

ARTICLE

Differences in methylation patterns in the methylation boundary region of IDS gene in hunter syndrome patients: implications for CpG hot spot mutations

Shunji Tomatsu^{*1}, Kazuko Sukegawa², Georgeta G Trandafirescu¹, Monica A Gutierrez¹, Tatsuo Nishioka¹, Seiji Yamaguchi³, Tadao Orii², Roseline Froissart⁴, Irene Maire⁴, Amparo Chabas⁵, Alan Cooper⁶, Paola Di Natale⁷, Andreas Gal⁸, Akihiko Noguchi¹ and William S Sly⁹

¹Department of Pediatrics, Saint Louis University, Pediatric Research Institute, St Louis, MO, USA; ²Department of Pediatrics, Gifu University School of Medicine, Gifu, Japan; ³Department of Pediatrics, Shimane University, Shimane, Japan; ⁴Centre d'Etude des Maladies Métaboliques, Hôpital Debrousse, Lyon France; ⁵Institut de Bioquímica Clínica, Barcelona, Spain; ⁶Willink Biochemical Genetics Unit, Royal Manchester Children's Hospital, Great Britain, UK; ⁷Department of Biochemistry and Medical Biotechnologies, University of Naples, Federico II, Italy; ⁸Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany; ⁹Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St Louis, MO, USA

Hunter syndrome, an X-linked disorder, results from deficiency of iduronate-2-sulfatase (IDS). Around 40% of independent point mutations at *IDS* were found at CpG sites as transitional events. The 15 CpG sites in the coding sequences of exons 1 and 2, which are normally hypomethylated, account for very few of transitional mutations. By contrast, the CpG sites in the coding sequences of exon 3, though also normally hypomethylated, account for much higher fraction of transitional mutations. To better understand relationship between methylation status and CpG transitional mutations in this region, the methylation patterns of 11 Hunter patients with transitional mutations at CpG sites were investigated using bisulfite genomic sequencing. The patient cohort mutation spectrum is composed of one mutation in exon 1 (one patient) and three different mutations in exon 3 (10 patients). We confirmed that in normal males, cytosines at the CpG sites from the promoter region to a portion of intron 3 were hypomethylated. However, specific CpG sites in this area were more highly methylated in patients. The patients with p.R8X (exon 1), p.P86L (exon 3), and p.R88H (exon 3) mutations had a hypermethylated condition in exon 2 to intron 3 but retained hypomethylation in exon 1. The same trend was found in four patients with p.A85T (exon 3), although the degree of hypermethylation was less. These findings suggest methylation patterns in the beginning of *IDS* genomic region are polymorphic in humans and that hypermethylation in this region in some individuals predisposes them to CpG mutations resulting in Hunter syndrome.

European Journal of Human Genetics (2006) 14, 838–845. doi:10.1038/sj.ejhg.5201615; published online 12 April 2006

Keywords: mucopolysaccharidosis II; methylation; CpG dinucleotide; IDS; recurrent mutation

*Correspondence: S Tomatsu, Department of Pediatrics, Saint Louis University Pediatric Research Institute, 3662 Park Ave., St Louis, MO 63110-2586, USA.
Tel: +1 314 577 5623, ext. 6213; Fax: +1 314 577 5398;
E-mail: tomatsus@slu.edu

Contract grant sponsor: Austrian Research Society for Mucopolysaccharidoses and Related Diseases, German MPS Society, Italian MPS Society, and International Morquio Organization (Carol Ann Foundation).
Received 14 September 2005; revised 14 February 2006; accepted 15 February 2006; published online 12 April 2006

Introduction

Methylation of cytosine is the only known endogenous modification of DNA in mammals and occurs by the enzymatic addition of a methyl group to the carbon-5 position of cytosine. In 98% of the genome, CpGs are present approximately once per 80 dinucleotides. In contrast, CpG islands, which comprise 1–2% of the genome, are approximately 200bp to several kb in length and have a frequency of CpGs approximately five times greater than the genome as a whole.^{1,2}

The epigenetic modification of eukaryotic DNA by the methylation of cytosine residues modulates access to and regulation of genetic information.³ This process is implicated in the regulation of gene expression, oncogenesis, imprinting, and X-inactivation. Moreover, one of the most significant and well-known aspects related to DNA methylation is mutability. DNA methylation is a major contributor to point mutations leading to human genetic disease as a consequence of deamination of 5-methylcytosine present within CpG dinucleotides.^{4,5} In spite of the relative paucity of the number of CpG dinucleotides and the presence of a sophisticated repair system, more than one-third of all point mutations causing human genetic diseases are derived from C-to-T or G-to-A transitions at a CpG site.⁴

The methylation status of specific CpG sites does exhibit significant inter-individual variation, although differences in site-specific methylation patterns between different ethnic groups have not so far been apparent in studies of a variety of DNA sequences.^{6–8} Most methylation analyses in housekeeping genes, imprinted genes, or oncogenes have been limited to the 5' flanking sequencers. However, direct and specific observations of CpG methylation at the sites of recurrent mutations in a portion of genes have been described.^{9–11} To gain a better understanding of the frequent transitional mutation events at CpG dinucleotides, studies on pattern of cytosine methylation covering the promoter region and the whole coding region have been performed.^{12,13} These studies showed a high mutation frequency at methylated CpG sites and definition of the boundaries between unmethylated and methylated regions. However, to date no study has compared differences at the methylation level of specific CpG sites between the normal population and a patient group in which CpG transitional mutations lead to an inherited genetic disease.

