

## Opitz G/BBB Syndrome in Xp22: Mutations in the *MID1* Gene Cluster in the Carboxy-Terminal Domain

Karin Gaudenz,<sup>1</sup> Erich Roessler,<sup>1,3</sup> Nandita Quaderi,<sup>4</sup> Brunella Franco,<sup>4</sup> George Feldman,<sup>1</sup> David L. Gasser,<sup>2</sup> Bärbel Wittwer,<sup>6</sup> Eugenio Montini,<sup>4</sup> John M. Opitz,<sup>7</sup> Andrea Ballabio,<sup>4,5</sup> and Maximilian Muenke<sup>1,2,3</sup>

Departments of <sup>1</sup>Pediatrics and <sup>2</sup>Genetics, The Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Philadelphia; <sup>3</sup>Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda; <sup>4</sup>Telethon Institute of Genetics and Medicine, and <sup>5</sup>Università Vita Salute, Milan; <sup>6</sup>Institut für Humangenetik, Westfälische Wilhelms-Universität, Münster, Germany; and <sup>7</sup>Departments of Pediatrics, Human Genetics, and Obstetrics and Gynecology, University of Utah, Salt Lake City

### Summary

The *MID1* gene in Xp22 codes for a novel member of proteins containing a RING finger, B-box, coiled-coil and a conserved C-terminal domain. Initially, three mutations in the C-terminal region were found in patients with Opitz G/BBB syndrome, a defect of midline development. Here we have determined the complete gene structure of the *MID1* gene and have analyzed all nine exons for mutations in a set of 40 unrelated Opitz G/BBB patients. We now report six additional mutations all clustered in the carboxy-terminal domain of the *MID1* protein. These data suggest that this conserved domain of the B-box proteins may play a fundamental role in the pathogenesis of Opitz syndrome and in morphogenetic events at the midline during blastogenesis.

### Introduction

The Opitz G/BBB syndrome (OS) is a multisystem disorder comprising primarily hypertelorism and hypospadias, but also additional anomalies—such as clefts of lip and palate, heart defects, imperforate anus, developmental delay, and tracheo-laryngeal and esophageal abnormalities, especially in males (Christian et al. 1969; Opitz et al. 1969a, 1969b; Opitz 1987). In the initial clinical reports, the G and BBB syndromes were thought to represent distinct entities until a family was described that segregated both sets of anomalies (Cordeiro and

Holmes 1978). It was appreciated at an early stage of syndrome delineation that the major organ systems affected in Opitz G/BBB syndrome involve midline structures; however, the molecular basis of Opitz syndrome was determined only recently (Quaderi et al. 1997).

Males with OS are frequently more severely affected than their sisters, whose only finding may be hypertelorism. The generally milder findings in affected females and the pedigrees of the initially reported families are consistent with X-linked inheritance. In contrast, additional reports of male-to-male transmission in OS were compatible only with autosomal dominant inheritance (Stoll et al. 1978; Farndon and Donnai 1983). Robin et al. (1995) resolved this discrepancy with evidence for genetic linkage to markers either to Xp22 or 22q11.2. This was the first instance of a classic multiple-congenital-anomalies syndrome shown to be genetically heterogeneous, with an X-linked and an autosomal form. Interestingly, the clinical spectrum of anomalies appeared to be virtually identical between the two modes of inheritance (Robin et al. 1996).

Further support for an Opitz gene on Xp22 came from detailed linkage studies (May et al. 1997) and molecular analysis of a key pericentric inversion of the X chromosome, inv(X)(p22.3q36) in a French OS family (Verloes et al. 1995). When genes in the vicinity of the Xp22.3 breakpoint were examined, a single gene, *MID1* (midline 1), was identified as the OS gene, on the basis of its location with respect to the inversion breakpoint and mutations in three of the X-linked families (Quaderi et al. 1997). Interestingly, in all three families the mutation was in the C-terminal region of the *MID1* protein; however, these studies did not analyze the entire *MID1* gene.

In order to confirm that mutations of the *MID1* gene cause the X-linked form of OS and to extend the spectrum of disease-causing mutations, we set out to examine a large cohort of OS patients. Here we report the entire gene structure of *MID1* and mutational analysis of all nine exons of the coding region in 40 unrelated patients.

Received February 17, 1998; accepted for publication July 16, 1998; electronically published August 14, 1998.

Address for correspondence and reprints: Dr. Maximilian Muenke, Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, 10 Center Drive, MSC 1852, Building 10, 10C101, Bethesda, MD 20892-1852. E-mail: mmuenke@nhgri.nih.gov

© 1998 by The American Society of Human Genetics. All rights reserved. 0002-9297/98/6303-0008\$02.00

**Table 1****Sequences of PCR Primers Used to Amplify All *MID1* Exons in OS Patients and Controls**

EXON	DOMAIN	PRIMER (5'-3')		PRODUCT (bp)	T <sub>m</sub> (°C)
		Forward	Reverse		
1	RING finger	3876, GACCTCCCTGTGCCTAATC	3856, GGCTGAGGGTGATGACATG	296	50
1	B box 1	3857, CATGTCATCACCCCTCAGCC	3797, ATGTGAGAGTCCGGAATTGG	328	58
1	B box 2	3859, CCAATTCGGACTCTCACA	3877, GTAGCTAGCAGTGAATACAC	251	56
2	Coiled-coil	3968, TATCTGAAGTATAAAAAGTATGC	3969, CAAAACCTTGATCTGGCAGT	225	54
3	Coiled-coil	3994, GCCAAATTAACACTTACCATTG	3995, AGTCCCTGGAGTGATTAGG	217	54
4	Coiled-coil	3974, AGACCACCCATGTAACCATC	3906, AATCACAGGGCAGATACAAG	251	54
5		3871, GCAGTGATGTTGAAAGTCATC	3889, AAGCAATACCTGTAAGGTAATG	197	52
6		3979, TCTGAAATGGTATCCCAGTTG	3978, ATCAATCTGAAGAAAGCCACC	261	54
7		3835, CTTGTGCCAAAGAAGTGCAC	3870, GAGCAGATAAGACATGACAG	246	56
8	C-terminal	3869, TTGAAAATATGCTATTGGTTATG	3825, AGTACAGAATGAGATGTGCC	245	55
9	C-terminal	3868, ATGTGCAACATGGCTCATTG	3769, GACACTTGTCCACACGGTG	334	56
9		3830, CACCGTGTGGAACAAGTGTC	3892, GTAATCCTGATTATGTTTCTC	348	57

