

A survey of proteins encoded by non-synonymous single nucleotide polymorphisms reveals a significant fraction with altered stability and activity

Abdellah ALLALI-HASSANI*¹, Gregory A. WASNEY*¹, Irene CHAU*¹, Bum Soo HONG*, Guillermo SENISTERRA*, Peter LOPPNAU*, Zhen SHI†, John MOULT†, Aled M. EDWARDS*, Cheryl H. ARROWSMITH*, Hee Won PARK*‡¹, Matthieu SCHAPIRA*‡¹ and Masoud VEDADI*²

*Structural Genomics Consortium, University of Toronto, 101 College Street, Room 839, MaRS Centre, South Tower, Toronto, ON, Canada M5G 1L7, †Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, MD 20850, U.S.A., and ‡Department of Pharmacology, University of Toronto, 101 College Street, Room 839, MaRS Centre, South Tower, Toronto, ON, Canada M5G 1L7

On average, each human gene has approximately four SNPs (single nucleotide polymorphisms) in the coding region, half of which are nsSNPs (non-synonymous SNPs) or missense SNPs. Current attention is focused on those that are known to perturb function and are strongly linked to disease. However, the vast majority of SNPs have not been investigated for the possibility of causing disease. We set out to assess the fraction of nsSNPs that encode proteins that have altered stability and activity, for this class of variants would be candidates to perturb cellular function. We tested the thermostability and, where possible, the catalytic activity for the most common variant (wild-type) and minor variants (total of 46 SNPs) for 16 human enzymes for which the three-dimensional structures were known. There were significant differences in the stability of almost half of the vari-

ants (48%) compared with their wild-type counterparts. The catalytic efficiency of approx. 14 variants was significantly altered, including several variants of human PKM2 (pyruvate kinase muscle 2). Two PKM2 variants, S437Y and E28K, also exhibited changes in their allosteric regulation compared with the wild-type enzyme. The high proportion of nsSNPs that affect protein stability and function, albeit subtly, underscores the need for experimental analysis of the diverse human proteome.

Key words: human arginine methyltransferase 1-like protein (HRMT1L3), human pyruvate kinase isoform 2, human sirtuin 5 isoform 1 (human SIRT5), protein stability, single nucleotide polymorphism (SNP).

INTRODUCTION

SNPs (single nucleotide polymorphisms) are the most common source of genetic variation in humans, accounting for approx. 90% of sequence differences [1–6]. Recent high-throughput genotyping technologies have enabled the measurement of SNP allele frequency differences between case and control populations on a genome-wide scale and have provided a wealth of information on genetic variation in the human population [7]. Public databases describe millions of human SNPs and frequency data for hundreds of thousands of these variants (http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi) [2,8–10].

One of the challenges in the analysis of human genetic variation is to distinguish functional from non-functional variants. The most straightforward instances are those in which the variants are both associated with a protein-coding gene and with a genetic disease. These may be synonymous variants, which encode the identical amino acid but regulate splicing or translation, or non-synonymous variants affecting protein sequence and consequently protein function or stability. Computational analysis of disease-causing variants predicts that approx. 75% of the mutations destabilize the protein, and only 7% directly affect biochemical

function, suggesting that change in protein stability is the most common mechanism underlying monogenic disease [11–13].

Disease-causing variants comprise the minority of SNPs in the human proteome. The more commonly found protein-associated SNPs have no obvious phenotype. This assertion is based not only on the lack of a phenotypic change, but also from predictions of functional effects using the 3D (three-dimensional) structure of the proteins as a guide (see the review by Teng et al. [14] and also SNPs3D [15], PolyPhen [16,17], TopoSNP [18], ModSNP [19], LS-SNP [20], SNPeff [21], MutDB[22], SNP@Domain [23], Snap [24], StSNP [25] and Bongo [26]). However, although most nsSNPs (non-synonymous SNPs) are thought to have no functional consequences, computational analysis predicts that a significant portion – up to 30% – are substantially less stable than the most common variant. If even a fraction of these variants did indeed affect protein function, then it is possible that they may contribute to the heterogeneity of the human population. On the other hand, some nsSNPs could affect protein function by affecting protein dynamics without changing protein stability [14].

To explore the frequency with which nsSNPs affect protein stability and activity, we selected 46 variants of 16 different human proteins for which the 3D structures are known, purified

Abbreviations used: 3D, three-dimensional; ARF4, ADP-ribosylation factor 4; DSF, differential scanning fluorimetry; DSLS, differential static light scattering; DTT, dithiothreitol; FBP, fructose 1,6-bisphosphate; HRMT1L3, human arginine methyltransferase 1-like protein; INMT, indolethylamine *N*-methyltransferase; ITD, isothermal denaturation; MWCO, molecular-mass cut-off; Ni-NTA, Ni²⁺-nitrilotriacetate; nsSNP, non-synonymous single nucleotide polymorphism; PAP, 3'-phosphoadenosine 5'-phosphate; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PEP, phosphoenolpyruvate; PK, pyruvate kinase; PKM2, pyruvate kinase muscle 2; SAHH, S-adenosylhomocysteine hydrolase; SIRT5, sirtuin 5 isoform 1; SNP, single nucleotide polymorphism; SULT, sulfotransferase; T_{agg} , aggregation temperature.

¹ These authors contributed equally to this work.

² To whom correspondence should be addressed (email mvedadi@uhnres.utoronto.ca).

them and screened them for stability by DLS (differential static light scattering), and where possible we determined their catalytic efficiency. DLS is an aggregation-based method for assessing protein stability. In this method, the point of inflection of a denaturation curve, termed T_{agg} (aggregation temperature), is determined by detecting scattered light from the denatured and aggregated protein. DLS and DSF (differential scanning fluorimetry) have been compared extensively and shown to produce similar results for the majority of proteins [27]. However, fluorescent background introduced by some proteins limits the use of DSF, whereas DLS is not sensitive to protein hydrophobicity [28,29]. Thermostability of globular proteins can be evaluated and discussed in terms of thermodynamic stability and kinetic stability when the protein denatures reversibly or irreversibly respectively [30–33]. DLS is based on the assumption that most proteins denature irreversibly and form aggregates during thermal denaturation. Light scattering as a measure of protein aggregation is a very sensitive technique; however, proteins that reversibly unfold and do not form aggregates are not amenable to this method of analysis [34]. The irreversible denaturation of proteins can be studied in most cases by measuring the rate of thermally induced activity loss at a set temperature, such as 37 °C, similar to *in vivo* conditions.

Among the selected variants were 21 variants of human cytosolic SULTs (sulfotransferases), which catalyse the transfer of a sulfonate group from PAPS (3'-phosphoadenosine 5'-phosphosulfate) to an acceptor group of substrates including drugs, hormones and xenobiotics [35–43]. We also included 16 variants of human PKM2 [PK (pyruvate kinase) muscle 2]. PK is a key enzyme in the glycolysis pathway and catalyses the formation of pyruvate and ATP from phosphoenolpyruvate and ADP (see Supplementary Figure S1 at <http://www.BiochemJ.org/bj/424/bj4240015add.htm>) [44]. Overall, four PK isoforms are expressed in human, namely PKL, PKR, PKM1 and PKM2. The PKL and PKR isoforms are expressed in liver and erythrocytes respectively, from a single gene by use of different promoters [45]. PKM1 and PKM2 isoforms are produced from the same gene by alternative RNA splicing [46]. PKM2 is found primarily in rapidly proliferating fetal tissues and tumour cells, but is replaced by PKM1 in normal adult tissues such as muscle and brain, for which large amounts of energy are required. This protein is essential for tumour cell growth and metabolism and therefore any change in its availability, level of activity or allosteric regulation may affect tumorigenesis [47–51]. PKM2 activity is allosterically activated by FBP (fructose 1,6-bisphosphate). A recent report showed that the activity of PKM2 can be regulated by tyrosine kinase signalling pathways via binding to a phosphotyrosine-containing peptide [50]. Binding of phosphotyrosine peptides to PKM2 results in the release of FBP and subsequent inhibition of enzymatic activity that lead to a build-up of glycolytic intermediates, which in turn can be used for fatty acid and nucleic acid synthesis, promoting proliferation in PKM2-expressing cancer cells. However, once FBP is released, the concentration of FBP at the time determines whether PKM2 goes into a low-activity state or rebinds FBP and is re-activated [50]. In this model, PKM2 only has the ability to undergo dynamic regulation by FBP if a tyrosine kinase pathway is activated (Supplementary Figure S1). Our results show that a great number of PKM2 SNPs affect protein availability and function. We characterized these PKM2 variants in more detail to understand the effects of human SNPs on allosteric activity of the enzyme, an important characteristic of PKM2 which is critical for tumour cell growth. Our results show that PKM2 S437Y and E28K variants exhibit changes in their allosteric regulation compared with the wild-type enzyme.

MATERIALS AND METHODS

Cloning

All genes were cloned by PCR amplification of cDNA templates of the wild-type genes using Pfu Ultra II (Stratagene). The PCR primers added 15 bp overhangs that were homologous with the ends of BseRI linearized pET28a-LIC. For site-directed mutagenesis, the genes were PCR-amplified as two or more fragments using mutagenic primers which spanned the altered codons and added 20 bp of homologous overhanging sequence between sequential fragments of the genes. PCR products were directionally assembled into pET28a-LIC in a single reaction using the In-Fusion cloning enzyme (Clontech). All genes were cloned into a modified pET vector, pET28a-LIC (GenBank® EF442785), and all mutations were confirmed by sequencing.