Hunter syndrome (or mucopolysaccharidosis type II) is an X-linked recessive disease caused by the deficiency of the enzyme, iduronate-2-sulphatase (IDS; EC3.1.6.13, MIM#309900). In humans, this deficiency leads to accumulation of heparan sulfate and dermatan sulfate in lysosomes and their excretion in the urine. Clinically, the phenotype varies from severe to mild, depending on the different mutations at the IDS gene. IDS cDNA and its full genomic DNA sequences have been isolated

and characterized.^{14–16} At least 307 independent point mutations (160 different nucleotide changes) have been described in Hunter patients^{17–28,32} (mutations are cited in www.hgmd.org).²⁸

At least 40% of point mutations were C-to-T transitional mutations at CpG sites.²⁸ The 56 CpGs represent only 3.6% of the IDS gene sequence. These findings indicate that the incidence of transitional mutational events at CpG sites in the IDS gene is about 30 times greater than expected. The methylation pattern in five normal healthy males showed that cytosine methylation at the CpG sites was extensive, except for sites in the promoter region to exon 3.²⁹ The CpG sites in exons 4–9, where over 90 independent transitional mutations were found, were completely methylated in healthy male control samples. Only two transitional events were observed in exons 1 and 2, which were unmethylated or hypomethylated. These findings support the importance of CpG cytosine methylation in predisposition to C-to-T transition mutation in the IDS gene. Exceptions were the CpG sites in exon 3, which were unmethylated or hypomethylated in healthy male controls, yet associated with a high rate of transitional mutations in Hunter patients. This lack of correlation between transitional mutations at CpG sites in exon 3 and the methylation status in this region in normal healthy males might be explained if there was some polymorphism in the degree of methylation of CpG sites in exon 3 in humans and CpG mutations in exon 3 occur predominantly in those patients who show greater than normal methylation in this region.

To test this hypothesis, we compared methylation patterns of specific CpG dinucleotides between normal healthy males and Hunter patients with CpG transitional mutations. Our results support the hypothesis that the CpG sites in this region are more highly methylated in exon 3 of Hunter patients with mutations in this region.

Materials and methods

Sample preparation

Genomic DNAs of 11 Hunter patients with the transitional mutations and five normal healthy males were isolated from peripheral blood samples and analyzed to determine the distribution and the percentage of methylated cytosine residues at the IDS gene locus encompassing the upstream region of initial codon ATG and exon 1 to a part of intron 3. The 11 patients included one with p.R8X in exon 1 (Patient 11, 28 years old, referred as JS²⁷), four with p.A85T (Patient 4, 15 years old, referred as HT65;²⁵ Patient 9, 6 years old, cousin of HT65; Patient 10, 5 years old, cousin of HT65; Patient 23, 9 years old, referred as H6²⁴), three with p.P86L (Patient 1, fetus, referred as Fe 1;³⁰ Patient 14, 1 year old, referred as CC;²⁷ Patient 28, 1 year old, referred as no. 6²³), and three (Patient 2, fetus, referred as Fe 2;³⁰ Patient 17, 6 years old, referred as No. 7;²² Patient 33, 2 years old,

referred as no. 8²³) with p.R88H in exon 3. Patients 4, 9, and 10 were in the same family tree. Mutation analysis of the patients was described previously.^{22–25,27,30} Peripheral genomic DNAs from five unrelated healthy males were also isolated.

Sodium bisulfite treatment and sequencing

Bisulfite conversion of genomic DNA was performed using the protocol described previously with minor modification,³¹ as we recently described for analysis of the β -glucuronidase and IDS genes.^{12,29}

The following regions were sequenced and analyzed: approximately 700 bp upstream of initiation codon ATG, 418 bp of exons 1–3 in the coding region, 741 bp of intron 1, 666 bp of intron 2, and 2273 bp of intron 3, covering 49, 23, 9, 11, and 12 CpG sites, respectively (Figure 1).

The following nucleotide (nt) numbering scheme was provided throughout the text. The +1 of nucleotide number was counted from the transcription starting point, namely the first nucleotide of the 5'UTR. Thus, the A of the transcription starting point was as 1 with this numbering system, while the A of ATG initiation codon was designated as 301. Accession numbers of GenBank used here were NM_000202.2 and AF_011889.1.

PCR primers and conditions are described in Table 1 and its legend. PCR primers were determined by using GenBank sequence modified for predicted changes following bisulfite treatment.

Cloning and sequencing

DNAs for PCR came from the 11 patients (one with p.R8X, four with p.A85T, three with p.P86L, and three with p.R88H in exon 3) and five normal male controls. The 12 PCR fragments (sense strand; F1–F6, antisense strand; aF1–aF6) were generated from each individual with the nested primer pairs (Table 1). The amplified fragments were run on a 2% agarose gel. Each isolated fragment was ligated into Easy T-vector and the ligation product was transformed according to the manufacturer's protocol (Promega). In all, 10 randomly selected clones from each transformant were sequenced. The percent methylation at individual CpG sites was determined from the frequency of clones having a cytosine residue at each site in the sequence analysis. In Figure 2, the average of methylation pattern analyzed in the same group (normal males, p.R8X, p.A85T, p.P86L, or p.R88H) is shown. Namely, methylation pattern in Figure 2 shows the average of 50 clones derived from five normal healthy male DNA, 10 clones from one p.R8X, 40 clones from four p.A85T patients, 30 clones from three p.P86L patients, and 30 clones from three p.R88H patients.

The methylation status of CpGs was divided into five groups according to percentage of methylation as previously described:^{12,13,29} briefly, unmethylated (0%),

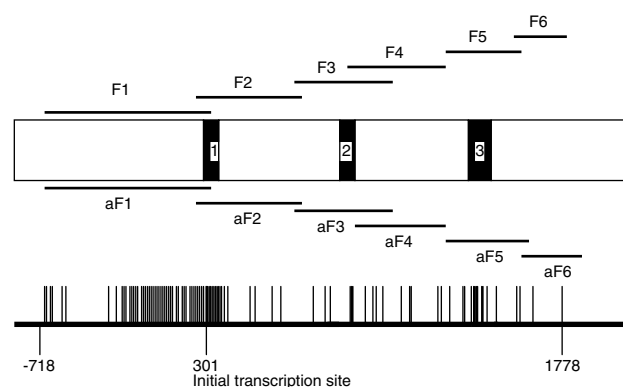


Figure 1 Locations of PCR fragments used in genomic sequencing analysis of the human IDS gene. Locations of PCR fragments used in genomic sequencing analysis of methylation status of the human IDS gene. The numbered boxes represent each exon and the horizontal bars show the position of respective PCR fragment. The upper bars analyze the sense strand and the lower bands analyze the antisense strand, respectively. Primer sets to amplify the fragments are described in Table 1. CpG density plot of the human *IDS* (–718 to 1778 nucleotide position) CpG sites and their distribution covering the entire gene are indicated as small vertical lines.