## Patients and Methods

### Patients

Twenty-two families with OS, as well as 18 individuals with sporadic OS (and, when available, their parents), participated in this study. Three of these OS families with mutations in the *MID1* gene have been described elsewhere (Quaderi et al. 1997). Here we have identified *MID1* mutations in five additional OS families (8/22 [36%]) and in one apparently sporadic OS patient (1/18 [6%]). Detailed clinical findings, including photographs, were reported for kindred OS3 (Brooks et al. 1992). All of the OS patients either were personally evaluated or had photographs and clinical records reviewed by one of the authors (J.M.O. or M.M.). Kindred OS14 consisted of two affected brothers, their affected mother, and an affected maternal aunt. The index case had hypertelorism with widow's peak, hypospadias and unilateral cryptorchidism, and an imperforate anus. Only the proband, his mother, and his unaffected father were available for molecular analysis. In kindred OS19 two brothers were affected, one with hypertelorism, severe hypospadias, and imperforate anus and the other with hypertelorism, unilateral cleft lip and palate, and hypospadias. Their sister and mother both had hypertelorism. DNA samples from all four of these individuals and from the unaffected father were available for molecular studies. Kindred OS28 consisted of a large family with numerous affected individuals. The proband had hypertelorism, cleft palate, esophageal-tracheal fistula, and ambiguous genitalia. For the molecular analysis, samples from two females with hypertelorism were available. In kindred OS36, the index case had hypertelorism with broad nasal bridge and widow's peak, a unilateral cleft lip, an esophageal cleft with dysphagia, a ventricular-septal defect, umbilical hernia, and hypospadias.

His mother had mild hypertelorism and a widow's peak. One apparently sporadic OS patient had hypertelorism and hypospadias. He was considered to be a sporadic case, since his parents, who were not available for molecular analysis, appeared to be clinically normal. Informed consent was obtained from all individuals, according to the standards set by local review boards.

### DNA Analysis

Genomic DNA from affected and unaffected members of OS families, apparently sporadic OS individuals, and their parents was extracted from lymphocytes or established lymphoblastoid cell lines by routine methods. DNA samples were amplified by PCR using 12 different *MID1*-specific primer pairs (table 1). Each PCR reaction (final volume 15  $\mu$ l) contained 60 ng DNA; 200  $\mu$ M each of dATP, dGTP, and dTTP; 62.5  $\mu$ M dCTP; 5  $\mu$ Ci  $\alpha$ [<sup>32</sup>P]-dCTP (800 Ci/mmol [10  $\mu$ Ci/ $\mu$ l]); 25 pmol each primer; 1% formamide; 10 mM Tris-HCl pH 8.3; 50 mM KCl; 1.5 mM MgCl<sub>2</sub>; 0.0001% (w/v) gelatin; and 1 U *AmpliTaq* polymerase (Perkin-Elmer). Reactions were performed in a Hybaid Omnigene Temperature Cycler, with the annealing temperatures (T<sub>m</sub>) specified in table 1. The program consisted of an initial denaturing step (95°C for 5 min) followed by 35 PCR cycles (95°C for 45 s, T<sub>m</sub> for 1 min, and 72°C for 30 s) and a terminal extension step (72°C for 10 min).

For SSCP analysis, the  $\alpha$ [<sup>32</sup>P]-dCTP-labeled PCR products were electrophoresed under two different conditions in a 0.5% MDE gel (AT Biochem) with or without 5% glycerol, at room temperature. Amplicons demonstrating SSCP bandshifts were analyzed by the Nucleic Acid/Protein Core Facility of the Children's Hospital of Philadelphia by use of an ABI 377A automated sequencer. DNA samples showing SSCP alterations were compared with those of other members from the same

family to assess allele segregation and with >100 X chromosomes from unrelated unaffected control individuals.

## Results

The *MID1* gene was localized to a mega-YAC in Xp22 that spans the short arm breakpoint of the inversion, inv(X)(p22.3q36), which segregated with OS in a 20-member multigenerational French family (Verloes et al. 1995; Quaderi et al. 1997). Complete sequencing of the intron-exon boundaries by use of cosmid templates demonstrated that the *MID1* coding region consists of nine exons distributed over a 300-kb interval. The genomic structure of the *MID1* gene is represented in figure 1A. The intron-exon boundaries are detailed in table 2.

For mutational study using SSCP analysis, we selected 18 cases of apparently sporadic and 22 familial OS cases without evidence of male-to-male transmission that we considered as compatible with X linkage. The small size of many of the OS families precluded linkage studies. The SSCP mutational screen employed 12 pairs of primers that encompassed all of the intron-exon boundaries and coding regions of the gene, as shown in table 1.

Our results show the conserved C-terminal domain to be the most common site involved in structural mutations in OS (fig. 1B). Among this panel of 40 OS patients a total of 12 sequence variations were detected as band shifts and were confirmed by direct sequencing of the amplicons. One of these band shifts is caused by a polymorphism in the intron immediately preceding exon 3 (table 2). Two of these band shifts are predicted to be polymorphisms (C1452T and A1884G), since they do not alter the amino acid at these positions, whereas the remaining nine sequence variations predict structural changes in the *MID1* protein, with a clustering of effects in the C-terminal domain of the protein (fig. 1B). None of these mutations were detected in >100 normal X chromosomes.

