ in both the 5' and 3' untranslated regions (5'UTR and 3'UTR), are transcribed (Figure 2). The 5'UTR is characterized by the presence of two different exons, alternatively spliced, and the 3'UTR differs in size because of the recognition of alternative polyadenylation signals. The functional significance of these mRNA splice variants remains unknown. The FAC gene, greater than 120 kb (Savoia, unpublished data) with 14 exons, maps to chromosome 9q22.3, between two genes involved in diseases with high predisposition to cancers, xeroderma pigmentosum and nevoid basal cell carcinoma syndrome (Gorlin syndrome). FAC homologues from mouse, rat, and bovine have also been isolated. The murine Fac gene encodes a protein with a 78% similarity to the human polypeptide and, when expressed in human FA-C cells, it is able to correct the hypersensitivity to cross-linking drugs.

The FAA gene was identified by both functional and positional cloning. The latter strategy consists of two steps: first a gene is mapped to a subchromosomal region and then coding sequences from that region are isolated and analyzed for the presence of genetic alterations. The FAA gene was localized at chromosome 16q24.3 at an interval of 1 Mb by linkage analysis in a panel of families belonging to complementation group FA-A and identified by mutation analysis in candidate genes. The FAA gene has an open reading frame of 4,365 bp encoding a 16-kD protein of 1,455 amino acids. The predicted protein has possible nuclear localization signals and a partial leucine zipper consensus sequence, homologies consistent with a function of FAA in the nucleus. Northern blot analysis displays an extensive hetero-
geneity of FAA with transcripts of several sizes. In particular, besides a prominent mRNA of 4.7 kb, there are transcripts of 7.5, 3.0 and 2.0 kb in size. In addition, differentially processed forms due to alternative splicing have been characterized (d’Apolito, unpublished data). Similar to FAC, the functional significance of the messenger heterogeneity is unknown.

The genomic structure of the gene revealed that the coding sequence is contained in 43 exons spanning approximately 80 kb of genomic DNA.

Mutation analysis
In the FAC gene, ten pathogenic mutations have so far been identified (Figure 2). Except for two aminoacid substitutions, the other sequence variations would predict premature termination of the FAC protein. In particular, the L554P mutation completely abolished the activity of the FAC protein as demonstrated by functional complementation analysis using site-directed in vitro mutagenesis. However, a wider spectrum of FAC mutations has to be expected since only one mutation has been identified in several compound heterozygous patients.

The most common mutation is IVS4+4A→T, a splice site mutation leading to a frameshift deletion of exon 4, which has been found in the majority of FA Ashkenazi-Jewish patients. This mutation was also found to be homozygous in dizygotic twins with the VACTERL association, suggesting that appropriate investigations should be undertaken for differential diagnosis. Another common alteration is 322delG, a frameshift deletion found in Northern European individuals. The genotype-phenotype correlation revealed that the former was associated with severe clinical manifestation, whereas the latter was unexpectedly associated with a milder phenotype. In fact, cell lines carrying the 322delG mutation express a FAC isoform of 50 kD which is an in frame amino terminal truncated protein able to partially correct the FA phenotype.

Seventy-four different mutations have thus far been described in the FAA gene (Figure 3). They include missense, nonsense, splicing, and frameshift mutations, which are widely distributed over the gene, suggesting that mutational hot spots are not involved. Most of the mutations are expected, as in FAC, to cause a premature termination of the FAA protein. The mutation screening also detected a consistent number of variants that are probably merely polymorphisms. It has been suggested that FAA is hypermutable, and that slipped-strand mispairing is important for the generation of germline and somatic mutations in a variety of cancer-related genes. However, a substantial proportion of mutations was not detected, at least in the European population. This might be partially explained by the possible existence of rearrangements in the FAA gene, which might not be detectable by common screening procedures. In fact, in an Italian patient a more detailed analysis led to the characterization of the first two genomic deletions, one encompassing the entire FAA gene, the other of 5.0 kb spans from intron 17 to intron 21. The latter is likely to be the result of an unequal crossover between Alu repeats. FAA mutations are listed in the Fanconi Anemia Mutation Database (http://www.rockefeller.edu/fanconi/mutate/default.html)
they were found only in single patients. This situation hampers a detailed genotype-phenotype correlation study, which would otherwise be helpful in determining possible associations with a mild or severe form of the disease. A further difficulty is due to a wide clinical heterogeneity in FA not only among patients from different families of the same complementation group A but also among patients within one family.51