slightly methylated ($0 < \chi \leq 25\%$), lightly methylated ($25 < \chi \leq 50\%$), moderately methylated ($50 < \chi \leq 75\%$), and heavily or completely methylated ($75 < \chi \leq 100\%$). The term 'partially methylated' was used when the results of sequencing at an individual CpG site showed even one methylated clone among investigated alleles and the term 'hypomethylated' was used for CpG sites defined as $\chi \leq 50\%$ in the text. The term 'hypermethylated' was used for CpG sites defined as $50\% < \chi$ in the text. The CpGs in the exons and the introns are described by number in the order of upstream to downstream direction, such as 'e1, e2, e3...' and 'i1, i2, i3...' (eg, the first CpG in exon 1 is e1 and the tenth CpG downstream e1 is e10). The CpGs in the promoter region are p-7 to p-1 (relative to the start of exon 1).

Results

Comparison of methylation patterns between Hunter patients and normal male controls

We previously established that there were demarcated methylated, unmethylated, and partially methylated regions of the IDS gene in normal healthy controls.¹² In the present study, the methylation status of the patients with transitional mutations at CpG sites in exon 1 or 3 was examined (Figure 2). Figure 2 describes the average of methylation profiles of five controls and the patients with the same mutation at both sense strand and antisense strands. The figure shows the classification of the average methylation status for the population of molecules at each CpG dinucleotide. Namely, the p.R8X comes from Patient 11, the p.A85T from Patients 4, 9, 10, and 23, the

Table 1 (a) Primers used for methylation study on IDS (sense strand) and (b) Primers used for methylation study on IDS (antisense strand)

(a) Primer	Sequence (sense strand)	Localization	Fragment
mH40	TTGAATTTTTGTTGGGAAAAAGTAAGTGGAGGTGT	Promoter	F1
mH41	TTGAAGAAATAGGGGATTTTATTGA	Promoter	F1
mH42R	AACCAAACCCAACCAAAAAAACCTCAA	Exon 1	F1
mH43R	ATAATCAAATTAACCTACATTTCAAATCCAAAA	Exon 1	F1
mH7'	GAAGGGTAGGGTTTTAGGTGAAAGGTATAGT	Promoter	F2
mH8'	TATTTTTAGTAGTTTTATATTTAGGGGTT	Promoter	F2
mH4R'	AAATTAACCTTTAATCATAACCTCCAACCTAAA	Intron 1	F2
mH5R'	AACTAAACCTACACCTTAACCTAACCAAAA	Intron 1	F2
mH9	AGGTTGGAGGTTATGATTAAGTTTAAATTTTGT	Intron 1	F3
mH10	TGTTAAGGTGTAGGGTTTGGGTTTGGGTT	Intron 1	F3
mH11R	CACAATCTATACCATCTAACAATAACTA	Intron 2	F3
mH12R	CCAACTATACTCCCTCTAACAAAATATCCCA	Intron 2	F3
mH49	TTTTGGGTTGTTATGGGGATAAGTTGGTGAGGTTT	Exon 2	F4
mH50	AGGTATGTTTGGGAATTTTATGTTGGGTGTGTGTTGT	Exon 2 – intron 2	F4
mH51R	CAAAACATATCTACCATCTCTTAAATAACCAAAA	Intron 2	F4
mH52R	TACAAACAAAATAATATCAAATCTCCTACCAATA	Intron 2	F4
mH13	TGGTTGAGTTTTGTATGATTTAGTTATTAAGT	Intron 2	F5
mH14	TTGTTTGGTTATTTAAGAGATGTTAGAT	Intron 2	F5
mH15R	TCAAACACCACAAACCAAAAAAACCCAA	Intron 3	F5
mH16R	AAAAAACTCTACAAAAACAACCCCATCCCCAAA	Intron 3	F5
mH15	GTTTGGGTTTTTTGGTTTGTGGTGTGTTGAATTT	Intron 3	F6
mH16	TGGGGATGGGGTTGTTTTGTAGAGTTT	Intron 3	F6
mH44R	ATAAACTAAATACACTCTCAACCTAA	Intron 3	F6
mH45R	ATTCCTAACAAATTCACCTCTAAAACCTAACCTA	Intron 3	F6
(b) Primer	Sequence (antisense strand)	Localization	Fragment
aH52	AATTCTCAAATTAATAAATCAAATATATTTACAA	Promoter	aF1
aH53	ATATAAACTAAATCCATCTCCCCACAATTAATA	Promoter	aF1
aH54R	GTGGTATTGGGGTTTTATGTTTTGTTTTGGTGGAGT	Exon 1	aF1
aH55R	GTTATTTTTTTTTATGAGGTTTTTTGAGAAAGTT	Exon 1	aF1
aH7	ACTCCACCAAAAACAAAAACATAAAAAACCAATA	Promoter	aF2
aH8	CTATACCCCTAACAATTCTACACCTAAA	Promoter	aF2
aH4R	GTGGTAGGGTTAGGTTTTGGTTATGGTT	Intron 1	aF2
aH5R	GTGTTAGGTTATTAAGAGGTTAAGGTTGGATT	Intron 1	aF2
aH9	AACCATAACCAAAAACCTAACCCCTACACCCAAA	Intron 1	aF3
aH10	AAACCTCTAATAACCTAACACCTACCATAAAA	Intron 1	aF3
aH11R	TTGTGTTATTTGATAATAGTTGAAGTTT	Intron 2	aF3
aH12R	TTGAAGGTTAAGTTTTTTTATGAGTGGAAAATT	Intron 2	aF3
aH48	AAACTAATAAAAATCCCCAAATATTACCAACTAA	Exon 2	aF4
aH49	CAACTAACATCCACAACCTCCTCTCCAAAATA	Exon 2	aF4
aH50R	AAATTAGATATTGGGATTTGGGTGTTATTTAGGT	Intron 2	aF4
aH51R	TAGTAAATAGTTTGGTAGTTGAGTTATGTAGAGT	Intron 2	aF4
aH13	TTACCTAAATAACACCCAAATCCCAATATCTAA	Intron 2	aF5
aH14	CTACATAACTCAACTACCAAACTATTTA	Intron 2	aF5
aH15R	TTAGGATGGGAATGTTGGATTAGATAT	Intron 3	aF5
aH16R	TTTAGTTAGAAGGGTTTTGTAAAGATAGTTT	Intron 3	aF5
aH15	GTTTGTGATATCTAAATCCAACATTCCTCCTA	Intron 3	aF6
aH16	AACTATCTTTACAAAACCTCTTCTAACTA	Intron 3	aF6