Expression and purification

The SULT wild-type and SNP constructs were transformed into *Escherichia coli* strain BL21-CodonPlus (DE3)-RIPL (Stratagene; catalogue number 230280), and the cells were grown in 2–8 litres of Terrific Broth at 37 °C with 50 µg/ml kanamycin and 30 µg/ml chloramphenicol to a D_{600} of ~1. Protein expression was induced by the addition of 0.5 mM isopropyl β-D-thiogalactoside with shaking overnight at 15 °C. Cells were harvested by centrifugation, resuspended in binding buffer [50 mM Hepes, 500 mM NaCl, 5 mM imidazole and 5% (v/v) glycerol, pH 7.5] with protease inhibitors (1 mM benzamidine HCl and 1 mM PMSF), frozen in liquid nitrogen, and stored at –80 °C. The thawed cells were lysed by incubation with 22.5 units/ml of benzonase and 0.5% CHAPS for 30 min at 4 °C, followed by sonication, in the presence of protease inhibitors, at a frequency of 8.5 for three cycles of 10 s on and 10 s off using Sonicator 3000 (Misonix; model S3000–001). The lysate was then clarified by centrifugation, and the supernatant was loaded on a DE52 column (Whatman catalogue number 4057200) pre-equilibrated with binding buffer. The flow-through was loaded on a Ni-NTA (Ni²⁺-nitrilotriacetate) Superflow column (Qiagen catalogue number 30450) pre-equilibrated with binding buffer. The column was washed with wash buffer (50 mM Hepes, 500 mM NaCl, 30 mM imidazole and 5% glycerol, pH 7.5), and the protein was eluted with elution buffer (50 mM Hepes, 500 mM NaCl, 250 mM imidazole and 5% glycerol, pH 7.5). EDTA (1 mM), 1 mM DTT (dithiothreitol), 2.5 mM CaCl₂ and thrombin (1 µg/mg of protein being cleaved) (Sigma catalogue number T6884–1KU) were added to the eluted protein, which was then injected into Slide-A-Lyzer Dialysis Cassettes [Pierce catalogue number 66110, 3500 MWCO (molecular-mass cut-off)]. Dialysis was carried out in 10 mM Hepes, 500 mM NaCl and 2.5 mM CaCl₂, pH 7.5, overnight at 4 °C. After dialysis, the protein was loaded on a Ni-NTA Superflow column pre-equilibrated with binding buffer. The flow-through was dialysed in 10 mM Hepes, 500 mM NaCl (pH 7.5) for 2 × 45 min at 4 °C. Protein concentration was then carried out by centrifugation using Amicon Ultra-15 Centrifugal Filter Devices (Millipore; catalogue number UFC-901024; 10000 MWCO), and the concentration of protein was determined by measuring A_{280} using Cary 50 Bio UV-Visible Spectrophotometer (Varian). Protein mass and purity was verified by MS (Agilent Technologies) and SDS/PAGE analysis.

DLS

DLS was performed as previously described in [27], in a 50 µl volume with a final concentration of 1 mM of compound per well, in 384-well plates (Nunc). The concentration of protein was the

same for all wells at 0.4 mg/ml. Ligand binding was detected by monitoring the increase in thermostability of proteins in the presence of ligands. Protein thermostability at pH 7.5 was studied using StarGazer technology, which monitors protein stability by its aggregation properties. Protein samples were heated from 27 to 80 °C at a rate of 1 °C/min in clear-bottom 384-well plates in 50 μ l of 100 mM Hepes (pH 7.5) and 150 mM NaCl. Protein aggregation was monitored by capturing images of scattered light every 30 s with a CCD (charge-coupled device) camera. The pixel intensities in a pre-selected region of each well were integrated to generate a value representative of the total amount of scattered light in that region. These total intensities were then plotted against temperature for each sample well and fit to the Boltzman equation by non-linear regression. The point of inflection of each 'denaturation' curve was identified as T_{agg} . Change in stability of a protein is shown as ΔT_{agg} .

Isothermal denaturation

ITD (isothermal denaturation) measurements were performed in parallel at 37 °C in 384-well plates using a fluorescence plate reader (FluoDia T70, Photon Technology International), with excitation at 465 nm and emission at 590 nm [52]. The measurement for each protein was repeated six times within the plate. All the fluorescence measurements were performed in 20 μ l of 0.1 M Hepes buffer, pH 7.5, containing 0.15 M NaCl, 5 times the final concentration of Sypro[®] Orange and 0.1 mg/ml of protein. The target temperature of 37 °C was reached 60 s after the plate had been placed in the instrument and the software started. Fluorescence data points were collected every 60 s for 2–4 h. The ITD curves were fit to a double-exponential function of the form:

$$Y = y_0 + A_1(1 - e^{-k_1t}) + A_2(1 - e^{-k_2t})$$

where k_1 and k_2 are two first-order rates of denaturation. In all cases, the rate constant k_1 was used to calculate the half lives ($t_{1/2}$) of the wild-type proteins and the SNPs.

Activity assays

Sulfotransferase activity assays

Enzyme assays were performed using a HPLC-based method previously described by Allali-Hassani et al. [53]. SULTs at 10–200 μ g/ml were assayed in the presence of 0.1 mM PAPS and different concentrations of each substrate in 100 mM Hepes, pH 7.5, by incubating the reaction mixture at 37 °C for a period of time from 15 to 120 min depending on how quickly PAPS was converted into PAP (3'-phosphoadenosine 5'-phosphate). The K_m values for characterized sulfotransferases are in the range of nanomolar to millimolar concentrations [35,43], with a significant variation in catalytic efficiency and substrate specificity. On the basis of these observations and considering possible substrate inhibition [36,37], we tested all sulfotransferases at substrate concentrations of 10, 25 and 100 μ M. The reactions were stopped by adding 2 vol. of urea (final concn. 5.3 M), and the mixture was filtered through a 5-kDa MWCO Amicon Ultrafree-MC filter (Millipore, catalogue number MAUF01010) to remove the protein. The ratio of PAP and PAPS was determined after separating them on HPLC using a 4.5 mm-internal diameter \times 50 mm-long WP QUAT, a strong ion-exchange column (J.T. Baker), using a gradient of triethylamine bicarbonate from 20 to 500 mM applied at 2 ml/min for 7 min. The progress of the reaction was monitored by reading the A_{259} , and the amount

of PAP produced was determined by integration of the resolved peaks using the HPLC software (Waters, <http://www.waters.com/>).

PK activity assays

Activity of the wild-type PK and its SNPs were assayed using a coupled reaction with lactic dehydrogenase [54]. Enzymatic reactions was started by adding 50 ng of enzyme (0.5 μ g/ml) to 50 μ l of reaction mixture [50 mM Hepes, pH 7.5, 50 mM KCl, 200 μ M NADH, 20 mM MgCl₂, 5 mM ADP, 7 units of lactic dehydrogenase and different concentrations of PEP (phosphoenolpyruvate; Sigma)]. The reaction progress was monitored for 3 min by following the decrease in A_{340} using a Synergy 2 microplate reader from BioTek. In the absence of FBP, wild-type PKM2 exhibited classic Michaelis–Menten kinetics, possibly because the protein bound to bacterial FBP during expression in *E. coli* [50]. K_m and V_{max} values were obtained by fitting the data to a hyperbolic function using Sigma Plot 9.

Arginine methyltransferase assay

The arginine methyltransferase assay was performed using a coupled assay [55]. In this assay, SAHH (*S*-adenosylhomocysteine hydrolase) and adenosine deaminase convert the methyltransferase reaction product *S*-adenosylhomocysteine into homocysteine and inosine. Homocysteine can be quantified in the presence of ThioGlo 1 (excitation at 384 nm; emission at 513 nm). The substrate peptide used in this assay was the first 24 residues of histone 4 (SGRGKGGKGLGKGGAKRHRKVLRD). Final protein concentration was 0.277 μ M. The SAHH clone was provided by Dr R. C. Trievel (Department of Biological Chemistry, University of Michigan, Ann Arbor, MI, U.S.A.).

Deacetylase assay for SIRT5 (sirtuin 5 isoform 1)

NAD⁺-dependent deacetylase activity of SIRT5 was assayed using the SIRT1 fluorimetric Kit (AK-555) from BIOMOL, monitoring the increase in fluorescence of the fluorophore (excitation: 360 nm; emission: 460 nm) upon deacetylation of the substrate. In this reaction, NAD⁺ is converted to nicotinamide and *O*-acetyl-ADP-ribose. Assays were performed at different concentrations of the substrate, KI-177 peptide [Arg-His-Lys-acetyl-Lys-methylcoumanlylamide], from 15 μ M to 4 mM in the presence of 5 mM NAD⁺. Final protein concentration was 16 μ M.

Fluorescence polarization

Fluorescence polarization assays were performed in 384-well plates, using the Synergy 2 microplate reader from BioTek. PFPKM2 peptide [GGAVDDD(PTyr)AQFANGG; where PTyr is phosphotyrosine] was synthesized, N-terminally labelled with fluorescein and purified by Tufts University Core Services (Boston, MA, U.S.A.). Binding assays were performed in a 10 μ l volume at a constant labelled peptide concentration (30 nM), by titrating PKM2 and SNPs (at concentrations ranging from low to high micromolar) into 50 mM Hepes, pH 7.5, 50 mM KCl, 0.01 % Tween-20 and 1 mM DTT. A maximum fluorescence polarization of approx. 200–280 mP (milli-polarization level) was detected upon binding of PKM2 and its different SNPs to the labelled peptide, whereas typically approx. 90–100 mP was detected for free labelled peptide. To determine K_d values, the data were fitted to a hyperbolic function using Sigma Plot software.

