Dimethylglycine Deficiency and the Development of Diabetes

Experimental studies have suggested possible protective effects of dimethylglycine (DMG) on glucose metabolism. DMG is degraded to glycine through a DMG-dehydrogenase (DMGDH)-catalyzed reaction, and this is the only known pathway for the breakdown of DMG in mammals. In this study, we aimed to identify the strongest genetic determinant of circulating DMG concentration and to investigate its associations with metabolic traits and incident diabetes. In the cohort with full metabolomics data (n = 709), low plasma levels of DMG were significantly associated with higher blood glucose levels (P = 3.9E−4). In the genome-wide association study (GWAS) of the discovery cohort (n = 5,205), the strongest genetic signal of plasma DMG was conferred by rs2431332 at the DMGDH locus, where the major allele was associated with lower DMG levels (P = 2.5E−19). The same genetic variant (major allele of rs2431332) was also significantly associated with higher plasma insulin (P = 0.019), increased HOMA insulin resistance (P = 0.019), and an increased risk of incident diabetes (P = 0.001) in the pooled analysis of the discovery cohort together with the two replication cohorts (n = 20,698 and n = 7,995). These data are consistent with a possible causal role of DMG deficiency in diabetes development and encourage future studies examining if inhibition of DMGDH, or alternatively, supplementation of DMG, might prove useful for the treatment/prevention of diabetes.

Using mass spectrometry (MS)-based metabolomic approaches, recent studies have identified associations between small molecules and insulin sensitivity and type 2 diabetes (1–4). Although these circulating metabolites may represent useful markers of disease susceptibility, their causal involvement in the development of diabetes is less certain. For example, although previous studies have shown that high glycine levels are associated with increased insulin sensitivity and decreased type 2 diabetes risk (5–7), a recent study failed to show association between a genetic variant within the CPS-1 gene that was genome-wide significantly associated with glycine, with increased insulin sensitivity and decreased risk of diabetes arguing against a causal involvement of glycine in diabetes development (8).

The clinical importance of assessing causality between circulating biomarkers and disease with the use of Mendelian randomization approaches has been clearly exemplified by genetic studies of LDL- and HDL-cholesterol (9). This proof-of-concept study showed that whereas genetic elevation of LDL-cholesterol associated with coronary artery disease (CAD), genetically lower levels of HDL-cholesterol did not, results that are in concert with the CAD preventive effect of statins and the lack of such effect of drugs elevating HDL-cholesterol by inhibition of cholesteryl ester transfer protein (10,11).

The tertiary amine dimethylglycine (DMG) is produced from betaine during the remethylation of homocysteine to methionine, catalyzed by betaine-homo-cysteine methyltransferase (BHMT). DMG is in turn catalyzed by DMG dehydrogenase (DMGDH) and sarcosine dehydrogenase metabolized to glycine. Epidemiological data have shown that plasma levels of betaine are associated in opposite directions with key components of the metabolic syndrome.
and high plasma levels of glycine have been suggested to be associated with increased insulin sensitivity. However, data on their intermediate metabolite DMG in relation to diabetes development is lacking.

The first aim of this study was to examine the association of plasma levels of DMG and glycine with glycemia and insulin resistance. Secondly, we aimed to replicate genetic associations with DMG (DMGDH locus) (13) and glycine (CPS-1) (8,14) and to investigate whether genetic alterations underlying DMG and glycine concentrations are associated with glucometabolic risk factors and incident diabetes.


RESEARCH DESIGN AND METHODS

Study Samples

Metabolite Cohort

Metabolites were profiled from EDTA plasma collected at the baseline examination of participants of the Malmö Diet and Cancer Study–Cardiovascular Cohort (MDC-CC; see below for a description of the MDC-CC) using previously described methodology (3,15) (see also the section METABOLIC PROFILING for details). These subjects were derived from a nested incident cardiovascular disease (CVD) case-control study (n = 506) (16) with case and control subjects matched by sex, age, and Framingham risk score (17), and a nested incident diabetes case-control study (n = 326) (3). From this pool, 27 subjects who participated in both studies were excluded, leaving 805 subjects. Of these, 709 were successfully genotyped for the genome-wide association study (GWAS) (see GENOTYPING), passed all quality control steps, and had complete data on all covariates (age, sex, BMI). These 709 subjects in the metabolite cohort were used in a GWAS analysis to identify genetic variants associated with the levels of DMG and glycine in plasma. In addition, the metabolite cohort was used to study the relationship between plasma levels of DMG and glycine in relation to metabolic traits (glucose, insulin, and HOMA of insulin resistance [HOMA-index]).