### *Predicted Polymorphisms*

The first of the two predicted polymorphisms occurs at nucleotide C1452T and predicts no change in the amino acid Thr422 in exon 6. This base-pair change was identified in a female with OS. All three of her affected children (two sons and one daughter) do not carry the C1452T change, suggesting that it is indeed a polymorphism (Cotton and Scriver 1998). Similarly, the A1884G sequence variation does not predict an alternative codon for Glu566 in exon 9 and, therefore, predicts no change in the primary structure of the *MID1* protein. The A1884G change was identified in a mother and her son with OS. Both of them had been shown to carry a 24-bp insertion in exon 8 of the *MID1* gene (the

family of these individuals have been designated as "OS20" by Quaderi et al. [1997]), suggesting that the sequence variation (A1884G) and the insertion are in the same *MID1* allele. Both of these polymorphisms are rare, since neither was seen in the normal controls.

A third polymorphism was detected 15 bases 5' of the intron/exon boundary of the start of exon 3, as a substitution of a T for the normal G. This sequence variant was observed only in the normal control samples and not in OS individuals and does not occur in the sequence context of a donor or acceptor splice site. A total of five normal individuals were detected, by SSCP analysis, as being either homozygous or heterozygous for this sequence variation, as confirmed by DNA sequencing.

### *Missense Mutations*

Two disease-associated missense mutations were observed in the familial group. The first of these occurred within the coiled-coil domain at nucleotide position T982C and predicts a substitution of an Arg residue for the normal Cys at position 266 of the amino acid sequence. This sequence change is presumed to cause Opitz syndrome in OS3, since it can be seen to segregate with the disease, from an affected mother to her affected son (fig. 2a). The second missense mutation occurred within the C-terminal domain at position T1793C and predicts a substitution of Thr for the normal Ile at amino acid 536. This amino acid position is adjacent to a highly conserved domain in the C-terminal region; however, this particular residue itself is not conserved among members of the B-box family. This sequence change also segregates with the phenotype of Opitz syndrome in OS36 (fig. 2d).

### *Deletions*

The deletion of an ATG codon at nucleotide 1498 in OS5 has been described elsewhere (Quaderi et al. 1997). A second deletion of a single G base pair at position 1134 in OS14 is predicted to result in a frameshift and termination of the protein, within three subsequent codons (figs. 1B and 2b). The resulting protein is predicted to be missing the end of the coiled-coil domain and the entire C-terminal domain.

### *Splice-Acceptor Mutation*

A single putative splicing mutation was detected in the 3' splice-acceptor sequence in OS19 (fig. 2c) and predicts a substitution of a GG for the invariant AG at the end of intron 6. This sequence is shown also in table 2, which tabulates the sequences of the intron-exon boundaries of the *MID1* gene. Although reverse transcription-PCR experiments have not yet defined the molecular consequences for this sequence change (since a

A

1 atttcgccgggttgcctttgtcttgcgggctcctgtcgggttcgggtttccgctctgaagaactgogacgogggctccgatgcagctcgcctccctgcggg 100  
> exon 1  
101 atgggtcatgggattctaaacatgaggcagatagctgatcagcttccctgggttttgcctgatgacacaagagagctttgcctgaag ATG GAA ACA 195  
1 M E T 3

196 CTG GAG TCA GAA CTG ACC TGC CCT ATT TGT CTG GAG CTC TTT GAG GAC CCT CTT CTA CTG CCC TGC GCA CAC AGC 270  
4 L E S E L T T C C P I C L F E D P L L P C H S 28

271 CTC TGC TTC AAC TGC GCC CAC CGC ATC CTA GTA TCA CAC TGT GCC ACC AAC GAG TCT GTG GAG TCC ATC ACC GCC 345  
29 L C F N C A H R I L V S H C A T N E S V E S I T A 53

346 TTC CAG TGC CCC ACC TGC CGG CAT GTC ATC ACC CTC AGC CAG CGA GGT CTA GAC GGG CTC AAG CGC AAC GTC ACC 420  
54 F Q C P T C R H V I T L S Q R G L D G L K R N V T 8

421 CTA CAG AAC ATC ATC GAC AGG TTC CAG AAA GCA TCA GTG AGC GGG CCC AAC TCT CCC AGC GAG ACC CGT CGG GAG 495  
79 L Q N I I D R F Q K A S V S G P N S P S E T R R E 103

496 CGG GCC TTT GAC GCC AAC ACC ATG ACC TCC GCC GAG AAG GTC CTC TGC CAG TTT TGT GAC CAG GAT CCT GCC CAG 570  
104 R A F D A N T M T S A E K V L Q F C D Q D P A Q 128

571 GAC GCT GTG AAG ACC TGT GTC ACT TGT GAA GTA TCC TAC TGT D GAG TGC CTG AAA GCC ACT CAC CCG AAT AAG 645  
129 D A V K T C V T C E V S Y C D G C L K A T H L V N K 153

646 AAG CCC TTT ACA GGC CAT CGT CTG ATT GAG CCA ATT CCG GAC TCT CAC ATC CGG GGG CTG ATG TGC TTG GAG CAT 720  
154 K P F T G H R L I E P I P D S H I R G L M C L E H 178

721 GAG GAT GAG AAG GTG AAT ATG TAC TGT GTG ACC GAT GAC CAG TTA ATC TGT GCC TTG TGT AAA CTG GTT GGG CGG 795  
179 E D E K V N M Y C V T D T D G C L I C A L C K L V G R 203

796 CAC CGC GAT CAT CAG GTG GCA GCT TTG AGT GAG CGC TAT GAC AAA TTG AAG CAA AAC TTA GAG AGT AAC CTC ACC 870  
204 H R D H Q V A A L S E R Y D K L K Q N L E S N L T 228  
> exon 2