Notes: PCR amplification was performed in reaction mixtures containing bisulfite-treated genomic DNA, 200 mM dNTPs, 1 mM primers, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂ and Taq DNA polymerase. PCR amplification was done under the following conditions: 94°C for 5 min for one cycle; 95°C, 45 s, 60–65°C for 30 s, 72°C for 45 s for 35 cycles. After purification of the 1st PCR product, the second PCR proceeded in reaction mixture from the first PCR product and the same reaction buffer. The primers used are a new set of two nested primers or a set of one primer used in the first PCR and one new nested primer. Namely, two sets of nested primers were synthesised to get one fragment. PCR amplifications were performed for both strands by use of each strand-specific primer. To include the promoter to part of intron 3, four PCR fragments were prepared for one strand (16 primers of PCR amplification for each strand). The 32 different primers were used for PCR amplification. The sequence information on human *IDS* for each primer was from GenBank (AF011889.1). The original GenBank sequence was modified with bisulfite treatment. Primers sequences were based on the modified sequences.

p.P86L from Patients 1, 14, and 28, and the p.R88H from Patients 2, 17, and 33. A total of 10 clones from each fragment were sequenced and the percentage of methylation

was calculated. In the case of the p.A85T, total 40 clones from the four patients were sequenced and used for calculation of the average of methylation. In all, 10 clones