The MDC

The MDC is a prospective population-based study (n = 30,447) where baseline examinations, including anthropometric measurements and blood sample donations including DNA, were performed between 1991 and 1996 (18). A complete description of the study population has been given elsewhere (19).

The MDC-CC

To study the epidemiology of carotid artery disease, a random sample of the MDC study subjects was selected (n = 6,103) during the years 1991–1994, and this sample is referred to as the MDC-CC (20,21). Fasting plasma samples were obtained in 5,405 subjects in the MDC-CC (20). In MDC-CC subjects for whom GWAS genotyping was performed and data were available on fasting glucose and insulin, cross-sectional analyses were performed between plasma DMG and glycine-associated single nucleotide polymorphisms (SNPs) (rs2431332 at the DMGDH locus and rs4673546 at the CPS-1 locus) and plasma glucose (n = 4,790), insulin (n = 4,740), and HOMA-index (n = 4,745). Furthermore, after exclusion of subjects in the MDC-CC with prevalent diabetes at the baseline examination (n = 242), we related rs2431332 and rs4673546 to the incidence of diabetes in 5,205 subjects during a median (interquartile range [IQR]) follow-up time of 16.6 (14.5–18.7) years using local and national registries as described below.

The MDC Replication Cohort

After excluding 1,044 subjects with prevalent diabetes at the baseline examination among MDC participants who were not included in the MDC-CC, we genotyped an additional 20,698 MDC subjects (the MDC replication cohort) for the rs2431332 and rs4673546 using the TaqMan system (see GENOTYPING) for the analysis of these genetic variants and incidence of diabetes during a median (IQR) follow-up time of 14.6 (12.1–17.0) years (see CLINICAL ASSESSMENT).

The Malmö Preventive Project

The Malmö Preventive Project (MPP, 1974–1992) was a preventive program with the aim to screen for cardiovascular risk factors, alcohol abuse, and breast cancer among inhabitants in Malmö, Sweden, born 1921–1949 (n = 33,346). A complete description of the study population has been given elsewhere (MPP baseline examination) (22). A reexamination was conducted during 2002–2006 (n = 18,240; 72% attendance rate, 63% men, mean age 69 ± 6 years) at which whole-blood DNA samples were collected. After excluding subjects who had also participated in the MDC and subjects with prevalent diabetes at the MPP baseline examination, we genotyped the rs2431332 and rs4673546 variants in 7,995 unique subjects of the MPP using the TaqMan system (see GENOTYPING) and related the genetic variants to incidence of diabetes (see CLINICAL ASSESSMENT) during a median (IQR) follow-up time of 28.8 (25.4–30.4) years after the MPP baseline examination.

The study protocols were approved by the Lund University Institutional Review Boards, Sweden. All study participants provided written informed consent.

Metabolic Profiling

Given the relatively high analytical demands and costs of liquid chromatography–MS (LC-MS) profiling, we were not able to analyze the entire MDC-CC cohort (3,16). Plasma metabolites were profiled in EDTA plasma samples drawn at the baseline examination in the MDC-CC. Samples were stored at −80°C and profiled using LC-MS as described in detail previously (3). Formic acid, ammonium acetate, LC-MS grade solvents, and valve-d8 were purchased from Sigma-Aldrich (St. Louis, MO). The remainder of the isotopically labeled analytical standards were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Plasma samples (10 μL) were prepared for LC-MS analyses by protein precipitation with the addition of nine volumes of 74.9:24.9:0.2 v/v/v acetonitrile/methanol/formic acid containing two additional stable isotope-labeled internal standards for valve-d8 and phenylalanine-d8. The samples were centrifuged for 10 min...
Chemistry. We used the HOMA-index (23).