Table 1 Assessing the thermostability of nsSNPs and wild-type proteins by DSLS in comparison with the computational data taken from the SNPs 3D database

The differences in T_{agg} of the SNPs and their wild-type proteins are shown as ΔT_{agg} . All values for T_{agg} and ΔT_{agg} are in °C. Any change in T_{agg} below 2 °C is considered within experimental error. Negative and positive values indicate destabilization and stabilization compared with wild-type respectively. Allele frequencies in this Table are the highest frequencies reported for these SNPs in at least one human population. For the computational methods, a positive score indicates a variant classified as not having a significantly deleterious effect on protein function *in vivo*, and a negative score indicates a variant with a deleterious effect. The larger the value, the more confident the assignment is. Benchmarking has shown accuracy to be significantly higher for scores > |0.5| and also to be significantly higher when both computational methods agree. The methods were trained on monogenic disease data where destabilization effects are 2 kcal/mol or larger, and comparison with change in melting temperature is therefore approximate [12,13]. Nonetheless, agreement between experiment and calculation is high for the cases where the two computational methods agree. There are a number of cases where the computational methods disagree, not providing a clear assignment. NP, ND and NA stand for not purified, not determined and not applicable respectively. dbSNP, SNP database.

Protein	dbSNP	Highest allele frequency	T_{agg}	ΔT_{agg}	Structural mapping	Scores obtained from SNPs3D	
						Sequence analysis	Structural analysis
INMT	NA	NA	50.2 ± 0.1	NA	NA	NA	NA
INMT_F254C	rs4720015	0.922	44.8 ± 0.1	-5.4	Buried	-1.82	-1.02
HRMT1L3	NA	NA	54.8 ± 0.3	NA	NA	NA	NA
HRMT1L3_L440V	rs3758805	0.136	51.6 ± 0.1	-3.2	Buried	-0.48	0.95
HRMT1L3_N508S	rs6483700	0.178	54.6 ± 0.2	-0.2	Exposed	0.99	0.4
HRMT1L3_S470C	rs11025585	< 0.003	49.0 ± 0.2	-5.8	Exposed	-2.36	0.91
SRM	NA	NA	53.9 ± 0.2	NA	NA	NA	NA
SRM_L149V	rs1049932	< 0.003	56.4 ± 0.2	2.6	Buried	1.3	0.76
SIRT5	NA	NA	51.8 ± 0.3	NA	NA	NA	NA
SIRT5_F285L	rs9464003	0.058	53.1 ± 0.1	1.3	Buried flexible loop	-0.1	-0.99
GAMT	NA	NA	53.2 ± 0.3	NA	NA	NA	NA
GAMT_T209M	rs17851582	0.053	52.4 ± 0.1	-0.8	Partially buried	1.33	1.38
ARF4	NA	NA	48.5 ± 0.1	NA	NA	NA	NA
ARF4_V68A	rs11550597	0.5	45.1 ± 0.3	-3.4	Buried	-0.61	-0.31
PPIC	NA	NA	50.9 ± 0.1	NA	NA	NA	NA
PPIC_N190S	rs451195	0.4	51.0 ± 0.2	0.1	Exposed	1.89	0.67
SULT1A1	NA	NA	53.1 ± 0.8	NA	NA	NA	NA
SULT1A1_E151Q	rs1042011	0.008	52.4 ± 0.9	-0.7	Exposed	1.22	0.99
SULT1A1_E151D	rs1042014	0.867	53.7 ± 1.1	0.6	Exposed	1.91	1.18
SULT1A1_R213H	rs9282861	0.243	51.9 ± 1.6	-1.1	Exposed	0.64	0.71
SULT1A1_S290T	rs1062482	Not available	54.3 ± 1.6	1.2	Exposed	2.22	1.14
SULT1A3	NA	NA	48.4 ± 1.1	NA	NA	NA	NA
SULT1A3_R213H	rs1042028	0.11	47.8 ± 0.2	-0.6	Exposed	0.63	1.13
SULT1A3_M220V	rs2230714	0.525	50.9 ± 0.7	2.4	Buried	0.72	1.23
SULT1A3_T242P	rs1064682	Not available	47.3 ± 1.5	-1.2	Catalytic loop	1.58	-0.53
SULT2A1	NA	NA	50.7 ± 1.1	NA	NA	NA	NA
SULT2A1_A63P	rs11569681	0.065	52.6 ± 0.6	1.9	Partially buried	2.39	-0.52
SULT2A1_K227E	rs11569680	Not available	51.3 ± 0.4	0.6	Catalytic loop	1.15	1.15
SULT2A1_A261T	rs11569679	0.095	48.8 ± 1.2	-2.0	Exposed	-1.26	0.92
SULT4A1	NA	NA	64.3 ± 1.8	NA	NA	NA	NA
SULT4A1_I108M	rs11542809	Not available	62.2 ± 2.1	-2.2	Buried	1.66	1.31
SULT4A1_V88F	rs11542810	Not available	NP	NA	Partially buried	1.93	-0.81
SULT4A1_N239S	rs17852026	Not available	NP	NA	Partially buried	-1.44	1.26
SULT4A1_R155Q	rs11542811	Not available	NP	NA	NA	-0.89	1.14
SULT1C1	NA	NA	48.9 ± 0.8	NA	NA	NA	NA
SULT1C1_R282T	rs17036112	Not available	48.2 ± 0.7	-0.7	Exposed	-0.09	0.99
SULT1C1_S255A	rs17036104	0.08	48.8 ± 1.5	-0.1	Catalytic loop	0.81	0.87
SULT1C1_R73Q	rs17036058	0.004	49.3 ± 1.4	0.4	Disordered loop	0.62	-1.39
SULT1C1_Y128H	rs17036091	0.014	38.6 ± 0.5	-10.3	Buried	-0.96	-1.58
SULT1C1_V68M	rs17036055	0.004	48.5 ± 0.5	-0.4	Disordered loop	0.61	1.08
SULT1C2	NA	NA	45.8 ± 1.0	NA	NA	NA	NA
SULT1C2_D5E	rs1402467	0.942	46.1 ± 1.4	0.3	Disordered N-terminus	ND	0.43
SULT1C3	NA	NA	38.9 ± 1.2	NA	NA	NA	NA
SULT1C3_A88T	rs11903659	0.069	39.4 ± 0.8	0.6	Disordered loop	-2.4	0.8
SULT1C3_M194T	rs6722745	0.911	41.4 ± 1.4	2.6	Exposed	1.3	0.9
SULT1C3_E165K	rs6716965	< 0.003	41.9 ± 1.4	3.0	Exposed	0.7	0.7
SULT1C3_G179R	rs2219078	0.83	NP	NA	Exposed	-1.88	-0.11
SULT1C3_Y148C	rs17035911	0.078	NP	NA	Buried, close to catalytic site	-0.83	-0.01
SULT1E1	NA	NA	48.2 ± 1.0	NA	NA	NA	NA
SULT1E1_P253H	rs11569712	0.004	46.9 ± 0.4	-1.4	Catalytic loop	1.06	0.59
PKM2_WT	NA	NA	51.8 ± 0.9	NA	NA	NA	NA
PKM2_Q310P	rs11558370	< 0.003	40.7 ± 0.2	-11.1	Buried, a polar surrounding	-3.16	-0.71
PKM2_S437Y	rs59430203	Not available	43.3 ± 0.4	-8.5	Part of the FBP binding site	ND	ND
PKM2_E28K	rs11558360	Not available	45.2 ± 1.0	-6.6	salt bridge	0.08	-0.43
PKM2_V292L	rs11558353	Not available	46.3 ± 0.3	-5.5	Hydrophobic core	-1.45	0.33
PKM2_C31F	rs11558375	< 0.003	47.2 ± 0.5	-4.6	Partly exposed	-0.52	-0.96
PKM2_R339P	rs2959910	< 0.003	47.5 ± 0.5	-4.3	Exposed loop	-3.08	-0.45
PKM2_V71G	rs11558350	< 0.003	47.6 ± 0.6	-4.2	Hydrophobic core	-2.14	-0.64
PKM2_V490L	rs11558358	Not available	49.0 ± 0.9	-2.8	Partly buried, hydrophobic	ND	ND

Table 1 Contd.

Protein	dbSNP	Highest allele frequency	T_{agg}	ΔT_{agg}	Structural mapping	Scores obtained from SNPs3D	
						Sequence analysis	Structural analysis
PKM2_E275D	rs11558364	Not available	49.1 ± 0.8	−2.7	Exposed	ND	ND
PKM2_K186N	rs11558351	Not available	50.0 ± 0.6	−1.8	Exposed	ND	ND
PKM2_G200C	rs11558354	Not available	50.7 ± 0.7	−1.1	Exposed	−0.57	−0.51
PKM2_R279S	rs11558369	Not available	51.3 ± 0.7	−0.5	Exposed, salt bridge	ND	ND
PKM2_E59D	rs11558366	Not available	53.3 ± 0.6	1.5	Exposed	ND	ND
PKM2_Q16H	rs11558365	Not available	53.7 ± 0.4	1.9	Exposed	ND	−0.77
PKM2_T45I	rs11558355	Not available	54.0 ± 0.6	2.2	Buried	ND	ND
PKM2_A54V	rs11558363	Not available	54.0 ± 0.9	2.2	Part of an exposed loop	ND	ND
PKM2_Q235Ter	rs11558352	< 0.003	NP	NA	NA	ND	ND
PKM2_R92P	rs11558357	< 0.003	NP	NA	Partially buried	−2.36	−0.1
PKM2_G46D	rs11558377	Not available	NP	NA	Buried	ND	ND
PKM2_G204V	rs17853396	Not available	NP	NA	Exposed	−0.92	−1.1
PKM2_V176E	rs55738887	0.9	NP	NA	Buried	ND	ND

Displacement experiments

Displacement of FPPKM2 peptide by FBP was carried out using the fluorescence polarization signal obtained upon peptide binding to PKM2 protein. Each protein (100 μ M) was incubated with 30 nM peptide and different concentrations of FBP were added (from 2 μ M to 2 mM). Displacement of peptide was followed by a decrease in the FP signal. To determine the displacement constant ($K_{displacement}$), the data were fitted to a hyperbolic function using Sigma Plot software.