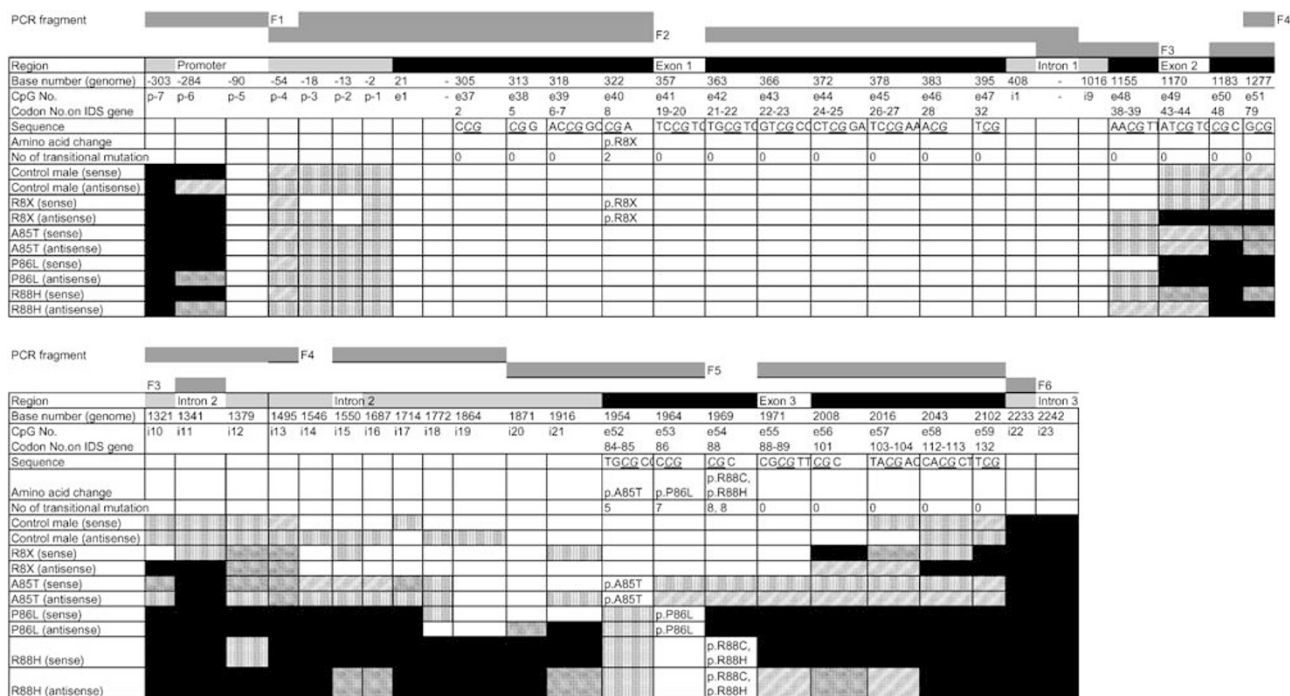


Figure 2 Methylation analysis of 5' gene flanking region and exon 1 to intron 3 at human IDS gene locus. Figure 2 describes the average of methylation profiles of five controls and the patients with the same mutation at both sense strand and antisense strands. The figure shows the classification of the average methylation status for the population of molecules at each CpG dinucleotide. Namely, the p.R8X comes from Patient 11, the p.A85T from Patients 4, 9, 10, and 23, the p.P86L from Patients 1, 14, and 28, and the p.R88H from Patients 2, 17, and 33. Methylation pattern of an individual patient is not shown here (available on request). It includes the promoter, exons 1–3, and part of intron 3, with methylation of both strands. The PCR fragments of F1–F6 for the sense strand (or aF1–aF6 for the antisense strand) were amplified from DNAs after bisulfite treatment and cloned. The PCR fragment incorporates CpG sites p-11 to i25 corresponding to the coordinates base numbers from –418 to 2378. Each fragment was independently cloned, and the methylation status from p-11 to i25 was not continuous over each fragment. The degree of methylation is indicated as follows: where the open rectangle is 0%, the vertical striped one is $0 < \chi \leq 25\%$, the diagonally striped one is $25 < \chi \leq 50\%$, the crosshatched one is $50 < \chi \leq 75\%$, and the black one is $75 < \chi \leq 100\%$. The sequence numbering used here is relative to the transcription starting site. The CpGs in the promoter, exons and the introns are described by number in the order of upstream to downstream direction, such as 'p-7, p-6, p-5.... p-1', 'e1, e2, e3....' and 'i1, i2, i3....' (eg, p-7 to p-1 is relative to the start of exon 1, the first CpG in exon 1 is e1 and the tenth CpG downstream e1 is e10). The accession number for the sequence is AF011889.1.

from each fragment of one individual control were sequenced and the percentage of methylation was calculated. Total 50 clones from the five controls were sequenced and used for calculation of the average of methylation.

Overall, we found that the sense strand of DNA of some patients with a transitional mutation had higher methylation compared to that in the normal male controls. Most of the difference was in exons 2 and 3. Greater methylation in the antisense strand correlated well with the greater methylation status of the sense strand except for a few sites (Figure 2). In both strands of patient DNA, all CpG dinucleotides between positions –418 and –303 (p-11–p-7) were hypermethylated as in normal DNA (Figure 1). This pattern was followed by the sequences where both strands had similar partial methylation of the first boundary, positions –284 to –2 (p-6–p-1). The extent of methylation in the first boundary in patient DNA was similar in pattern with that from normal controls. In the

next region downstream, all the CpG sites from nucleotide positions 21–1016 (e1–i9) were completely unmethylated in both strands, identical with the finding of the normal male controls.

From CpG site i22 (nucleotide 2232) in the IDS gene, all CpG sites were completely methylated in both strands without exceptions, as in controls. However, upstream from this point through exon 3, differences were seen between DNA from normals and Hunter patients. The second boundary area between CpGs e48 and e59 (around 1000 bp), which was broader than the first boundary area (200–300 bp), showed a pattern of higher methylation in Hunter patients than controls.

The average methylation level at CpG sites (e48–e59) in patients with p.A85T, p.P86L, or p.R88H (exon 3) mutations was compared (Figure 2). The CpG sites in the three p.P86L patients were the most methylated, while those from p.A85T patients were less hypermethylated. Except for e48 and i18–i21, all CpG sites in the three patients with

a p.P86L mutation were completely methylated. The CpG sites in p.R88H patients were hypermethylated except for e48, e49, e52, and e53, while the CpG sites in p.A85T patients were less hypermethylated at CpG sites of e48, e49, and i14–i21. Thus, the CpG sites near the p.A85T mutation were less hypermethylated, but still more highly methylated than controls (Figure 2).

When the individual methylation patterns in four patients with a p.A85T mutation were compared, the methylation patterns in Patients 4, 9, and 10, who derived from the same family tree showed differences from each other at several CpG sites (e50, e51, i11, i13, i17). Increased methylation in Patient 4 was observed at the CpG sites (i12–i15) of the sense strand (not antisense strand). Methylation was also increased at CpG sites (e48, e50, e51, i11, i13, e53, e54, e56, e57) in unrelated Patient 23. Individual hypermethylation patterns in three patients with the p.P86L mutation and three patients with the p.R88H mutation were similar except for a few CpG sites. However, some individual CpG sites in each patient showed a clear difference in the percentage of methylation between the strands (data not shown). For example, reduced methylation of the antisense strand relative to the sense was noted in position e55 in Patients 17 and 33. These strand differences in methylation status at certain sites could reflect hemi-methylation. The methylation pattern of Patient 11 with a p.R8X mutation in exon 1 showed the area surrounding the mutation site to be completely unmethylated, unlike the other 10 patients with mutation sites in exon 3 (Figure 2). However, the secondary boundary area (e48–i13, e56–e59) in this patient was more methylated than in normal male controls.

Discussion

The correlation between CpG hypermethylation and transitional mutations was described previously.^{9–13} Except for studies in tumors, no prior study has compared the methylation pattern on the causative gene in affected patients with that of normal controls. We have demonstrated that specific CpG sites in areas of exon 2 to intron 3 were more methylated in Hunter patients with a CpG transitional mutation in exon 3 than to the corresponding CpG sites in normal males. This finding is consistent with our hypothesis that the pattern and level of methylation of the IDS gene would be different between these Hunter patients and normal controls.

The CpG sites in exons 1 and 2 in normal controls were hypomethylated. The CpG sites in exon 3 were also hypomethylated in normal controls, but were associated with the highest rate of transitional mutations producing Hunter syndrome except for the CpG sites in exon 9. The finding that all four mutations in exon 3 happened recurrently (over five times) is surprising given that this

exon is normally hypomethylated in normal controls. An explanation for this discrepancy is suggested by our results indicating that the patients with p.P86L and p.R88H mutations showed hypermethylation in exon 2 to intron 3. The same trend was true for the patients with p.A85T, although the degree of hypermethylation was less in this region than in patients with other CpG transitional mutations. It is impossible to provide direct proof that these CpG sites in Hunter patients were actually methylated before introduction of the mutation, since the nucleotide change leading to the disease had already occurred. However, the higher methylation status in the surrounding CpG sites compared to the normal controls suggests that methylation of this region preceded the mutation and may have predisposed it to CpG transitional mutations.