Mutations in the FAA and FAC genes account for an estimated 75–80% of all FA patients. Therefore, the classification of patients in subtypes by cell fusion, which is extremely laborious and time-consuming, may be replaced by screening for mutations in both genes. The search for mutations can be preceded by linkage analysis on chromosomes 16q24.3 and 9q22.3 to exclude unclassified families from subtypes FA-A and FA-C, respectively.78 Strong allelic association was detected with microsatellites at 16q24.3 in the Afrikaner FA patients, suggesting that a founder mutation has contributed to the high incidence of FA in this population.59–61

Mosaicism

Mosaicism, resulting from post-zygotic mutation, is a condition in which an individual has 2 or more cell lines of different genetic constitution. Approximately 25% of patients with Fanconi anemia have spontaneously occurring mosaicism due to two different populations of lymphocytes, one of which is hypersensitive to cross-linking agents while the other normally responds to these agents.81 If in one FA allele any event led to a normal copy of the gene, this event would confer wild type phenotype on the cell. This condition has already been described in other pathologies, such as Duchenne muscular dystrophy, Bloom’s syndrome, and adenosine deaminase deficiency.82–84

The molecular basis of the phenotypic reversion in FA has not yet been completely determined. However, evidence for two different mechanisms has been reported: mitotic recombination and somatic mutation. In one patient, a mitotic intragenic crossing over between the maternally and paternally inherited FAC chromosomes has determined the segregation of a wild-type allele.81 In another patient, a de novo insertion downstream of the germline FAA mutation restores the open reading frame possibly leading to a protein altered only for few amino acids (Savoia, unpublished results). Although mosaicism is under investigation, it is sometimes associated with a relatively mild hematological course of the disease with therefore several consequences and implication for diagnosis, bone marrow transplantation and gene therapy.

Molecular biology

FAA and FAC proteins have no significant homology to any known polypeptide, or to each other, and may therefore represent elements of a novel pathway. A variety of biochemical and cellular approaches have been used to determine the defective function in FA by studying different biological aspects thought to underlie the clinical and cellular features of the disease. Since several approaches could exhibit epiphenomena of the basic defect, the research has been focusing directly on FA genes and their products. The study of gene expression and regulation, the characterization of proteins, the development of animal models are some of the aspects that may provide insights into the molecular defect of the disease.

The FAA and FAC genes are ubiquitously expressed, although at different levels.54,59 In human adult tissues, the highest expression was detected in testis, thymus, and bone marrow which are actively dividing and differentiating tissues (d’Apolito, unpublished data). Similarly, transcription is elevated in human fetal tissues. However, while the FAC transcripts were found in similar abundance in all tissues, FAA mRNA level was higher in liver and kidney than in brain and lung, suggesting that the FAA and FAC expression is under the control of different mechanisms at least in fetal tissues (d’Apolito, unpublished data). Consistent with these data, the expression of the mouse homologous FAC gene during development was found coupled with cellular proliferation and differentiation and the congenital malformations of FA patients.85

The analysis of the promoter regions of both FAA and FAC identified highly GC-rich regulatory sequences, a feature of several housekeeping genes characterized by wide expression in human tissues.63,66 Functional analysis in FAC showed a complex transcriptional regulation possibly under the control of two separate promoters.86 Further studies, leading to the identification of the trans-acting factors, will provide advances in the knowledge of the FA genes. Recently, it has been described that the p53 binds the FAC promoter and, if overexpressed, it regulates the transcription positively.87 However, although p53 modulate the expression, it is not a prerequisite for the FAC transcription, suggesting a possible role only under certain conditions.