RESULTS AND DISCUSSION

As a result of recent improvements in high-throughput genotyping technologies, allele frequency for tens of thousands of SNPs have been made available [7]. Many variants have been identified that affect human physiology through genome-wide association studies using large cohorts. We selected 16 human enzymes with known biochemical activities and 3D structures in order to assess the frequency and mechanism by which SNPs affect the biochemical function of the proteins. For each protein target, all the SNPs with allele frequencies > 5% in at least one ethnic population were selected (19 SNPs in total), as well as 37 additional lower-frequency alleles for a total of 56 SNPs. Importantly, prior literature knowledge or medical relevance was not a factor in selecting these SNPs. We subcloned the recombinant SNPs, which included 26 variants of eight different SULTs, 21 variants of PKM2 and nine variants of seven additional enzymes [INMT (indolethylamine *N*-methyltransferase), HRMT1L3 (human arginine methyltransferase 1-like protein), SRM (spermidine synthase), SIRT5, GAMT (guanidinoacetate *N*-methyltransferase isoform b), ARF4 (ADP-ribosylation factor 4) and PPIC (peptidylprolyl isomerase C)]. A total of 46 of these SNPs as well as their 'wild-type' proteins were purified to > 95% purity and analysed for thermal stability using DSLS (Table 1) [27–29,34]. Ten SNPs could not be sufficiently purified under the same conditions as the wild-type protein, owing to poor expression levels, aggregation and/or precipitation, suggesting that these proteins may be unstable relative to the wild-type protein (Table 1).

Many common nsSNPs encode proteins with altered stability

The stability of almost half (48%) of the purified variants was different from the respective wild-type protein, with 16 variants being less stable and six being more stable than their wild-type

proteins (Figure 1 and Table 1). A change in midpoint of the DSLS melting curve (ΔT_{agg}) of $\geq 2^\circ\text{C}$ was considered significant, with changes in stability as large as 11°C . The results for the proteins that exhibited a change in stability by DSLS were confirmed using CD (results not shown). Overall, the stability of almost half of the SNPs was altered, with 35% significantly less, and 13% significantly more, stable (Figure 1). If we consider the ten SNPs that could not be purified, the total fraction of SNPs with altered stability may be even greater. The magnitudes of the changes in stability were not related to the frequency of the SNP alleles, or to the enzyme class to which they belonged.

Predictive computational analysis of human SNPs also suggests that a change in protein stability is the major effect of nsSNPs [12,13,15,16]. In view of the many thousands of nsSNPs in the human population, it is important to develop computational methods for assessing *in vivo* impact. Although such methods will not substitute for thorough experimental study, the analyses should be useful for obtaining a statistical picture of the extent to which these SNPs influence phenotype and for prioritizing SNPs for experimental study. Comparison of the current experimental results (Table 1) with two computational methods of SNP analysis [12,13] (<http://www.snps3d.org>), serves as a *bona fide* test of the computational methods. Such a comparison (Figure 2 and Table 1), using a two-degree threshold, shows that all but one of the six cases where both methods predict a stability impact and all but one of the 15 cases where both methods predict non-deleterious impact are correct. However, there are a number of cases where the predictions from the two methods differ. Thus, on this small sample, when the two computational methods agree, prediction accuracy is high [13].

Altered protein stability changes the half-life of unfolding

Association of loss of *in vivo* activity due to SNPs with protein destabilization has previously been reported [12,56–58]. In order to test if the change in stability of SNPs could affect the half-life of unfolding, we employed isothermal denaturation at 37°C to measure the half-lives of those SNPs with altered stabilities [28,52]. The half-lives of all tested mutants were altered and the changes correlated with the thermostability for tested variants (Figure 3, Table 1). Furthermore, it is also possible that a variant without a significant difference in T_{agg} or T_m may also have a different half-life at 37°C relative to wild-type [29,52]. This suggests that, at physiological temperatures, these and possibly some other SNPs may have lower functional levels. To investigate the effect that a decrease in protein half-life has on the level of enzyme activity

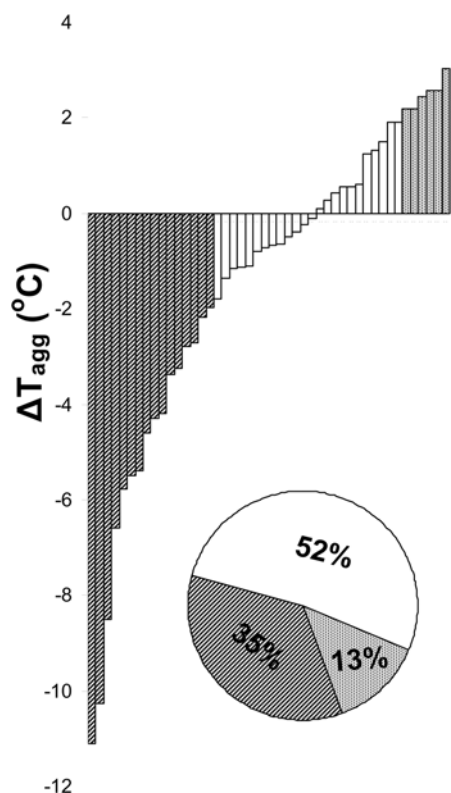


Figure 1 Distribution of the change in protein stability for 46 nsSNPs

The distribution of ΔT_{agg} , the difference between the T_{agg} of each SNP and its wild-type counterpart ($\Delta T_{agg} = T_{agg} \text{ SNP} - T_{agg} \text{ wt}$), for 46 nsSNPs is plotted; the corresponding Pie chart shows the distribution of the effect. White areas indicate that the difference between T_{agg} of the SNP and its wild-type is not significant ($< 2^\circ\text{C}$). Negative and positive values indicate destabilization and stabilization effects of SNPs respectively. Proteins are arranged from the most destabilized (largest negative ΔT_{agg} value) to the most stabilized (largest positive ΔT_{agg} value).

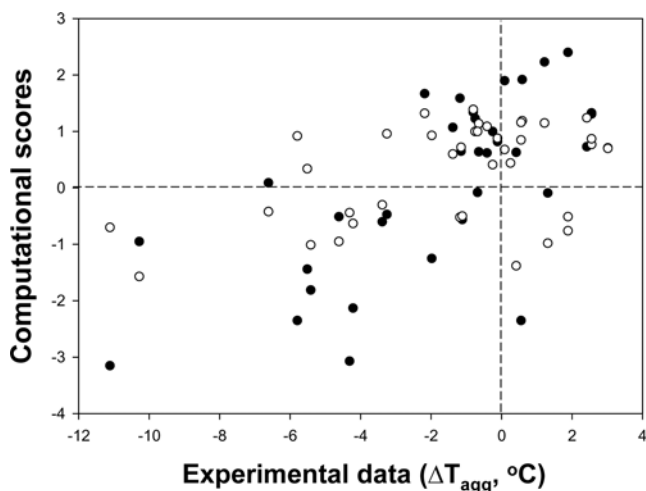


Figure 2 Comparison between experimental results and predicted effects of nsSNPs

Experimentally determined ΔT_{agg} values are plotted against scores assigned by sequence (●) and structure (○)-based predictive methods [13]. Any change in T_{agg} below 2°C is considered within experimental error. For the computational methods, a positive score indicates a variant classified as not having a significantly deleterious effect on protein function *in vivo*, and a negative score indicates a variant with a deleterious effect. The larger the value, the more confident the assignment is. Benchmarking has shown accuracy to be significantly higher for scores $> |0.5|$. Computational scores and ΔT_{agg} values are shown in Table 1.

over time, we monitored the activity of the wild-type PKM2 and PKM2_E28K after incubation at 37°C from 2 to 33 min. Our results indicated that the activity of PKM2_E28K sharply decreases during the first 2 min of incubation (up to 40%), and is almost inactive after the first 10 min. In contrast, the wild-type protein retained more than 80% of its activity after 2 min of incubation and was still partially active ($> 20\%$) even after 33 min of incubation. Thus a change in stability of the SNPs may translate to a decrease in effective enzyme activity *in vivo* (see Supplementary Figure S2 at <http://www.BiochemJ.org/bj/424/bj4240015add.htm>). SNPs with lower stability can also result in insoluble aggregates that are prominent markers of aging and associated with cellular degeneration in many age-related diseases [56,59]. Although the effects of these mutations might be buffered in the cell [56], it is also possible that these variants may be subtle modifiers of normal physiology and disease.