Blood samples were analyzed for fasting blood glucose or use of antihypertensive therapy. The MDC-CC fasting glucose of $\geq 6.1 \text{ mmol/L}$, when available.

Incident diabetes (occurring after the MDC and MPP baseline examinations) was retrieved through record linkage of the Swedish personal identification number with national and regional registers as follows: the Malmö HbA1c register that analyzed all HbA1c samples at the Department of Clinical Chemistry obtained in institutional and noninstitutional care in Malmö are from 1988 and onwards (24), the Swedish National Diabetes Register (25), The Regional Diabetes 2000 register of the Scania region (26), the Swedish National Patient Register covering all somatic and psychiatric hospital discharges and hospital-based outpatient care (27), Swedish Cause of Death register (28), and Swedish Prescribed Drug Register (prescription of antidiabetes medication) (29). In addition, in the parts of MDC (30) and MPP (31) that were reexamined after the two baseline examinations in the two studies, incident diabetes was also based on a measured fasting plasma glucose value $\geq 7.0 \text{ mmol/L}$ or a plasma glucose value $\geq 11.1 \text{ mmol/L}$ after an oral glucose tolerance test.

Genotyping
GWAS genotyping was performed using the HumanOmniExpressExome BeadChip and iScan System (Illumina, San Diego, CA) in the MDC-CC. All other genotyping (of rs2431332 and rs4673546) was performed using “assay by design” TaqMan probes and a real-time PCR assay on an ABI 7900HT (Applied Biosystems), according to the manufacturer’s instructions. Twenty percent of the samples were run in duplicates with $\geq 99.5\%$ concordance. All genotypes were called by two different investigators. None of the SNPs (rs2431332 and rs4673546) showed any significant deviation from Hardy-Weinberg equilibrium ($P > 0.05$).

Statistics
Plasma levels of glycine, DMG, glucose, insulin, and HOMA-index were transformed to their natural logarithm ($\ln$) (due to skewed distribution) and then standardized to multiples of 1 SD. Linear regression analyses adjusted for age, sex, and BMI were performed cross-sectionally between metabolite levels (DMG and glycine) and glucometabolic traits (glucose, insulin, and HOMA-index) in the metabolite cohort. A GWAS was then performed for plasma levels of DMG and glycine in the metabolite cohort using additive genetic models of inheritance and linear regression adjusting for age, sex, and BMI.

Each of the two top genetic signals for DMG and glycine were tested for an association with fasting glucose, insulin, and HOMA-index in the MDC-CC as dependent variables using linear regression with the genotype as the explanatory variable adjusted for age, sex, and BMI.

Finally, we used logistic regression analysis to test the association between the strongest genetic signals of plasma DMG (rs2431332) and glycine (rs4673546) in the GWAS and the incidence of diabetes during follow-up.
in a model adjusted for age, sex, and BMI in three unique sets of individuals free from diabetes at the MDC and MPP baseline examinations, the MDC-CC, the MDC replication cohort, and the MPP. The GWAS was performed using PLINK software. All other analyses were performed in SPSS 20.0 software (IBM Corp., Armonk, NY). A two-tailed P value of <0.05 was considered statistically significant.

RESULTS

Characteristics of the study participants in the metabolite cohort, as well as the MDC-CC, MDC replication, and MPP cohorts, are reported in Table 1 and Supplementary Table 1.

DMG and Glycine Plasma Levels and Association With Metabolic Traits

Lower plasma levels of DMG were significantly associated with higher glucose levels but neither with insulin levels nor the HOMA-index. Low plasma levels of glycine were significantly associated increased glucose, insulin, and HOMA-index (Table 2).

Identification of the Strongest DMG- and Glycine-Associated Variance in the Genome

We identified a genome-wide significant association between the SNP rs2431332 (major allele associated with lower plasma levels of DMG) at the DMGDH locus and its enzyme substrate DMG (P = 2.5E-15), representing the strongest signal of the GWAS (Table 3). Furthermore, we identified the strongest signal of the GWAS for plasma glycine at a region near the CPS-1 gene, rs4673546, the major allele associated with higher plasma levels of glycine (P = 1.0E-12) (Table 3).