It is also noteworthy that the methylation status of some specific CpG sites shows marked inter-individual variation. This was true even for CpG sites in exon 3 in these patients with p.A85T from the same pedigree. These studies of the variability of the methylation level at IDS gene locus (especially, the boundaries between unmethylated and methylated region) need to be extended to additional Hunter patients and control peripheral DNAs. It would be ideal to also extend them to both sex germlines, if possible.

It is interesting that the unmethylated status of CpG sites in the area of the p.R8X mutation site (exon 1) was no different between the patient with p.R8X and normal males while the patient showed hypermethylation in the area of exon 2 to a part of intron 3. Whether this p.R8X mutation might have occurred on an ectopic methylated CpG site remains unknown. A second p.R8X mutation has been reported previously, but this patient's DNA was not available to be examined.³²

Why mutations occurred only at these four CpG sites within exons 1–3 (not at the other 18 CpG sites in this segment of the coding region of the IDS gene) remains a puzzle. Absence of mutations in the other 10 CpG sites in exon 1 (other than p.R8X) could be explained by the paucity of methylation and possibly also by the absence of an effect of these possible transitions affecting only the leader peptide region on the clinical phenotype. Absence of transitional mutations at CpG sites in exon 2 is of more interest, since all four CpG sites are somewhat more methylated in Hunter patients than in normal controls. However, these also may involve a region of the IDS gene that is not critical for function. Absence of CpG transitional mutations at methylated CpG sites (e55–e59) in exon 3 again could reflect a small impact of these mutations on the structure of the IDS protein, since these sites encode amino acid residues that are not conserved among the sulfatase family proteins (Figure 3). By contrast, the CpG sites of e52–e54 are in codons close to cysteine 84, the highly-conserved catalytic site common to all sulfatase proteins. Cysteine 84 is the site of the post-translational

IDS		M	PPRTGRGLL	WGLVLSVC	---	VALGSET	QAN-STTDAL	NVLLIIIVDDL	R-PSLGCYGD	KLVRSNPIDQ	66
GALNS			MAAVVAATR	WQQLLVL--	---	SAAGM	GASG-APQPP	NILLLLMDM	GWDDLGVYGE	PSRETPNLDR	61
GS6	MRLPLAP	GRLR-RGSPR	HLPSCSPAL-	---LLLVLGG	---	CLGVFGV	AAGT---RPR	NVLLLTDDQ	D-EVLGGMTF	LKKTALIGE	76
HSS			MSCPVPACCA	---LLLVLGL	---	-----C	RA-----RPR	NALLLLADDG	G-FESGAYNN	SAIATPHLDA	52
HSULF1			MKY	SCCALVLAVL	GTELLGSLCS	TVRSRFRGR	IQQERKNIRP	NIILVPTDDQ	D-VELGSLQV	MNKTRKIMEH	72
HSULF2			MGPP	SLVLCLLSAT	VFSLGGSSA	FLSHHRLKGR	FQRDRRNIRP	NIILVLTDDQ	D-VELGSMQV	MNKTRRIMEQ	73
ARSA				MGAPR	--SLLLAL--	-----AAGLA	VA---RPP	NIVLIFADDL	GYDGLGCYGH	PSSTPNLDQ	51
ARSB	MG	PRGAASLPRG	PGPRLRLLPV	VLPLLLLLLL	---	APPGSGA	GE---SRPP	HLVFLADDL	GWNDVGFHG-	SRIRTPHLDA	74
ARSC			MPLRKMKIPF	---LLLFFL-	---	WEAES	HE---ASRP	NIILVMADDL	GIGDPGCGYN	KTIRTPNIDR	57
ARSD			MRSAARRGRA	APAARDSLPV	---LLFLCLL	---LKTCEPK	TAN---AFKP	NILLIMADDL	GTGDLGCYGN	NLRTPNIDQ	71
ARSE			MLHLHHS	CLCFRSLWLA	M---LAVLSSL	---APSASS	DIS---ASRP	NILLMADDL	GIGDIGCGYN	NMRTPNIDR	68
ARSF			MRPRRPLVFM	S---LVCALLN	---	TWPGHTG	CM---TTRP	NIVLLMVDDL	GIGDLGCYGN	DTMRTPHIDR	61
ARSG			MGWFLFK	VLLAGVFSFG	F---LYPLVDF	---	CISGKTR	GQ-----KP	NFVIIILADDM	GWDLGANWA	ETKTDANLDK 66
/											
IDS											
GALNS											
GS6											
HSS											
HSULF1											
HSULF2											
ARSA											
ARSB											
ARSC											
ARSD											
ARSE											
ARSF											
ARSG											
/											
IDS	LASHSLFQ	AFAQQAVCAP	SRVSFLTGR	PDTRRLYDFN	SYWRVH-AG-	NFS-----	--TIPQYFKE	-NGYVTMSVG	KVF-HP-G	140	
GALNS	MAEGLLPN	FYSANPLCSP	SRAALLTGRL	PIRNGFYTTN	AHARNA-YTP	QEIYGGIPDS	EQLLPELLKK	-AGYVSKIYV	KW--HL-G	144	
GS6	M---GMTFSS	AYVPSALCCP	SRASILTGYK	P-HNHVVMN	TLEGN---CS	SKSWQKIQEP	N-TFPAILRS	MCGYQTFPAG	KYLNEY-G	155	
HSS	LARRSLLFRN	AFTSVSSCSP	SRASLLTGLP	QHONGMYGLH	QDVHFNFS--	-----FDK	VRSLPILLSQ	-AGVRTGLIG	KK--HV-G	127	
HSULF1	---GGATFIN	AFVTPMCCP	SRSSMLTGKY	-VHNHNVYTN	---NEN---CS	SPSWQAMHEP	R-TFAVYLN	-TGYRTAFFG	KYLNEYNG	147	
HSULF2	---GGAHFIN	AFVTPMCCP	SRSSILTGYK	-VHNHNTYTN	---NEN---CS	SPSWAQHES	R-TFAVYLN	-TGYRTAFFG	KYLNEYNG	148	
ARSA	LAAGGLRFTD	FYVPVSLCTP	SRAALLTGRL	PVRMGMYP--	---GVLV--	PSSRGGLEPE	EVTVAEVLAA	-RGYLTGMAG	KW--HL-G	127	
ARSB	LAAGVLLDN	YYT-QPLCTP	SRSQLLTGRY	QIRTGLQHQ-	---IWPV---	QPS--CVPLD	EKLLPQLLKE	-AGYTTMHV	KW--HL-G	149	
ARSC	LASGGVLTQ	HLAASPLCTP	SRAAFMTGRY	PVRSGMASWS	---RTGVPL-F	TASSGGLPTD	EITFAKLLK	-QGYSTALIG	KW--HL-G	138	
ARSD	LAEEGVRLTQ	HLAASPLCTP	SRAAFMTGRY	SFRSGMDA-S	---NGYRALQW	NAGSGLPEN	ETTFARILQ	-HGYATGLIG	KW--HQ-G	152	
ARSE	LAEDGVLTQ	HISAASLCTP	SRAAFMTGRY	PVRSGMVS-S	---IGYRVLQW	TGASGGLPTN	ETTFAKILKE	-KGYATGLIG	KW--HL-G	149	
ARSF	LAREGVRLTQ	HISAASLCTP	SRAAFMTGRY	PVRSGMVS-S	---GNRRVIQN	LAVPAGLPLN	ETTLAALLK	-QGYSTGLIG	KW--HQ-G	142	
ARSG	MASEGMRFVD	FHAASSTCSP	SRASLLTGRL	GLRNGVTR-N	-----FA	VTSVGGLEPLN	ETTLAEVLQ	-AGYVTGLIG	KW--HL-G	141	