Catalytic activity of variants

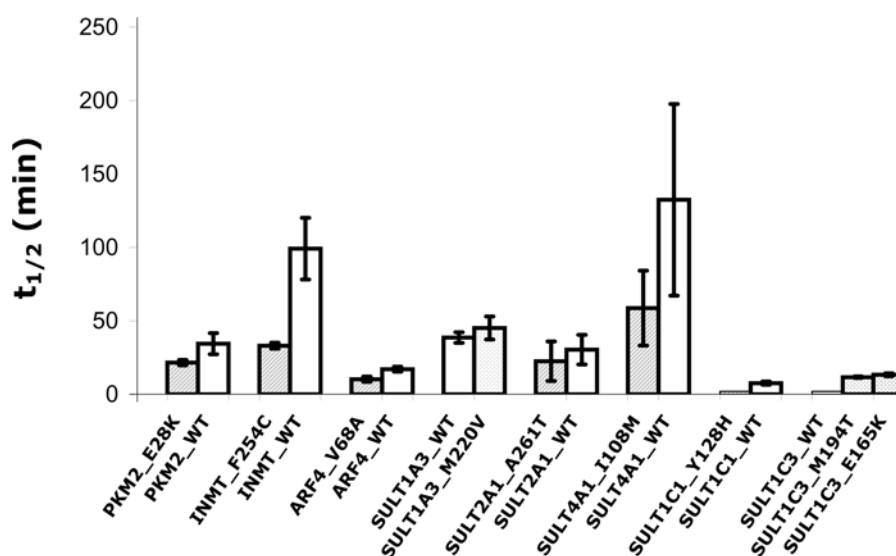
In order to assess the extent to which SNPs affect the catalytic activity of enzymes, we performed detailed enzymatic analysis on several enzyme variants, including those that did not display reduced thermal stability (Tables 2 and 3). Although a decrease in protein stability was in many cases associated with loss of activity (Figure 4), we observed cases in which thermally destabilized variants did not show altered activity, as well as cases in which the variant had altered catalytic activity, but no observable difference in stability. As discussed in the next section, these results can most often be explained through examination of the 3D protein structure.

PKM2 is a spliced isoform of PK that is expressed primarily in proliferating tissues and in cancer cells. Binding of FBP to PKM2 increases the catalytic efficiency of the enzyme allosterically by promoting the formation of the active tetramer [48]. Recently, it was reported that a phosphorylated tyrosine-containing peptide can bind to PKM2 and result in dissociation of FBP from the allosteric site, which reduces the activity of the enzyme. It was suggested that the concentration of FBP would be a determining factor in re-activating PKM2 to produce more energy [50]. Kinetic parameters were determined for wild-type PKM2 and all its variants (Table 2). Seven PKM2 variants showed more than a 40% decrease in activity. Significant increases in K_m for PEP were observed for five variants, including S437Y, T45I and V71G, with 5-, 9- and 15-fold increase respectively. The presence of 1 mM FBP significantly decreased the K_m value and consequently increased the catalytic efficiency of V71G, T45I, C31F and V292L by 12-, 6-, 3- and 2-fold respectively. S437Y was less sensitive to FBP binding, consistent with the location of Ser⁴³⁷ in the FBP-binding pocket. All 16 PKM2 SNPs bound to the phosphotyrosine peptide that was previously reported to decrease the activity of PKM2 by causing FBP dissociation. This effect is expected to be overcome at high FBP concentration, leading to recovery of full enzymatic activity. Although the K_d values for phosphotyrosine peptide binding were similar for all these SNPs, interestingly the presence of high concentrations of FBP resulted in peptide dissociation from the wild-type protein and all SNPs except for S437Y and E28K. This indicates that the level of PKM2 activity cannot be fine-tuned by relative concentrations of FBP and phosphotyrosine peptide in these two variants. The S437Y and E28K variants appear to disrupt the communication between the FBP and phosphotyrosine peptide binding sites. For S437Y, this effect may be due to an altered ability to bind FBP, because S437 is predicted to hydrogen-bond with FBP in the allosteric site of PKM2 (Figure 5).

Table 2 Kinetic analysis of PKM2 SNPs

The enzymes were assayed as described in the Materials and methods section. K_d peptide is the K_d for phosphotyrosine peptide [50], $K_{displacement}$ is the rate of phosphotyrosine peptide displacement by FBP. Kinetic parameters were calculated using SigmaPlot 9. K_m values are determined for PEP. Catalytic efficiency is the ratio of K_{cat}/K_m . ND, no displacement was detected.

Protein	No FBP added		1 mM FBP		Catalytic efficiency (%)		K_d peptide (μ M)	$K_{displacement}$ (μ M)
	K_m (μ M)	V_{max} (μ mol/min per mg)	K_m (μ M)	V_{max} (μ mol/min per mg)	No FBP added	FBP (1 mM)		
PKM2_WT	139 ± 3	100 ± 7	150 ± 24	110 ± 7	100	100	2.5	86
PKM2_Q310P	65 ± 20	4 ± 0.4	72 ± 8	4 ± 0.5	9	8	2.2	76
PKM2_S437Y	714 ± 21	105 ± 0.7	503 ± 23	91 ± 4	20	25	4.9	953
PKM2_E28K	241 ± 6	44 ± 0.3	93 ± 0.7	43 ± 0.7	25	63	3.7	ND
PKM2_V292L	347 ± 30	51 ± 2	132 ± 2	53 ± 0.7	20	55	2.3	85
PKM2_C31F	519 ± 5	113 ± 3	184 ± 4	124 ± 0.7	30	92	2.2	73
PKM2_R339P	55 ± 11	4 ± 0.1	82 ± 0	4 ± 0	10	7	4.2	79
PKM2_V490L	211 ± 4	111 ± 0.7	150 ± 3	108 ± 1	73	98	2.3	77
PKM2_E275D	185 ± 4	52 ± 0.7	109 ± 4	51 ± 0.7	39	64	2.9	73
PKM2_K186N	143 ± 2	69 ± 0.7	107 ± 0.7	68 ± 1	67	87	3.6	41
PKM2_V71G	2127 ± 26	70 ± 2	171 ± 3	52 ± 0.7	5	41	1.4	84
PKM2_G200C	170 ± 9	55 ± 1	126 ± 1	56 ± 0.3	45	61	2.2	42
PKM2_R279S	193 ± 14	97 ± 4	161 ± 1	98 ± 1	70	83	2.3	77
PKM2_E59D	260 ± 6	108 ± 0.7	208 ± 15	109 ± 1	58	71	3.6	78
PKM2_Q16H	199 ± 12	100 ± 0.3	164 ± 2	101 ± 2	70	84	3.1	69
PKM2_T45I	1273 ± 2	59 ± 1	217 ± 1	54 ± 0	6	34	1.6	65
PKM2_A54V	154 ± 9	96 ± 2	125 ± 0.7	99 ± 1	87	108	2.3	76

**Figure 3 Half-lives of SNPs and wild-type proteins**

Half-lives of some SNPs that showed stability different from that of their wild-type protein ($\geq 2^\circ\text{C}$) were determined at 37°C using isothermal denaturation as described in the Materials and methods section. Values are the average of six measurements. Wild-type (WT) proteins are shown as white bars.

The human cytosolic sulfotransferases comprise a family of 12 phase II enzymes involved in the metabolism of drugs and hormones, the bioactivation of carcinogens and the detoxification of xenobiotics [37,60]. Phenotypic changes in drug metabolism due to SNPs in SULTs contribute to inter-individual differences in drug sensitivity and are a well-studied area of pharmacogenomics [61,62]. Enzymatic activity of SULT variants were assayed using selected substrates [38]. No significant ($> 50\%$) change in overall activities of SULT variants were observed, except for SULT1A3_T242P, which showed higher activity compared with its wild-type counterpart (Table 3). Detailed kinetic analysis showed that the K_m values of SULT1A3_T242P for substrates such as dopamine, isoprenaline and tyramine increased by 7-, 5- and 21-fold respectively. The V_{max} values for these

compounds were 4–5 times higher than that of the wild-type protein. In the present study we report that SULT1C1_Y128H, SULT4A1_I108M and SULT2A1_A261T are destabilized by 10.3, 2.2 and 2°C respectively, relative to their wild-type counterparts. Although we did not see any significant changes in the activities of these proteins, their half-lives were reduced, raising the possibility that, under physiological conditions, these variants may be functionally impaired.

Two variants of HRMT1L3 were also significantly less stable than the wild-type enzyme, and their catalytic efficiencies were slightly lower than their wild-type counterpart (Table 3). Protein arginine methyltransferase has been reported to play an important role in transcriptional regulation by methylation of histone H4 at Arg³ [63]. A change in protein levels or activity of HRMT1L3 may

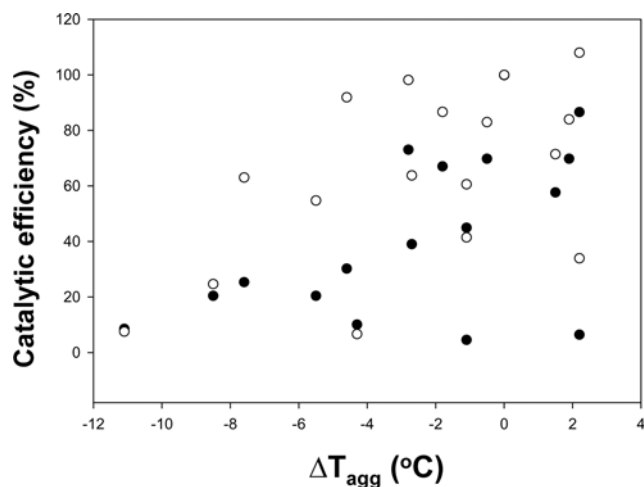


Figure 4 Correlation between the catalytic efficiency and thermostability of PKM2 variants

Catalytic efficiency of PKM2 variants with (○) and without (●) 1 mM FBP added are plotted against their ΔT_{agg} values. Catalytic efficiency values are shown in Table 2.

therefore have an epigenetic effect. The SIRT5 variant F285L, which was as stable as the wild-type SIRT5, was completely inactive (Table 3).