Effects of Metabolite-Associated SNPs on Metabolic Traits

In linear regression analysis adjusted for age, sex, and BMI, the major allele of rs2431332 (associated with lower DMG levels) showed significant direct relationships with insulin (β = 0.047, P = 0.019) and the HOMA-index (β = 0.046, P = 0.019) and a borderline significant association with glucose (β = 0.040, P = 0.067). The SNP in the CPS-1 locus (major allele of rs4673546) showed no significant associations with the measured glucometabolic traits (glucose, P = 0.872; insulin, P = 0.690; and HOMA-index, P = 0.986) (Table 4).

DMG Deficiency and Association With Diabetes Development

During follow-up of the MDC-CC (n = 5,205), there were 711 cases of incident diabetes. The logistic regression analyses, adjusted for age, sex, and BMI, showed that the major allele of rs2431332 was significantly associated with an increased risk of diabetes (odds ratio 1.149 [95% CI 1.002–1.315], P = 0.048) (Table 5). In contrast, the major allele of rs4673546 showed no significant association with diabetes development (0.988 [0.854–1.144], P = 0.875). We then attempted to replicate the association between the major allele of rs2431332 and incident diabetes in two large, independent populations. In the MDC replication cohort (n = 20,698), there were 2,180 cases of incident diabetes during follow-up and a borderline significant association between the major allele of rs2431332 and incident diabetes (1.078 [0.998–1.164], P = 0.057) (Table 5). In the MPP (n = 7,995), there were 1,310 cases of incident diabetes during follow-up, and the major allele of the rs2431332 was significantly associated with diabetes incidence (1.116 [1.006–1.238], P = 0.038). Pooled analysis across all three cohorts (MDC-CC, MDC replication, and MPP) was significantly associated with incident diabetes (1.101 [1.041–1.165], P = 0.001) (Fig. 1).

DISCUSSION

Most recently a GWAS identified a novel genetic variant at the DMGDH loci to be strongly (P = 6.6E-33) associated with the levels of DMG in plasma (13). Nonetheless, due to the exploratory nature of the study (GWAS of 217 plasma metabolites), possible correlations of metabolite-associated genetic variants with metabolic traits were not accounted for in the study (13). We were able to confirm this previously described genome-wide significant association between variants at DMGDH and its enzyme.
An increase in the DMG concentration in urine of treated rats was tested, showing that plasma levels of DMG significantly increased with physiologic stress, such as illness, as well as during times of increased physical activity. He experienced chronic muscular fatigue beginning in adolescence, accompanied by elevated levels of the muscle form of creatine kinase in serum. However, no details regarding the metabolic status of this patient were noted (32,33).

The literature regarding the effect of DMG on glucose metabolism in humans is scarce. A recent metabolic profiling study of 377 Framingham Offspring Study participants without diabetes, in which the effect of an oral glucose challenge on plasma levels of 110 metabolites was tested, showed that plasma levels of DMG significantly decreased during an oral glucose challenge (34). Interestingly, an animal study has demonstrated that the insulin-, glucose-, and triglyceride-lowering effect by DMGDH deactivation is strongly associated with fasting insulin and clamp-based measures of insulin sensitivity, although no associations were seen for the glycine-associated SNP and measurements of insulin resistance (8). These findings are concordant with our findings in which glycine plasma levels were significantly associated with glucose and measurements of insulin resistance but the glycine-associated SNP at the CPS-1 locus was not. These findings suggest that lower glycine levels might be secondary to DMG metabolism and/or to the diabetes disease process itself.

Apart from being the end product of DMG catalyzed by DMGDH and sarcosine dehydrogenase, glycine, on one hand, is also synthesized from serine, threonine, and hydroxyproline via interorgan metabolism involving primarily the liver and kidneys (36), which makes the interpretation of the SNP/metabolite associations more complex. DMG, on the other hand, is formed from the peroxisome proliferator-activated receptor α (PPAR-α) agonist WY14 643 is accompanied with a significant reduction in DMGDH function, together with a substantial increase in the DMG concentration in urine of treated rats compared with control rats, suggesting an extension of the role of PPAR-α beyond the control of fatty acid and glucose in rodents (35).