Figure 3 Multiple amino acid alignment among human IDS and other sulfatase proteins. Definition of abbreviations, GenBank accession numbers, symbols, and terminology are as follows. The gray colored letter indicates a CpG site. The symbol * marks a mutational site described in the text. The symbol/(and not I) indicates the boundary of exons. Abbreviations and accession numbers: GALNS, human N-acetylgalactosamine-6-sulfatase (P34059); IDS, human iduronate-2-sulfatase (P22304); G6S, human N-acetylglucosamine-6-sulfatase (NP_002067); HSS, human heparan sulfaminidase (NP_000190); HSULF1, human sulfatase 1 (NP_055985); HSULF2, human sulfatase 2 (Q8IWU5); ARSA, human arylsulfatase A (NP_000478); ARSB, human arylsulfatase B (P15848); ARSC, human arylsulfatase C (NP_000342); ARSD, human arylsulfatase D (P51689); ARSE, human arylsulfatase E (NP_000038); ARSF, human arylsulfatase F (NP_004033); ARSG, human arylsulfatase G (NP_055775).

modification, which is required for producing catalytically active sulfatases.³³⁻³⁵ Mutations at this site impair activation of IDS and result in clinical disease.

The current study provides the first evidence for a definite difference in the methylation pattern at certain CpG sites between normal controls and patients with Hunter disease. It raises the possibility that individuals who inherit the pattern of higher methylation of exon 2 to intron 3 are at greater risk for IDS mutations producing Hunter syndrome.

Some questions still remain: (1) What determines individual differences in overall methylation pattern of certain genomic regions like the IDS which appears to be polymorphic? How frequent is the hypermethylation in exons 2 and 3 in the normal population? Hunter syndrome occurs in roughly 1 in 10⁵ births and 10% of those cases result from CpG mutations in this region. How large a population may be at risk due to hypermethylation in this region is an interesting question. A much larger study of control patients will be needed to estimate this frequency.

It is known that differences in methylation can be inherited, suggesting either the persistence of certain methylation at all stages of development or the encryption of methylation pattern information.^{10,11} It will be interesting to determine whether the hypermethylation of exons 2 and 3 in these type patients with Hunter syndrome is found in other tissues – for example, cultured fibroblasts,

and whether it is true in the maternal DNA of such proband.

Acknowledgements

This work was supported by grants from the Austrian Research Society for Mucopolysaccharidoses and Related Diseases and Austrian MPS Society, German MPS Society, Italian MPS Society, International Morquio Organization, Jacob Randall Foundation, and Bennett Foundation.

References

- Bird AP: CpG-rich islands and the function of DNA methylation. *Nature* 1986; **321**: 209–213.
- Gardiner-Garden M, Frommer M: CpG islands in vertebrate genomes. *J Mol Biol* 1987; **196**: 261–282.
- Martienssen RA, Richards EJ: DNA methylation in eukaryotes. *Curr Opin Genet Dev* 1995; **5**: 234–242.
- Cooper DN, Youssoufian H: The CpG dinucleotide and human genetic disease. *Hum Genet* 1988; **78**: 151–155.
- Bottema CD, Ketterling RP, Vielhaber E *et al*: The pattern of spontaneous germ-line mutation: relative rates of mutation at or near CpG dinucleotides in the factor IX gene. *Hum Genet* 1993; **91**: 496–503.
- Kochanek S, Toth M, Dehmel A, Renz D, Doerfler W: Inter-individual concordance of methylation profiles in human genes for tumor necrosis factors alpha and beta. *Proc Natl Acad Sci USA* 1990; **87**: 8830–8834.
- Bottema CD, Ketterling RP, Yoon HS, Sommer SS: The pattern of factor IX germ-line mutation in Asians is similar to that of Caucasians. *Am J Hum Genet* 1990; **47**: 835–841.