871 AAC CTT ATT AAG AGG AAC ACA GAA CTG GAG ACC CTT TTG GCT AAA CTC ATC CAA ACC TGT CAA CAT GTT GAA GTC 945  
229 N L I K R N A N T M T S A E K V L Q T C Q H V E V 253

946 AAT GCA TCA CGT CAA GAA GCC AAA TTG ACA GAG GAG TGT GAT CTT CTC ATT GAG ATC ATT CAG CAA AGA CGA CAG 1020  
254 N A S R Q A A K L T E E C/R L L I E I I Q Q R R Q 278

1021 ATT ATT GGA ACC AAG ATC AAA GAA GGG AAG GTG ATG AGG CTT CGC AAA CTG GCT CAG CAG ATT GCA AAC TGC AAA 1095  
279 I I G T K I K E G K V M R L R K A L A Q Q I A N C K 303

1096 CAG TGC ATT GAG CGG TCA GCA TCA CTC ATC TCC CAA GCG GAA CAC TCT CTG AAG GAG AAT GAT CAT CGC CGT TTC 1170  
304 Q C I E R S L I S Q A E H S L K E N D H A R F 328

1171 CTA CAG ACT GCT AAG AAT ATC ACC GAG AGA GTC TCC ATG GCA ACT GCA TCC TCC CAG GTT CTA ATT CCT GAA ATC 1245  
329 L Q T A K N I T E R V S M A T A S S Q V L I P E I 353

1246 AAC CTC AAT GAC ACA TTT GAC ACC TTT GCC TTA GAT TTT TCC CGA GAG AAG AAA CTG CTA GAA TGT CTG GAT TAC 1320  
354 N L N D T N F M V K A I N Q A G S R S S E P G K 478

1321 CTT ACA GCT CCC AAC CCT CCC ACA ATT AGA GAA GAG CTC TGC ACA GCT TCA TAT GAC ACC ATC ACT GTG CAT TGG 1395  
379 L T A P N P F P T I R E E L C T A S Y D T I T V H W 403

1396 ACC TCC GAT GAT GAG TTC AGC GTG GTC TCC TAC GAG CTC CAG TAC ACC ATA TTC ACC ACT GGA CAA GCC AAC GTC GTT 1470  
404 T S D D E F T N V V S Y E L Q Y T I F T G Q A N V 428

1471 AGT CTG TGT AAT TCG GCT GAT AGC TGG ATG ATA GTA CCC AAC ATC AAG CAG AAC CAC TAC ACG GTG CAG GGT CTG 1545  
429 S L C N S V M I V P N I K Q N H Y T V H G L 453

1546 CAG AGC GGC ACC AAG TAC ATC TTC ATG GTC AAG GCC ATC AAC CAG GCG GGC AGC CGC AGC AGT GAG CCT GCG AAG 1620  
454 Q S G T K Y I F M V K A I N Q A G S R S S E P G K 478

1621 TTG AAG ACA AAC AGC CAA CCA TTT AAA CTG GAT CCC AAA TCT GCT CAT CGA AAA CTG AAG GTG TCC CAT GAT AAC 1695  
479 L K T N S V F I/T D S G R H Y W I E P A P H L K V S H D N 503

1696 TTG ACA GTA GAA CGT GAT GAG TCA TCA TCC AAG AAG AGT CAC ACA CCT GAA CGC TTC ACC AGC CAG GGG AGC TAT 1770  
504 L T V E R D E S S S K K S H T P E R F T S Q G S Y 528

1771 GGA GTA GCT GGA AAT GTG TTT ATT GAT AGT GGC CGG CAT TAT TGG GAA GTG GTC ATA AGT GGA AGC ACA TGG TAT 1845  
529 G V A G N V F I/T D S G R H Y W I E P A P H L K V S H D N 503

1846 GCC ATT GGT CTT GCT TAC AAA TCA GCC CCG AAG CAT GAA TGG ATT GGG AAG AAC TCT GCT TCC TGG GCG CTC TGC 1920  
554 A I G L A Y K S A P K H E W I G K N S A S W A L C 578

1921 CGC TGC AAC AAT AAC TGG GTG GTG AGA CAC AAT AGC AAG GAA ATC CCC ATT GAG CCT GCC CCC CAC CTC CGG CGC 1995  
579 R C N N N N W V R H N S K E I P I E P A P H L R R 603

1996 GTG GGC ATC CTG CTG GAC TAT GAT AAC GGC TCT ATC GCC TTT TAT GAT GCT TTG AAC TCC ATC CAC CTC TAC ACC 2070  
604 V G I L L D Y D N G S I A F Y D A L N S I H L Y T 628

2071 TTC GAC GTC GCA TTT GCG CAG CCT GTT TGC CCC ACC TTC ACC GTG TGG AAC AAG TGT CTG ACG ATT ATC ACT GGG 2145  
629 F D V A F A Q P V C P T F T V W N K C L T I I T G 653

2146 CTC CCT ATC CCA GAC CAT TTG GAC TGC ACA GAG CAG CTG CCG TGA gcgctcggccacatggagctgcttctggggaacagtaag 2230  
654 L P I P D H L D C T E Q L P \* 668

2231 gttcaggccactattttaggggactgagaagcagaggcttcatgagtgtaagaatctcaccagaagtgtcccgaatcggctcagatagggtcaaaa 2330

2331 caagagattcctcctcttcttactgtgtcttattagtaacgggcttataaattcttcttaattttttgtatttagagaaaaatctatagattattat 2430

2431 aagagaaaacataatcaggattacaacttttaggaattacttggttttgcacattaaagaggccataaagttatcagctattacaacctcatttcoaca 2530

2531 caatctgtgggcttacaaaaaaacaaactttttagtatttggatgttactcatctcttaacctgataccccatgatgatccatggtaggtcttctca 2630

2631 cctcgatggggoataaacaggatgtgtttgaacctagtggggaggaaacaggcttcttactctgggttaatttgaagtgttttaattgtgatgtcaaaa 2730