Structural analysis of variants

We examined the protein structural effects for each SNP reported here (Table 1). The majority of destabilizing mutations can be

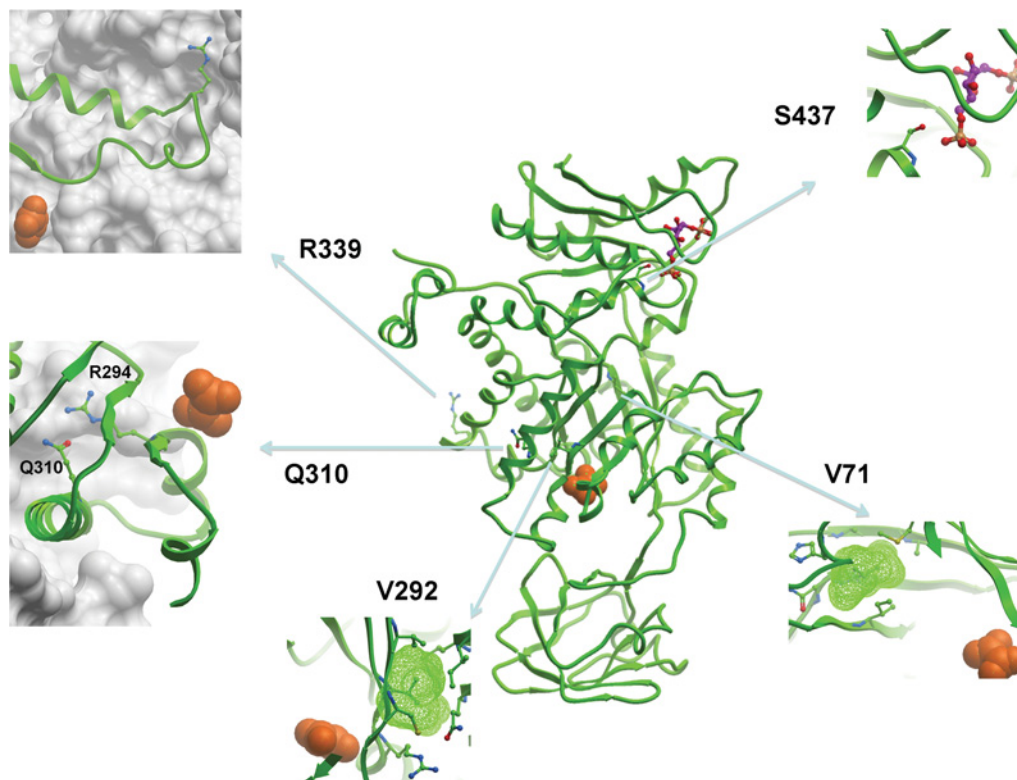


Figure 5 Structural mapping of some of the PKM2 SNPs

From the top right corner, clockwise: S437Y inhibits FBP binding, V71G and V292L destabilize core hydrophobic packing (green hydrophobic mesh), Q310P and R339P disrupt interactions with a neighbouring subunit (white surface) and induce modifications that may propagate to the active site. Colour coding: FBP, magenta; active site oxalate, orange.

Table 3 Kinetic analysis of SULT1A3, HRMT1L3 and SIRT5 SNPs

The enzymes were assayed as described in the Materials and methods section. Kinetic parameters were calculated using SigmaPlot 9. H4(1-24) is SGRGKGGKGLGKGGAKRHRKVLRD. WT, wild-type.

Enzyme	Substrate	K_m (μM)	V_{max} (nmol/min per mg)
SULT1A3 WT	Dopamine	2.2 ± 0.5	202 ± 12
	Isoprenaline	4.9 ± 0.5	162 ± 6
	Tyramine	22 ± 3	117 ± 3
SULT1A3_T242P	Dopamine	15 ± 2	868 ± 40
	Isoprenaline	27 ± 4	720 ± 36
	Tyramine	471 ± 51	665 ± 34
SULT1A3_M220V	Dopamine	2.5 ± 0.5	168 ± 8
	Isoprenaline	3.2 ± 0.3	135 ± 4
	Tyramine	29 ± 5	111 ± 5
HRMT1L3 WT	H4 (1-24)	1.6 ± 0.04	31.4 ± 0.14
HRMT1L3 L440V	H4 (1-24)	4.8 ± 0.92	19.5 ± 1.4
HRMT1L3 N508S	H4 (1-24)	1.8 ± 0.28	26.4 ± 1.7
HRMT1L3 S470C	H4 (1-24)	3.6 ± 0.07	22.1 ± 1.3
SIRT5 WT	KI-177	2300 ± 300	0.3 ± 0.02
SIRT5 F285L	KI-177	Not active	Not active

linked to disruption of the packing of buried residues in the core of the protein or to elimination of one or more hydrogen bonds. Here we will focus mainly on the structural explanation of protein stability and activity for two well-studied enzyme classes, SULTs and PKM2, in which our data reveal novel SNP activities.

PKM2 variants with altered stability or activity can be grouped into three categories; polymorphisms affecting buried residues, surface residues and those involved in binding sites or

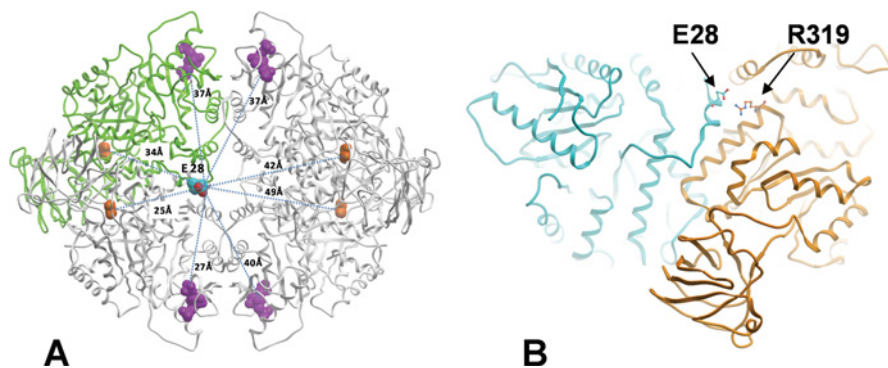


Figure 6 Structural mapping of Glu²⁸ in the PKM2 tetramer

Glu²⁸ is located at the interface between two PKM2 subunits (A) and induces significant phenotypic modifications, in spite of being remote from the allosteric and catalytic sites of the PKM2 tetramer. Glu²⁸ makes an electrostatic bridge with neighbouring Arg³¹⁹, which is disrupted in the E28K variant (B). One PKM2 subunit is shown in green. The other three are coloured grey. FBP, magenta; active site oxalate, orange.

subunit interactions. Buried residues can affect catalytic efficiency for diverse reasons: Val⁷¹ and Thr⁴⁵ are deeply buried in the hydrophobic core of PKM2 (Figure 5). Though V71G does not create steric clashes and T45I may be accommodated by local adjustment of neighbouring side-chains, both mutants have a dramatically lower affinity for PEP, suggesting that altered hydrophobic packing may destabilize the overall structure, thereby affecting PEP binding (Table 2). Interestingly, PEP affinity is rescued to near wild-type levels in the presence of 1 mM FBP, which indicates that FBP stabilizes the active conformation of the mutants. Another interesting buried residue is Gln³¹⁰, which is located on a helix inserted into the groove of another PKM2 subunit (Figure 5) and makes polar interactions with Arg²⁹⁴, a residue lining the catalytic site. Q310P is expected to break the helical topology, affect the quaternary arrangement of the tetramer, disrupt the interaction with Arg²⁹⁴, and induce conformational rearrangements at the active site (Figure 5). This is in agreement with the loss of 90% of the catalytic efficiency experimentally observed (Table 2). Val²⁹² is a buried residue located next to Ala²⁹³, which directly lines the catalytic pocket (Figure 5). The V292L mutant introduces a local steric strain that is likely to propagate to the active site, which would corroborate with the decreased affinity for PEP and lower activity. Some buried SNPs do not affect the catalytic activity of the enzyme: Val⁴⁹⁰ is buried in the vicinity of the FBP-binding site and is less stable than wild-type PKM2. However, the fact that the V490L conservative mutation hardly affects the catalytic efficiency or FBP's ability to displace the phosphorylated peptide strongly suggests that neighbouring residues can accommodate a leucine residue at position 490.

Glu²⁷⁵ is a surface residue surrounded with lysine and arginine residues, and stability or activity of E275D variant would not have been predicted to be different from that of wild-type PKM2. The observed decrease in stability and approx. 2-fold loss in catalytic efficiency may be attributable to a mild local rearrangement necessary to accommodate the shorter aspartate side chain, which may propagate to neighbouring Glu²⁷², a residue that co-ordinates a catalytic Mg²⁺ in the crystal structure of rabbit PK (PDB code 1a49). This is an example which indicates that preservation of physico-chemical properties at the site of mutation does not always result in small effects [64].