FAC is a soluble cytoplasmic protein associated with a multimeric complex.88–91 It has been found to bind a member of the chaperone family, GRP94, and to cyclin-dependent kinase, cdc2.92–93 FAC has also been detected in the cytoplasmic compartment and recently, it has been demonstrated that FAA and FAC proteins interact with each other and that, whereas unbound FAA and FAC are localized particularly in the cytoplasm, FAA–FAC complex is detected in both cytoplasm and nucleus.94–95

Analysis of mutations in FA patients has played an important role since it may identify functional domains. The C-terminal exon is necessary for the biological function of the FAC protein, which is not able to restore the cross-linking hypersensitivity when lacking at least the five C-terminal amino acids or when the fifth last leucine is replaced by proline.69,72
This indicates that the C-terminus is important for FAC activity, for instance in the binding with other proteins. In fact, the L554P mutant failed to interact to both FAA and cdc2 proteins.93,95

The development of a mouse model represents a powerful tool to understand the function of genes and to test new therapies. Two murine models have been created by knockout strategies.96-97 Both mutants showed similar phenotype: absence of developmental abnormalities and hematological defects, increased chromosomal aberrations in response to cross-linking agents, and G2 cell-cycle arrest. Homozygous male and female mice had compromised gametogenesis with markedly impaired fertility. The knockout mice replicates some human disease features, but they did not show pancytopenia during the first year of life. In addition, mice showed a hypersensitivity to interferon-γ (IFN-γ), a feature that may be associated with the bone marrow failure in FA.97

Possible function of the FA proteins

The basic defect in FA is not known. However, on the basis of clinical and cellular features and recent studies, some of the possible pathways, such as cell-cycle regulation, apoptosis or genomic stability, defective in FA are being delineated.

Evidence for a defect in cell cycle is supported by the prolonged G2 phase and the reduced proliferation that is enhanced by treatment with chemical cross-linking agents observed in FA cells. Studies of FAC expression during cell cycle showed that transcription is increased during S phase, is maximal at the G2/M transition, and declines during M phase.93 In addition, the FAC protein binds cdc2, a cyclin-dependent kinase that together with cyclins A and B directs the G2 progression.98 The G2 arrest and reduced proliferation of FA cells can be partially corrected by overexpression of SFAH, a member of the cyclin family necessary for cell cycle progression.99 Tissues with cells in active proliferation exhibit an increased FAC expression that, if repressed, correlates to cell growth inhibition at least in hematopoietic progenitors.85,100 Similar to other genetic diseases characterized by genomic instability, such as ataxia telangiectasia or Li-Fraumeni, FA may be the result of a defect in cell cycle regulation.101,102

Altered apoptosis has been indicated as another possible defect in FA. Congenital malformation and hematological abnormalities might be due to the tendency of FA cells for apoptosis especially after treatment with cross-linking agents that may cause DNA damage.103 In fact, FAC-retroviral transduced cells show significantly lower levels of apoptosis, so that FAC appears to be an inhibitor of apoptotic death.104 Among the possible mechanisms regulating apoptosis, the p53 independent-pathway seems not to be affected, even if controversial results have been reported.105-107 The fas activation by IFN-γ signals, another pathway controlling apoptosis, was thought to be defective because of the hypersensitivity to IFN-γ of hematopoietic progenitors from knockout mouse.97

In fact, overexpression of FAC protein partially reduce the number of these cells in apoptosis.108

Since FA cells are characterized by chromosomal instability and hypersensitivity to DNA cross-linking agents, a defective DNA repair has been suggested as a possible mechanism responsible for FA. An abnormally high production of intragenic deletions in the hypoxanthine phosphoribosyl-transferase (HPRT) gene has been reported to be an important feature of FA phenotype, suggesting that FA proteins play a role in the maintenance of genomic integrity.109 Among the different causes responsible for genomic rearrangements, the action of a site-specific mechanism has been suggested to be involved in FA cells.110 In another recent study, elevated levels of homologous recombination seem to contribute to genomic instability.111 For these reasons, it might be that the FA defect lies in a pathway that involves recombinational events during DNA replication or repair processes.