We could also confirm previous results showing a genome-wide significant association between plasma glycine levels and a genetic variant near the CPS-1 locus (8,14). In one of these studies, plasma levels of glycine were strongly associated with fasting insulin and clamp-based measures of insulin sensitivity, although no associations were seen for the glycine-associated SNP and measurements of insulin resistance (8). These findings are concordant with our findings in which glycine plasma levels were significantly associated with glucose and measurements of insulin resistance but the glycine-associated SNP at the CPS-1 locus was not. These findings suggest that lower glycine levels might be secondary to DMG metabolism and/or to the diabetes disease process itself.

Apart from being the end product of DMG catalyzed by DMGDH and sarcosine dehydrogenase, glycine, on one hand, is also synthesized from serine, threonine, and hydroxyproline via interorgan metabolism involving primarily the liver and kidneys (36), which makes the interpretation of the SNP/metabolite associations more complex. DMG, on the other hand, is formed from the peroxisome proliferator-activated receptor α (PPAR-α) agonist WY14 643 is accompanied with a significant reduction in DMGDH function, together with a substantial increase in the DMG concentration in urine of treated rats compared with control rats, suggesting an extension of the role of PPAR-α beyond the control of fatty acid and glucose in rodents (35).

We could also confirm previous results showing a genome-wide significant association between plasma glycine levels and a genetic variant near the CPS-1 locus (8,14). In one of these studies, plasma levels of glycine were strongly associated with fasting insulin and clamp-based measures of insulin sensitivity, although no associations were seen for the glycine-associated SNP and measurements of insulin resistance (8). These findings are concordant with our findings in which glycine plasma levels were significantly associated with glucose and measurements of insulin resistance but the glycine-associated SNP at the CPS-1 locus was not. These findings suggest that lower glycine levels might be secondary to DMG metabolism and/or to the diabetes disease process itself.

Apart from being the end product of DMG catalyzed by DMGDH and sarcosine dehydrogenase, glycine, on one hand, is also synthesized from serine, threonine, and hydroxyproline via interorgan metabolism involving primarily the liver and kidneys (36), which makes the interpretation of the SNP/metabolite associations more complex. DMG, on the other hand, is formed from the degradation of the BHMT-catalyzed reaction, and in turn, catalyzed by DMGDH and degraded to glycine; this is the only known pathway to this metabolite in mammals. Also, because DMG is synthesized from betaine during the remethylation of homocysteine to methionine, catalyzed by BHMT, it is rational that epidemiological studies have indicated metabolic protective proprieties of plasma betaine and choline (the precursor of betaine) (12). Our

### Table 3—The strongest genetic signals of plasma DMG and glycine in a GWAS of participants in the MDC-CC

<table>
<thead>
<tr>
<th>Locus</th>
<th>Trait</th>
<th>SNP/exon</th>
<th>Ch</th>
<th>Position</th>
<th>Major/minor allele</th>
<th>MAF</th>
<th>β</th>
<th>SE</th>
<th>P value</th>
<th>Gene function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMGDH</td>
<td>DMG</td>
<td>rs2431332</td>
<td>5</td>
<td>78321109</td>
<td>A/G</td>
<td>0.2507</td>
<td>-0.467</td>
<td>0.058</td>
<td>2.5E-15</td>
<td>DMGDH catalyzes oxidative demethylation of DMG to sarcosine</td>
</tr>
<tr>
<td>CPS-1</td>
<td>Glycine</td>
<td>rs4673546</td>
<td>2</td>
<td>211565692</td>
<td>C/T</td>
<td>0.1824</td>
<td>-0.453</td>
<td>0.062</td>
<td>1.0E-12</td>
<td>CPS-1 catalyzes the conversion of ammonia and bicarbonate to carbamyl phosphate</td>
</tr>
</tbody>
</table>

Ch, chromosome; MAF, minor allele frequency. *The β values per allele (major allele coded) in relation to metabolite levels expressed on a standardized scale. Regressions are adjusted for age, sex, and BMI.