2731 agttgatcagatcaactaaaatggagagcaagacagagaatgaaaagagtgattttggacctcggacctggcgtggcgttaaatctttacctctcata 2830

2831 gctgatgggataatgttggaaagaagggttgaatcctttggccacattttgcoctgctctcagggttaagggttctggaagaacattaagaatga 2930

2931 gatgcaattgaaaatagtcattttgaatcctattgattattcaaaaatcaggctgattgtctttatcagaggtaggattctgtttatagatagaat 3030

3031 ctactttatcctctccttttaagtcttcttagacctgtaaaattctctcaactatttaagtctcctatttcccgctccccatataatcttcc 3130

3131 tttgtctccgggctgagtaaaaaactgttctgtcacaaatagcagcaccactttggattgattttgctctccaggacatcagcacatggccctgat 3230

3231 cagcactaccacatccaaacataagtcactgaaaaaacacttaatttttagtggtaatgacaaggacattgtataaagtaactatttctagattca 3330

3331 tgctcaaaagtattataacagacctttataaacacatcttgaagagtgagaagtcctctctatagctatagtttacaatagagttgtaagac 3430

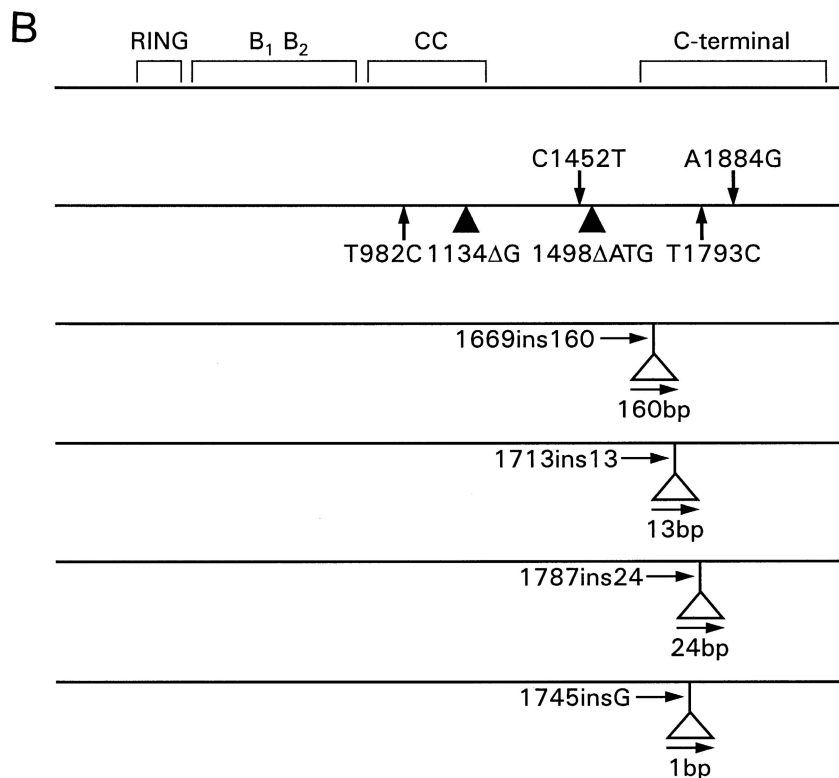
3431 caaaaaaaaaaaaaaaaaaaaaa

RING FINGER

B BOXES

COILED-COIL DOMAIN

C-TERMINAL DOMAIN



**Figure 1** Spectrum of *MID1* mutations in familial and sporadic Opitz G/BBB syndrome (OS). A, Nucleotide and predicted amino acid sequence of the *MID1* cDNA (modified from Quaderi et al. 1997). Conserved Cys and His residues in the RING and B-box domains are circled. Missense mutations are indicated by the nucleotide changes (*upper row in each set of paired rows*) above and resulting amino acid changes (*lower row in each set of paired rows*), below the wild-type sequence. Polymorphisms are indicated by nucleotide changes (*in boldface*) above the sequence: C1452T (Thr422) and A1884G (Glu566). Missense mutations are indicated in an identical fashion: T982C leads to a Cys266Arg change (in OS3; see fig. 2a), and T1793C predicts a Ile536Thr change (in OS36; see fig. 2d). The splice-site mutation and the intron 2 sequence variant are described in table 2. Deletions are indicated by an inverted “V” (^) above the sequence: a 1-bp deletion, 1134delG, leads to a frameshift and premature stop codon at position 1156 (in OS14; see fig. 2b), and an in-frame 3-bp deletion, 1498ΔATG, predicting a single amino acid deletion, ΔM438, in OS5, which has been described elsewhere (Quaderi et al. 1997). A single nucleotide insertion is indicated by a “V” above the sequence: 1745insG, in OS16 (Quaderi et al. 1997). The initial positions of the tandemly duplicated sequences are identified by an asterisk: \*1 = 1669ins160 (in OS28), \*2 = 1713ins13 (in sporadic case OS; see fig. 2e), and \*3 = 1787ins24 (in OS20 [Quaderi et al. 1997]). B, Diagrammatic representation of nucleotide sequence of the various altered *MID1* genes in OS, with respect to the functional domains: the RING finger (Freemont et al. 1991), the B-box (Reddy et al. 1991), the coiled-coil (Reddy et al. 1992), and the C-terminal domain (Quaderi et al. 1997). The positions of codon variants are indicated by the arrows. Blackened triangles denote the position of the microdeletions, the positions of the direct repeats are indicated by the horizontal arrows, and the sizes of the repeats are given below the unblackened triangles.

second blood sample was not available), it is tempting to speculate that this mutation results in either an unstable transcript or a significant alteration in the protein structure encoded by exons 7–9, because of use of an alternative upstream or downstream splice-acceptor site.