Although the molecular defect of FA remains unknown, it might be that FA proteins have a similar function, or work together, for instance, as part of a multimeric complex or participating in a common pathway.112 The recent report on the interaction between FAA and FAC proteins and consequent translocation of the complex to the nucleus support the model suggested by D’Andrea and Grompe.113 In FA, a signal transduction pathway is defective. FAC, as a sensor protein, interacts in the cytoplasm with FAA. This translocator, transferred to nucleus, is able to activate an inducer for the expression of proteins involved in the pathogenesis of FA.

Perspectives

The cloning of the FAA and FAC genes has provided an opportunity to understand the molecular basis of FA because only, with these gene, can new approaches be taken to explore the role of the altered proteins. Although the FA defect remains unknown, some progress has been made in several key areas, including gene expression, characterization and interaction of proteins, animal models and determination of pathways defective in FA. However, together with the clarification of the role of FAA and FAC proteins, the cloning of all the FA genes will be represent an important prerequisite in the understand pathogenesis of FA. Both functional and positional strategies are technically difficult: the former tends to yield many false-positive cDNA clones whereas the latter is hampered by the need of a large sample of families eligible for linkage analysis. This limitation can be overcome by an alternative approach leading to gene localization by mediated chromosome transfer, as reported for FAD gene.83

The identification of the genes also allows the screening for mutations. Knowledge of the mutation
spectrum in the FAA and FAC genes may contribute significantly to pathogenesis studies in FA. The identification of pathogenic alterations in FA patients allows also for the assignment of patients to the respective complementation group and thus removes laborious complementation studies. For patients who do not belong to FA-A or FA-C, complementation analysis is still the only way to assess their subtype. In addition, since FA is a clinically heterogeneous disease and diagnosis on the basis of symptoms alone is difficult and often unreliable, knowledge of the mutations allows for an accurate diagnosis and for rapid prenatal identification of an affected fetus. Although the diagnosis is facilitated by the study of induced chromosomal instability, the DEB test is not reliable in individuals with mosaicism and in testing carriers from general population. In both cases, screening for mutations is a powerful approach for molecular analysis.

The cloning of the FAA and FAC gene is leading to the establishment of protocols for gene therapy. FA is a good candidate disease for this new therapeutic trial because in vitro transfected cells have a selective growth advantage on FA cells. In addition, different target cells, either stem cells or cord blood and peripheral blood cells, can be easily purified and in vitro manipulated for gene transfer. Retroviral vectors were able to fully complement the cellular defect of both FA-C and FA-A cells. However, several problems remain unsolved, including the integration of the gene, the level of integration necessary for the correction of the disease, the long-term expression of the transfecting gene, the selective growth advantage of transfected cells, and the function of the FA proteins.

Contributions and Acknowledgments

AS was responsible for the conception of this review article. All the authors contributed to the analysis of the literature and writing of the paper. The authors wish to thank Achille Iolascon for helpful discussions and comments and Adriana Zatterale of the Italian Registry of Fanconi anemia (RIAF) for important contributions.

Disclosures

Conflict of interest: none.

Redundant publications: no substantial overlapping with previous papers.

Manuscript processing

Manuscript received January 27, 1998; accepted March 31, 1998.

References

23. Taylor AM, Byrd PJ, McConville CM, Thacker S. Gene-
32. Fundia A, Gorla N, Larripa I. Spontaneous chromosome aberrations in Fanconi's anemia patients are located at fragile sites and acute myeloid leukemia breakpoints. Hereditas 1994; 120:47-50.
33. Joenje H, Hoehn H. Fanconi anemia mutation caus-
35. Touraine RL, Bertrand Y, Foray P, Gilly J, Philippe N. In vivo hyper-
sensitivity to oxygen of Fanconi anemia (FA) cells is linked to ex vivo evidence for oxidative stress in FA homozygotes and heterozygotes. Blood 1997; 89: 1111-2.
ation of diepoxybutane sensitivity and cell cycle block-
41. Rackoff WR, Orazi A, Robinson CA, et al. Prolonged administration of granulocyte colony-stimulating fac-
44. Duckworth-Rysiecki G, Comish K, Clarke CA, Buch-

Molecular basis of Fanconi anemia