Since PKM2 is known to be regulated via quaternary structural allosteric mechanisms [48,65], residues located at the interface of two PKM2 subunits may induce non-native quaternary motions that could affect catalytic activity. Arg³³⁹ is located at the exit of a

loop that lines the catalytic site and is at the interface with another PKM2 subunit. The R339P mutation may induce quaternary restraints that directly propagate along the loop into the catalytic site, thereby critically affecting the catalytic activity of the enzyme. Cys³¹ is remote both from the catalytic site and the allosteric activation site. It is located at the interface with another PKM2 subunit, and the C31F mutant may destabilize the quaternary arrangement of PKM2, which could explain the observed decrease in affinity for PEP (Table 2). This destabilization is probably not strong, since 1 mM FBP can almost fully revert to wild-type catalytic efficiency.

The E28K variant of PKM2 is intriguing. Glu²⁸ is at the interface of two PKM2 monomers and makes electrostatic interactions with Arg³¹⁹ in a neighbouring subunit (Figure 6). Glu²⁸ is over 30 Å (1 Å = 0.1 nm) away from the closest FBP-binding site, and over 25 Å away from the closest catalytic site. Nevertheless, the E28K variant is reduced in its catalytic efficiency by 75% with no addition of FBP and even by 37% in the presence of 1 mM FBP. Moreover, unlike the wild-type enzyme, FBP, which results in a 2.5-fold gain in catalytic efficiency, is unable to displace a phosphotyrosine peptide in the E28K mutant, even though the peptide affinity is unchanged. This suggests that the E28K mutation affects allosteric regulation of PKM2 activity, possibly by affecting FBP binding. On the other hand, the S437Y mutant directly affects FBP binding (Figure 5). Ser⁴³⁷ lines the FBP-binding site and the hydroxy group of Ser⁴³⁷ is hydrogen-bonded to FBP. S437Y is predicted to preclude FBP binding, and thus prevent stabilization of the PKM2 active form. This is consistent with the observed inability of FBP to displace a phosphotyrosine peptide and with the dramatically reduced affinity of the protein for PEP. For S437Y, this is clearly related to the involvement of Ser⁴³⁷ in hydrogen-bonding with FBP in the allosteric site of PKM2 [65,66]. Genetic variations in PKM2 that results in a change in level of activity and allosteric properties of PKM2 may alter the susceptibility to cancer as well as affecting embryonic development.

Our SNP data reported here also reveal changes in activity and stability for several of the SULTs. SULT1C1_Y128H was dramatically destabilized, probably due to the elimination of a buried hydrogen bond with His¹⁷⁶, which in turn contributes to the packing of helix 7 on to the central β-sheet. The observed change in specificity of SULT1A3_T242P can also be explained through structural analysis. Thr²⁴² is in one of the flexible loops, part of which contributes to the catalytic site of this enzyme [38]. Thus our data for this SNP suggests that a rigid proline residue

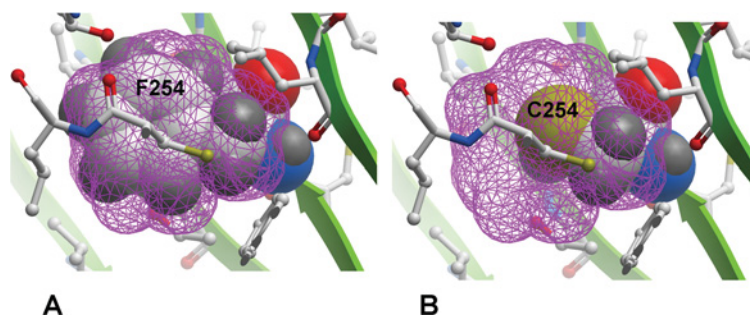


Figure 7 Proposed structural interpretation of observed destabilization of F254C

Phe²⁵⁴ is buried in the core of the INMT structure (PDB code 2a14) and fully occupies the cavity generated by surrounding residues (**A**, magenta mesh). A cysteine residue makes reduced Van der Waals contacts, which may result in decreased stability (**B**).

at this position may favour a change in conformation of the loop, resulting in a change in catalytic activity and that individuals with this SNP may metabolize substrates at high concentrations more efficiently.

Other variants

Three of the SNPs with increased stability had conservative mutations of one hydrophobic residue for another at buried positions within the protein core (Table 1). The greater stability of these SNPs may therefore be the result of improved side-chain packing within the core of the protein. INMT_F254C and ARF4_V68A are both significantly destabilized and can be classified as mutations in which a large hydrophobic residue buried in the interior of the protein is changed to a smaller hydrophobic residue (see, for example, Figure 7). The creation of 'void volumes' in the interior of proteins comes at a high energetic cost. For example, studies on 'large-to-small' substitutions in the hydrophobic core of T4-lysozyme and barnase resulted in mutants that were destabilized by several kcal/mol (1 cal \approx 4.184 J) [67–69]. ARF4 specifically recognizes the C-terminal sorting signal of rhodopsin, regulating its incorporation into specialized post-Golgi rhodopsin transport carriers targeted to the rod outer segments of retinal photoreceptors [70]. Rhodopsin C-terminal mutants that affect its interaction with ARF4 contribute to the pathology of retinitis pigmentosa, raising the possibility that the ARF4_V68A variant might be a disease-modifying allele. INMT catalyses the production of bufotenine and *N,N*-dimethyltryptamine, which are hallucinogenic dimethylated indolethylamines, from serotonin and tryptamine. A change in cellular availability of this enzyme might contribute to the pathology of schizophrenia or other psychiatric illnesses [71].

Concluding remarks

In conclusion, our studies have focused on exploring the possibility that relatively common non-synonymous protein variants might display altered thermal stability or catalytic activity. Although we were only able to test a small number of human proteins for pragmatic reasons, we observed that a remarkably large number showed altered stability and function. Although the effects of these mutations might be buffered in the cell, it is also possible that some of these alleles may be subtle modifiers of normal physiology and disease.

AUTHOR CONTRIBUTIONS

Abdellah Allali-Hassani performed all enzyme activity and fluorescence polarization assays, Gregory Wasney performed all DSLs and CD experiments and purified proteins,

Irene Chau purified proteins and cloned PKM2 SNPs, Bum Soo Hong contributed to structural analysis, Guillermo Senisterra performed ITD experiments, Peter Loppnau cloned most of the SNPs and Zhen Shi performed computational analysis. John Moulton contributed to data analysis and wrote the manuscript, Aled Edwards contributed to overall project design and wrote the manuscript, Cheryl Arrowsmith contributed to overall project design and wrote the manuscript, Hee Won Park contributed to structural analysis and wrote the manuscript, Mattieu Schapira performed structural analysis and contributed to the manuscript and Masoud Vedadi designed the project, directed the experimental efforts, analysed the data and led the preparation of the manuscript.

ACKNOWLEDGEMENTS

We thank Dr Ian Dunham for his help in identifying SNPs of our targets and Dr Trievel for providing the SAHH clone.

FUNDING

The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust.

REFERENCES

- Li, W. H. and Sadler, L. A. (1991) Low nucleotide diversity in man. *Genetics* **129**, 513–523
- International HapMap Consortium (2003) The International HapMap Project. *Nature* **426**, 789–796
- Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J. et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**, 1077–1082
- Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Shaw, N., Lane, C. R., Lim, E. P., Kalyanaraman, N. et al. (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* **22**, 231–238
- Halushka, M. K., Fan, J. B., Bentley, K., Hsie, L., Shen, N., Weder, A., Cooper, R., Lipshutz, R. and Chakravarti, A. (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genet.* **22**, 239–247
- Collins, F. S., Brooks, L. D. and Chakravarti, A. (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res.* **8**, 1229–1231
- Zaitlen, N., Kang, H. M., Eskin, E. and Halperin, E. (2007) Leveraging the HapMap correlation structure in association studies. *Am. J. Hum. Genet.* **80**, 683–691
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. and Sirotkin, K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311
- Fredman, D., Siegfried, M., Yuan, Y. P., Bork, P., Lehvaslaiho, H. and Brookes, A. J. (2002) HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources. *Nucleic Acids Res.* **30**, 387–391