#### *Head-to-Tail Direct Repeats in the C-Terminal Domain as the Cause of Frameshift Truncations*

No fewer than four direct repeats of the C-terminal coding region were observed in familial OS, with the first two (1745insG, in OS16; and 1787ins24, in OS20) having been described elsewhere (Quaderi et al. 1997). The 1745insG results in a frameshift with premature

termination within 16 codons, whereas the 1787ins24 causes an in-frame addition of eight amino acids, FIDSGRHL. Here we report two additional frameshift mutations resulting from direct repeats in the coding sequence. The first of these occurred in OS28 and resulted from the tandem duplication of 160 bp at position 1669. It is not until six codons into the second repeat that a termination codon is encountered, suggesting that the protein is intact for the first 528 amino acids and then terminates prematurely in the C-terminal domain. As shown in figure 2e, one of the apparently sporadic OS patients (whose parents were unavailable) has a direct repeat, of 13 bp at position 1713, that predicts

**Table 2****MID1 Splice Junctions**

EXON	SPICE SITE <sup>a</sup>		EXON SIZE (bp)
	3'	5'	
1	tcttgtttccag/ATAGCTGATCAG	GACAAATTGAAG/gttagtccgatc	716
2	tttttgaacag/CAAACTTAGAG	CAACATGTTGAA/gtgagtatctct	96
3	cttatcctgtag/GTCAATGCATCA	AAAGAAGGGGAAAG/gtatgatttcta	108
4	cttttgattaag/GTGATGAGGCTT	TATCACCCGAGAG/gtgagtgcaggg	149
5	tttgattacag/AGTCTCCATGGC	ATTACCTTACAG/gtattgtctttt	128
6	tgagccccacag/CTCCCAACCCTC	CCAACGTCGTTA/gtgagtatctca	144
7	ttggctttgcag/GTCTGTGTAATT	TGAAGACAAACA/gtaagtgtctgtg	162
8	attgttttcag/GCCAACCATTTA	TGGAAGCACATG/gtaagcaactta	209
9	ttgttccgcag/GTATGCCATTGG	...	...

<sup>a</sup> Exon sequences are represented by uppercase letters, and intron sequences are represented by lowercase letters. The only substantial difference between our results and those of Perry et al. (1998) involve the beginning of exon 1, which is attributed to alternative splicing of the 5' UTR; only minor sequence differences within a few of the introns are observed. In OS19, a putative 3' splice acceptor-site mutation (underlined in the sequence that follows) was identified: ttggctttgcag/GTCTGTGTAATT (see fig. 2c). An A at position -2 is invariant in all higher eukaryotes. At position -15 upstream of exon 3, a sequence variant—a T substituted for the normal G (underlined in the sequence that follows) within the intron 2 sequence, ccatt(g/t)tactatcctgtag/GTCAATGCATCA—was observed in five normal control individuals.

termination of the protein as soon as the translational machinery encounters the TGA at the beginning of the second repeat.

## Discussion

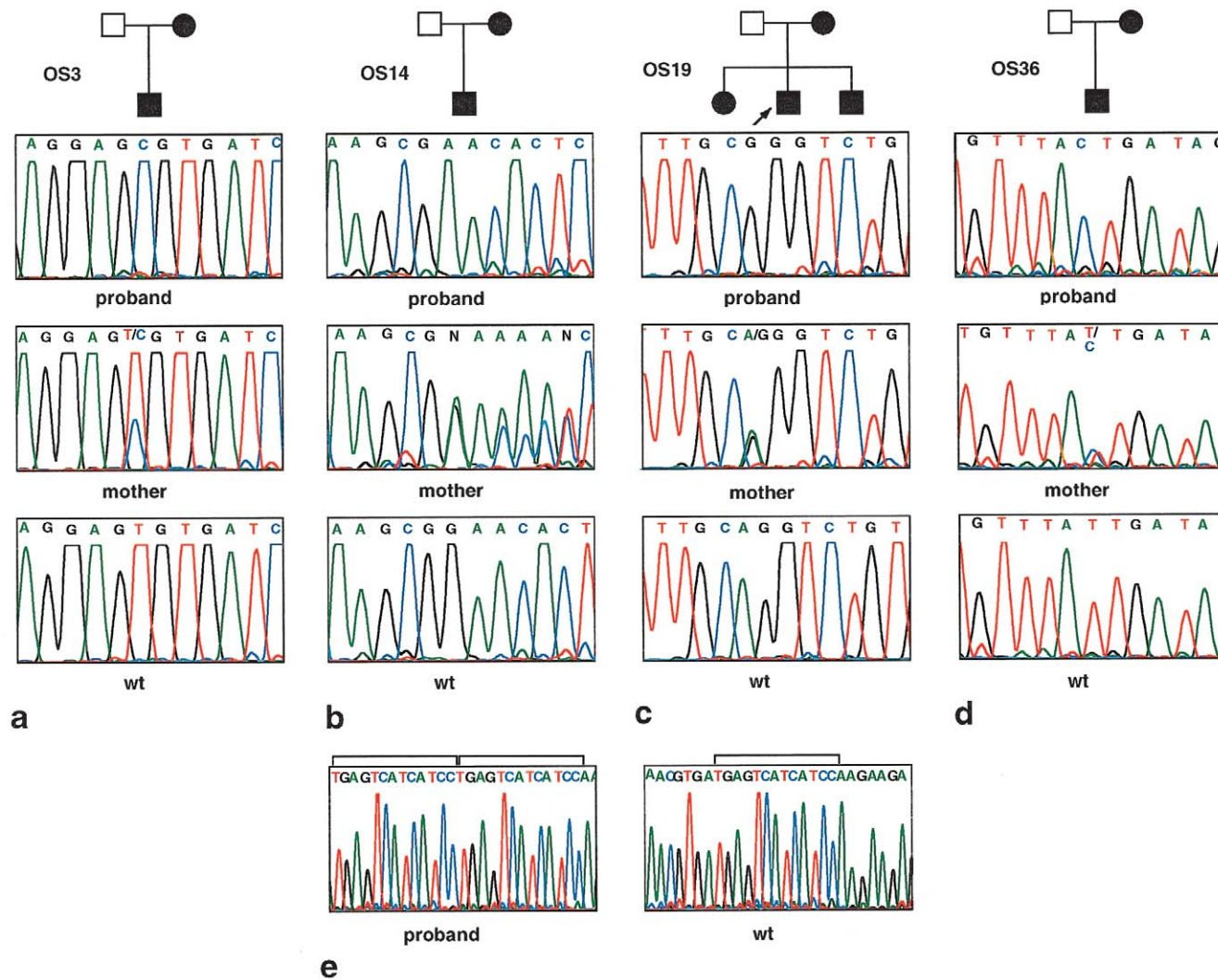
Elsewhere, we have described the identification of the *MID1* gene (Quaderi et al. 1997), the latest member of a family of proteins with a characteristic RING-finger motif that consists of a cluster of conserved Cys/His residues that ligate to zinc atoms in a novel tetrahedral "cross-brace," clearly distinct from classical zinc-finger motifs (Borden and Freemont 1996). A gene identical to *MID1* has been identified independently by another group and has been named "FXV" (Perry et al. 1998). The two gene structures are identical with respect to the coding region and differ only in the alternative splicing of the 5' UTR (table 2).