- 8 Behn-Krappa A, Holker I, Sandaradura dS, Doerfler W: Patterns of DNA methylation are indistinguishable in different individuals over a wide range of human DNA sequences. *Genomics* 1991; **11**: 1–7.
- 9 Andrews JD, Mancini DN, Singh SM, Rodenhiser DI: Site and sequence specific DNA methylation in the neurofibromatosis (NF1) gene includes C5839T: the site of the recurrent substitution mutation in exon 31. *Hum Mol Genet* 1996; **5**: 503–507.
- 10 El Maarri O, Olek A, Balaban B *et al*: Methylation levels at selected CpG sites in the factor VIII and FGFR3 genes, in mature female and male germ cells: implications for male-driven evolution. *Am J Hum Genet* 1998; **63**: 1001–1008.
- 11 Millar DS, Krawczak M, Cooper DN: Variation of site-specific methylation patterns in the factor VIII (F8C) gene in human sperm DNA. *Hum Genet* 1998; **103**: 228–233.
- 12 Tomatsu S, Orii KO, Islam MR *et al*: Methylation patterns of the human beta-glucuronidase gene locus: boundaries of methylation and general implications for frequent point mutations at CpG dinucleotides. *Genomics* 2002; **79**: 363–375.
- 13 Tomatsu S, Nishioka T, Montano AM *et al*: Mucopolysaccharidosis IVA: identification of mutations and methylation study in GALNS gene. *J Med Genet* 2004; **41**: e98.
- 14 Wilson PJ, Morris CP, Anson DS *et al*: Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA. *Proc Natl Acad Sci USA* 1990; **87**: 8531–8535.
- 15 Wilson PJ, Meaney CA, Hopwood JJ, Morris CP: Sequence of the human iduronate 2-sulfatase (IDS) gene. *Genomics* 1993; **17**: 773–775.
- 16 Timms KM, Lu F, Shen Y *et al*: 130 kb of DNA sequence reveals two new genes and a regional duplication distal to the human iduronate-2-sulfate sulfatase locus. *Genome Res* 1995; **5**: 71–78.
- 17 Bunge S, Steglich C, Beck M *et al*: Mutation analysis of the iduronate-2-sulfatase gene in patients with mucopolysaccharidosis type II (Hunter syndrome). *Hum Mol Genet* 1992; **1**: 335–339.
- 18 Sukegawa K, Tomatsu S, Tamai K *et al*: Intermediate form of mucopolysaccharidosis type II (Hunter disease): a C1327 to T substitution in the iduronate sulfatase gene. *Biochem Biophys Res Commun* 1992; **183**: 809–813.
- 19 Hopwood JJ, Bunge S, Morris CP *et al*: Molecular basis of mucopolysaccharidosis type II: mutations in the iduronate-2-sulphatase gene. *Hum Mutat* 1993; **2**: 435–442.
- 20 Jonsson JJ, Aronovich EL, Braun SE, Whitley CB: Molecular diagnosis of mucopolysaccharidosis type II (Hunter syndrome) by automated sequencing and computer-assisted interpretation: toward mutation mapping of the iduronate-2-sulfatase gene. *Am J Hum Genet* 1995; **56**: 597–607.
- 21 Goldenfum SL, Young E, Michelakakis H, Tsagarakis S, Winchester B: Mutation analysis in 20 patients with Hunter disease. *Hum Mutat* 1996; **7**: 76–78.
- 22 Rathmann M, Bunge S, Beck M, Kresse H, Tylki-Szymanska A, Gal A: Mucopolysaccharidosis type II (Hunter syndrome): mutation 'hot spots' in the iduronate-2-sulfatase gene. *Am J Hum Genet* 1996; **59**: 1202–1209.
- 23 Froissart R, Maire I, Millat G *et al*: Identification of iduronate sulfatase gene alterations in 70 unrelated Hunter patients. *Clin Genet* 1998; **53**: 362–368.
- 24 Gort L, Chabas A, Coll MJ: Hunter disease in the Spanish population: molecular analysis in 31 families. *J Inherit Metab Dis* 1998; **21**: 655–661.
- 25 Isogai K, Sukegawa K, Tomatsu S *et al*: Mutation analysis in the iduronate-2-sulphatase gene in 43 Japanese patients with mucopolysaccharidosis type II (Hunter disease). *J Inherit Metab Dis* 1998; **21**: 60–70.
- 26 Karsten S, Voskoboeva E, Tishkanina S, Pettersson U, Krasnopolskaja X, Bondeson ML: Mutational spectrum of the iduronate-2-sulfatase (IDS) gene in 36 unrelated Russian MPS II patients. *Hum Genet* 1998; **103**: 732–735.
- 27 Vafiadaki E, Cooper A, Heptinstall LE, Hatton CE, Thornley M, Wraith JE: Mutation analysis in 57 unrelated patients with MPS II (Hunter's disease). *Arch Dis Child* 1998; **79**: 237–241.
- 28 Stenson PD, Ball EV, Mort M *et al*: Human gene mutation database (HGMD): 2003 update. *Hum Mutat* 2003; **21**: 577–581.
- 29 Tomatsu S, Orii KO, Bi Y *et al*: General implications for CpG hot spot mutations: methylation patterns of the human iduronate-2-sulfatase gene locus. *Hum Mutat* 2004; **23**: 590–598.
- 30 Villani GR, Daniele A, Balzano N, Di Natale P: Expression of five iduronate-2-sulfatase site-directed mutations. *Biochim Biophys Acta* 2000; **1501**: 71–80.
- 31 Clark SJ, Harrison J, Paul CL, Frommer M: High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* 1994; **22**: 2990–2997.
- 32 Filocamo M, Bonuccelli G, Corsolini F, Mazzotti R, Cusano R, Gatti R: Molecular analysis of 40 Italian patients with mucopolysaccharidosis type II: New mutations in the iduronate-2-sulfatase (IDS) gene. *Hum Mutat* 2001; **18**: 164–165.
- 33 Schmidt B, Selmer T, Ingendoh A, von Figura K: A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. *Cell* 1995; **82**: 271–278.
- 34 Cosma MP, Pepe S, Annunziata I *et al*: The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell* 2003; **113**: 445–456.
- 35 Dierks T, Schmidt B, Borissenko LV *et al*: Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell* 2003; **113**: 435–444.