### Table 4—rs2431332 and rs4673546 correlations to metabolic traits in the MDC-CC

<table>
<thead>
<tr>
<th>Glucometabolic trait</th>
<th>Standardized β-coefficient (SE)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>0.040 (0.022)</td>
<td>0.067</td>
</tr>
<tr>
<td>Insulin</td>
<td>0.047 (0.020)</td>
<td>0.019</td>
</tr>
<tr>
<td>HOMA-index</td>
<td>0.046 (0.020)</td>
<td>0.019</td>
</tr>
<tr>
<td>Glucose</td>
<td>-0.004 (0.024)</td>
<td>0.883</td>
</tr>
<tr>
<td>Insulin</td>
<td>0.009 (0.022)</td>
<td>0.690</td>
</tr>
<tr>
<td>HOMA-index</td>
<td>0.000 (0.022)</td>
<td>0.986</td>
</tr>
</tbody>
</table>

The β values per allele (major allele coded) in relation to insulin, glucose, and HOMA-index expressed on a standardized scale. Regressions are adjusted for age, sex, and BMI.

### Table 5—rs2431332 and correlation to incident diabetes in three independent Swedish cohorts

<table>
<thead>
<tr>
<th>rs2431332</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDC-CC</td>
<td>1.149</td>
<td>1.002–1.317</td>
<td>0.048</td>
</tr>
<tr>
<td>MDC</td>
<td>1.078</td>
<td>0.998–1.164</td>
<td>0.057</td>
</tr>
<tr>
<td>MPP</td>
<td>1.116</td>
<td>1.006–1.238</td>
<td>0.038</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>1.101</td>
<td>1.041–1.165</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Odds ratio per allele (major allele coded). Regressions are adjusted for age, sex, and BMI.
finding that genetic variation (major allele of rs2431332) within the DMGDH locus was associated with lower DMG concentrations, which in turn was associated with higher blood glucose and was also associated with increased insulin resistance as well as with increased risk of incident diabetes in multiple independent large cohorts, urges for further studies examining if a possible inhibition of DMHDH or supplementation of DMG might prove useful for the treatment/prevention of diabetes.

A significant strength of the current study is the use of three independent well-characterized prospective cohorts that have been monitored longitudinally for diabetes incidence for decades using nationwide registers with 100% coverage and high known accuracy (18–21). A weakness of the study is the lack of measurements of sarcosine, the intermediate product between DMG and glycine. Furthermore, our samples consisted predominantly of individuals of European ancestry, and therefore, the results of this study may not be generalizable to other racial/ethnic groups.

Given the genetic correlation between rs2431332 and measures of insulin resistance, it was surprising that plasma DMG did not correlate with insulin resistance indexes. One reason may be insufficient power. Another reason may be confounding influences on the relationship between environmentally determined DMG, which explains the major part of DMG variance, and insulin resistance compared with the unconfounded relationship between genetically determined DMG (rs2431332) and insulin resistance.

Finally, we do not have the data to elucidate the molecular consequences of rs2431332; however, because it is not in the coding region of DMGDH, it most likely increases gene expression and thus total activity.

Our results offer an indication for a potential causal role of DMG in the etiology of diabetes. Further studies are needed to investigate the mechanisms underlying this association and possibilities for preventive interventions.

Funding. M.M. and O.M. were supported by grants from the Swedish Medical Research Council, the Medical Faculty of Lund University, Skåne University Hospital, the Albert Pålsson Research Foundation, the Crafoord Foundation, the Emhild Lundströms Research Foundation, the Region Skåne, the Hulda and Conrad Mossfelt Foundation, the Southwest Skånes Diabetes Foundation, the King Gustaf V and Queen Victoria’s Foundation, the Lennart Hanssons Memorial Fund, Knut and Alice Wallenberg Foundation, and the Marianne and Marcus Wallenberg Foundation. P.N. was supported by grants from the Swedish Medical Research Council, the Swedish Heart and Lung Foundation, the Medical Faculty of Lund University, Skåne University Hospital, and the Emhild Lundströms Research Foundation. T.J.W. and R.E.G. were supported by National Institutes of Health NO1-HC-25195 and R01-DK-HL081572, the Leducq Foundation, and the American Heart Association.

The funding organizations had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation or approval of the manuscript.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

References