This RING-finger family includes >80 members of diverse cellular function, including putative transcriptional activators, developmental genes, protooncogenes, and viral replication genes (Freemont et al. 1991; Freemont 1993; Saurin et al. 1996). A subset of this large gene family contains a second zinc-binding motif called the "B box" (Reddy and Etkin 1991; Borden et al. 1996). The B-box motif consists of a second zinc-coordinated protein fold with a configuration of Cys/His clusters distinct from the RING domain and from typical zinc fingers. Some members of the B-box family have one or two zinc-binding domains. Most members of the B-box family, including *MID1*, contain a third conserved domain, which is rich in hydrophobic amino acids, in a tandem series of  $\alpha$ -helices that form the coiled-coil do-

main (Reddy et al. 1992). The function of the coiled-coil domain is thought to include an important protein-protein interaction surface. Proteins with the complete tripartite structure include *MID1* (Quaderi et al. 1997), *PML*, *rfp*, *T18*, and *rpt-1*, which regulates the interleukin-2 receptor  $\alpha$ -chain (Patarca et al. 1988). Additional examples include the lupus autoantigen Ro/SS-A (Chan et al. 1991), *efp* (an estrogen responsive gene) (Inoue et al. 1993), *xnf7* (*Xenopus* nuclear factor 7) (Reddy et al. 1991; Borden et al. 1993), and a nuclear protein *PwA33*, which binds to amphibian lampbrush chromosomal loops (Bellini et al. 1993). Many, but not all, of these gene products are nuclear in intracellular location and are found in very large macromolecular assemblages,  $\geq 600$  kD (Borden and Freemont 1996). Presumably, the protein-interaction domains inherent in the tripartite structure influence these protein aggregations. At present, the composition of these macromolecular complexes is unknown.

Mutations in *MID1* account for 8/22 (36%) of the OS familial cases. The fact that only 1/18 (6%) of the apparently sporadic cases could be attributed to a *MID1* mutation could be due to phenocopies, locus heterogeneity (e.g., the autosomal dominant locus for OS on chromosome 22q11.2) or the insensitivity of our screening method (SSCP analysis). Direct DNA sequencing of these patients has not yet been performed.

None of the mutations in the *MID1* gene associated with OS in this study alter the structure of either the RING or B-box domains. A missense mutation T982C was detected in the coiled-coil domain and could potentially affect the function of this domain. The 1134 $\Delta$ G



**Figure 2** Representative chromatograms of *MID1* DNA-sequence mutations from individuals with familial and sporadic Opitz G/BBB syndrome. The pedigree structure of four X-linked OS families is shown, with unaffected (*unblackened symbols*) and affected individuals (*blackened symbols*). DNA-sequence chromatograms of the affected index case, the affected mother, and a female control homozygous for the wild-type sequence (wt) are shown below each pedigree (*a-d*). Mothers are heterozygous for the wild-type and mutated alleles. *a*, T982C, leading to Cys266Arg. *b*, 1134delG. Note the presence of double peaks in the mother, 3' of the deletion, due to the single-base phase shift of the sequences of the normal and mutant alleles. *c*, A(-2)G splice-site mutation. In the pedigree diagram, the proband is indicated by an arrow. *d*, T1793C, leading to Ile536Thr. *e*, Patient with sporadic OS, with 1713ins13 tandem duplication (indicated by horizontal square brackets). The parents of this individual were unavailable for study.

mutation is predicted to truncate the protein within the coiled-coil domain and would therefore also be missing the entire C-terminal domain. Similarly, the splice acceptor mutation A(-2)G in intron 6 would be expected to alter the structure of the C-terminal domain through alternative splicing. It is this C-terminal domain that is the most common region affected by mutations in our OS families. This novel domain has a high degree of homology among all of the members of the RING/B-box family. The present study supports the critical role of this C-terminal region in the pathogenesis of the OS

malformations. The function of this domain is presently unknown.

Interestingly, all of the mutations found to date in OS patients spare the N-terminal region of the protein, which contains the zinc-finger and B-box domains. This suggests that a complete loss of function of the *MID1* protein may be lethal, or may result in a phenotype different from OS. However, if this were true, we should hypothesize that the X-chromosome inversion disrupting the 5' UTR of *MID1* in OS patients does not abolish *MID1* function completely. The analysis of *MID1* pro-

motor(s) and alternative splicing of the 5' UTR, as well as the creation of transgenic mice by homologous recombination—carrying MID1 mutations, will address these issues.

## Acknowledgments

We are deeply grateful to all OS families for their participation in this study. We thank Jeffrey E. Ming and Michelle Kaupas for their critical help with this manuscript. This work is supported by the Italian Telethon Foundation; by European Community, under grants BMH4-CT96-1134 and BMH4-CT96-0889 (both to A.B.); and both by National Institutes of Health (NIH) grants HD28732 and HD29862 (both to M.M.) and by the Division of Intramural Research, National Institute of Human Genome Research, NIH (support to M.M.).

## References

- Bellini M, Lacroix J-C, Gall JG (1993) A putative zinc-binding protein on lampbrush chromosome loops. *EMBO J* 12: 107–114
- Borden KLB, Freemont PS (1996) The RING finger domain: a recent example of a sequence-structure family. *Curr Opin Struct Biol* 6:395–401
- Borden KLB, Lally JM, Martin SR, O'Reilly NJ, Solomon E, Freemont PS (1996) In vivo and in vitro characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia protooncoprotein PML. *Proc Natl Acad Sci USA* 93:1601–1606
- Borden KLB, Martin SR, O'Reilly NJ, Lally JM, Reddy BA, Etkin LD, Freemont PS (1993) Characterization of a novel cysteine/histidine-rich metal binding domain from *Xenopus* nuclear factor XNF7. *FEBS Lett* 335:255–260
- Brooks JK, Leonard CO, Coccaro PJ Jr (1992) Opitz (BBB/G) syndrome: oral manifestations. *Am J Med Genet* 43: 595–601
- Chan EKL, Harnel JC, Buyon JP, Tan EM (1991) Molecular definition and sequence motifs of the 52 kD component of human SS-A/Ro autoantigen. *J Clin Invest* 87:68–76
- Christian JC, Bixler D, Blythe SC, Merritt AD (1969) Familial telecanthus with associated congenital anomalies. *Birth Defects* 5(2): 82–85
- Cordeiro JF, Holmes LB (1978) Phenotypic overlap of the BBB and G syndromes. *Am J Med Genet* 2:145–152
- Cotton RGH, Scriver CR (1998) Proof of “disease causing” mutation. *Hum Mutat* 12:1–3
- Farndon PA, Donnai D (1983) Male to male transmission of the G syndrome. *Clin Genet* 24:446–448
- Freemont PS (1993) The RING finger: a novel protein sequence motif related to the zinc finger. *Ann N Y Acad Sci* 684: 174–192
- Freemont PS, Hanson IM, Trowsdale J (1991) A novel cysteine-rich sequence motif. *Cell* 64:483–484
- Inoue S, Orimo A, Hosoi T, Kondo S, Toyoshima H, Kondo T, Ikegami A, et al (1993) Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proc Natl Acad Sci USA* 90:11117–11121
- May M, Huston S, Wilroy RS, Schwartz C (1997) Linkage analysis in a family with Opitz G/BBB syndrome refines the location of the gene on Xp22 to a 4 cM region. *Am J Med Genet* 68:244–248
- Opitz JM (1987) G syndrome (hypertelorism and esophageal abnormality, and hypospadias, or hypospadias-dysphagia, or “Opitz-Frias” or “Opitz G” syndrome) perspective and bibliography in 1987. *Am J Med Genet* 28:275–285
- Opitz JM, Frias II, Gutenberger JE, Pellet JR (1969a) The G-syndrome of multiple congenital anomalies. *Birth Defects* 5(2): 95–102
- Opitz JM, Summitt RL, Smith DW (1969b) The BBB syndrome: familial telecanthus with associated congenital anomalies. *Birth Defects* 5 (2): 86–94
- Patarca R, Schwartz J, Singh RP, Kong Q-T, Murphy E, Anderson Y, Sheng F-YW, et al (1988) rpt-1, an intracellular protein from helper/inducer T cells that regulates gene expression of interleukin 2 receptor and human immunodeficiency virus type 1. *Proc Natl Acad Sci USA* 85:2733–2737
- Perry J, Feather S, Smith A, Palmer S, Ashworth A (1998) The human FXI gene is located within Xp22.3: implications for evolution of the mammalian X chromosome. *Hum Mol Genet* 7:299–305
- Quaderi NA, Schweiger S, Gaudenz K, Franco B, Rugarli EI, Berger W, Feldman GJ, et al (1997) Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. *Nat Genet* 17:285–291
- Reddy BA, Etkin LD (1991) A unique bipartite cysteine-histidine motif defines a sub family of potential zinc-finger proteins. *Nucleic Acid Res* 19:6330
- Reddy BA, Etkin LD, Freemont PS (1992) A novel zinc-finger coiled-coil domain in a family of nuclear proteins. *Trends Biochem Sci* 17:344–345
- Reddy BA, Kloc M, Etkin LD (1991) The cloning and characterization of a maternally expressed novel zinc finger nuclear phosphoprotein (xnf7) in *Xenopus laevis*. *Dev Biol* 148:107–116
- Robin NH, Feldman GJ, Aronson AL, Mitchell HF, Weksberg R, Leonard CO, Burton BK, et al (1995) Opitz syndrome is genetically heterogeneous with one locus on Xp22 and a second locus on 22q11.2. *Nat Genet* 11:459–461
- Robin NH, Opitz JM, Muenke M (1996) Opitz G/BBB syndrome: clinical comparison of families linked to Xp22 and 22q and a review of the literature. *Am J Med Genet* 62: 305–317
- Saurin AJ, Borden KL, Boody MN, Freemont PS (1996) Does this have a familiar RING? *Trends Biochem Sci* 21:208–214
- Stoll C, Geradel A, Berland H, Roth MP, Dott B (1978) Male to male transmission of the hypertelorism/hypospadias (BBB) syndrome. *Am J Med Genet* 20:221–225
- Verloes A, David A, Odent S, Toutain A, André M-J, Lucas J, Le Marec B (1995) Opitz G/BBB syndrome: chromosomal evidence of an X-linked form. *Am J Med Genet* 59:123–128