- 10 Hirakawa, M., Tanaka, T., Hashimoto, Y., Kuroda, M., Takagi, T. and Nakamura, Y. (2002) JSNP: a database of common gene variations in the Japanese population. *Nucleic Acids Res.* **30**, 158–162
- 11 Wang, Z. and Moul, J. (2001) SNPs, protein structure, and disease. *Hum. Mutat.* **17**, 263–270
- 12 Yue, P., Li, Z. and Moul, J. (2005) Loss of protein structure stability as a major causative factor in monogenic disease. *J. Mol. Biol.* **353**, 459–473
- 13 Yue, P. and Moul, J. (2006) Identification and analysis of deleterious human SNPs. *J. Mol. Biol.* **356**, 1263–1274
- 14 Teng, S., Michonova-Alexova, E. and Alexov, E. (2008) Approaches and resources for prediction of the effects of non-synonymous single nucleotide polymorphism on protein function and interactions. *Curr. Pharm. Biotechnol.* **9**, 123–133
- 15 Yue, P., Melamud, E. and Moul, J. (2006) SNPs3D: candidate gene and SNP selection for association studies. *BMC Bioinformatics* **7**, 166
- 16 Sunyaev, S., Ramensky, V., Koch, I., Lathe, 3rd, W., Kondrashov, A. S. and Bork, P. (2001) Prediction of deleterious human alleles. *Hum. Mol. Genet.* **10**, 591–597
- 17 Ramensky, V., Bork, P. and Sunyaev, S. (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* **30**, 3894–3900
- 18 Stitzel, N. O., Binkowski, T. A., Tseng, Y. Y., Kasif, S. and Liang, J. (2004) topoSNP: a topographic database of non-synonymous single nucleotide polymorphisms with and without known disease association. *Nucleic Acids Res.* **32**, D520–D522
- 19 Yip, Y. L., Scheib, H., Diemand, A. V., Gattiker, A., Famiglietti, L. M., Gasteiger, E. andairoch, A. (2004) The Swiss-Prot variant page and the ModSNP database: a resource for sequence and structure information on human protein variants. *Hum. Mutat.* **23**, 464–470
- 20 Karchin, R., Diekhans, M., Kelly, L., Thomas, D. J., Pieper, U., Eswar, N., Haussler, D. and Sali, A. (2005) LS-SNP: large-scale annotation of coding non-synonymous SNPs based on multiple information sources. *Bioinformatics* **21**, 2814–2820
- 21 Reumers, J., Schymkowitz, J., Ferkinghoff-Borg, J., Stricher, F., Serrano, L. and Rousseau, F. (2005) SNPeff: a database mapping molecular phenotypic effects of human non-synonymous coding SNPs. *Nucleic Acids Res.* **33**, D527–D532
- 22 Dantzer, J., Moad, C., Heiland, R. and Mooney, S. (2005) MutDB services: interactive structural analysis of mutation data. *Nucleic Acids Res.* **33**, W311–W314
- 23 Han, A., Kang, H. J., Cho, Y., Lee, S., Kim, Y. J. and Gong, S. (2006) SNP@Domain: a web resource of single nucleotide polymorphisms (SNPs) within protein domain structures and sequences. *Nucleic Acids Res.* **34**, W642–W644
- 24 Li, S., Ma, L., Li, H., Vang, S., Hu, Y., Bolund, L. and Wang, J. (2007) Snap: an integrated SNP annotation platform. *Nucleic Acids Res.* **35**, D707–D710
- 25 Uzun, A., Leslin, C. M., Abyzov, A. and Ilyin, V. (2007) Structure SNP (StSNP): a web server for mapping and modeling nsSNPs on protein structures with linkage to metabolic pathways. *Nucleic Acids Res.* **35**, W384–W392
- 26 Cheng, T. M., Lu, Y. E., Vendruscolo, M., Lio, P. and Blundell, T. L. (2008) Prediction by graph theoretic measures of structural effects in proteins arising from non-synonymous single nucleotide polymorphisms. *PLoS Comput. Biol.* **4**, e1000135
- 27 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, Jr, P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H. et al. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15835–15840
- 28 Hong, B. S., Senisterra, G., Rabeh, W. M., Vedadi, M., Leonard, R., Zhang, Y. M., Rock, C. O., Jackowski, S. and Park, H. W. (2007) Crystal structures of human pantothenate kinases. Insights into allosteric regulation and mutations linked to a neurodegeneration disorder. *J. Biol. Chem.* **282**, 27984–27993
- 29 Senisterra, G. A. and Finerty, Jr, P. J. (2009) High throughput methods of assessing protein stability and aggregation. *Mol. bioSystems* **5**, 217–223
- 30 Baker, D. and Agard, D. A. (1994) Kinetics versus thermodynamics in protein folding. *Biochemistry* **33**, 7505–7509
- 31 Powers, E. T., Deechongkit, S. and Kelly, J. W. (2005) Backbone-backbone H-bonds make context-dependent contributions to protein folding kinetics and thermodynamics: lessons from amide-to-ester mutations. *Adv. Protein Chem.* **72**, 39–78
- 32 Blancas-Mejia, L. M., Tellez, L. A., del Pozo-Yauner, L., Becerril, B., Sanchez-Ruiz, J. M. and Fernandez-Velasco, D. A. (2009) Thermodynamic and kinetic characterization of a germ line human $\lambda 6$ light-chain protein: the relation between unfolding and fibrillogenesis. *J. Mol. Biol.* **386**, 1153–1166
- 33 Jaswal, S. S., Truhlar, S. M., Dill, K. A. and Agard, D. A. (2005) Comprehensive analysis of protein folding activation thermodynamics reveals a universal behavior violated by kinetically stable proteases. *J. Mol. Biol.* **347**, 355–366
- 34 Senisterra, G. A., Markin, E., Yamazaki, K., Hui, R., Vedadi, M. and Awrey, D. E. (2006) Screening for ligands using a generic and high-throughput light-scattering-based assay. *J. Biomol. Screen.* **11**, 940–948
- 35 Schrag, M. L., Cui, D., Rushmore, T. H., Shou, M., Ma, B. and Rodrigues, A. D. (2004) Sulfotransferase 1E1 is a low K_m isoform mediating the 3-O-sulfation of ethinyl estradiol. *Drug Metab. Dispos.* **32**, 1299–1303
- 36 Barnett, A. C., Tsvetanov, S., Gamage, N., Martin, J. L., Duggleby, R. G. and McManus, M. E. (2004) Active site mutations and substrate inhibition in human sulfotransferase 1A1 and 1A3. *J. Biol. Chem.* **279**, 18799–18805
- 37 Gamage, N., Barnett, A., Hempel, N., Duggleby, R. G., Windmill, K. F., Martin, J. L. and McManus, M. E. (2006) Human sulfotransferases and their role in chemical metabolism. *Toxicol. Sci.* **90**, 5–22
- 38 Allali-Hassani, A., Pan, P. W., Dombrowski, L., Najmanovich, R., Tempel, W., Dong, A., Loppnau, P., Martin, F., Thonton, J., Edwards, A. M. et al. (2007) Structural and chemical profiling of the human cytosolic sulfotransferases. *PLoS Biol.* **5**, e97
- 39 Liang, G., Miao, X., Zhou, Y., Tan, W. and Lin, D. (2004) A functional polymorphism in the SULT1A1 gene (G638A) is associated with risk of lung cancer in relation to tobacco smoking. *Carcinogenesis* **25**, 773–778
- 40 Pachouri, S. S., Sobti, R. C., Kaur, P., Singh, J. and Gupta, S. K. (2006) Impact of polymorphism in sulfotransferase gene on the risk of lung cancer. *Cancer Genet. Cytogenet.* **171**, 39–43
- 41 Nagar, S., Walther, S. and Blanchard, R. L. (2006) Sulfotransferase (SULT) 1A1 polymorphic variants *1, *2, and *3 are associated with altered enzymatic activity, cellular phenotype, and protein degradation. *Mol. Pharmacol.* **69**, 2084–2092
- 42 Hildebrandt, M. A., Salavaggione, O. E., Martin, Y. N., Flynn, H. C., Jalal, S., Wieben, E. D. and Weinsilboum, R. M. (2004) Human SULT1A3 pharmacogenetics: gene duplication and functional genomic studies. *Biochem. Biophys. Res. Commun.* **321**, 870–878
- 43 Brix, L. A., Barnett, A. C., Duggleby, R. G., Leggett, B. and McManus, M. E. (1999) Analysis of the substrate specificity of human sulfotransferases SULT1A1 and SULT1A3: site-directed mutagenesis and kinetic studies. *Biochemistry* **38**, 10474–10479
- 44 Fothergill-Gilmore, L. A. and Michels, P. A. (1993) Evolution of glycolysis. *Prog. Biophys. Mol. Biol.* **59**, 105–235
- 45 Noguchi, T., Yamada, K., Inoue, H., Matsuda, T. and Tanaka, T. (1987) The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters. *J. Biol. Chem.* **262**, 14366–14371
- 46 Noguchi, T., Inoue, H. and Tanaka, T. (1986) The M1- and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. *J. Biol. Chem.* **261**, 13807–13812
- 47 Imamura, K. and Tanaka, T. (1972) Multimolecular forms of pyruvate kinase from rat and other mammalian tissues. I. Electrophoretic studies. *J. Biochem. (Tokyo)* **71**, 1043–1051
- 48 Mazurek, S., Boschek, C. B., Hugo, F. and Eigenbrodt, E. (2005) Pyruvate kinase type M2 and its role in tumor growth and spreading. *Semin. Cancer Biol.* **15**, 300–308
- 49 Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., Wei, R., Fleming, M. D., Schreiber, S. L. and Cantley, L. C. (2008) The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **452**, 230–233
- 50 Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M. and Cantley, L. C. (2008) Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature* **452**, 181–186
- 51 Ferguson, E. C. and Rathmell, J. C. (2008) New roles for pyruvate kinase M2: working out the Warburg effect. *Trends Biochem. Sci.* **33**, 359–362
- 52 Senisterra, G. A., Soo Hong, B., Park, H. W. and Vedadi, M. (2008) Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. *J. Biomol. Screen.* **13**, 337–342
- 53 Allali-Hassani, A., Campbell, T. L., Ho, A., Schertzer, J. W. and Brown, E. D. (2004) Probing the active site of YjeE: a vital *Escherichia coli* protein of unknown function. *Biochem. J.* **384**, 577–584
- 54 Smith, S. B. and Freedland, R. A. (1979) Activation of pyruvate kinase by 6-phosphogluconate. *J. Biol. Chem.* **254**, 10644–10648
- 55 Collazo, E., Couture, J. F., Bulfer, S. and Trievel, R. C. (2005) A coupled fluorescent assay for histone methyltransferases. *Anal. Biochem.* **342**, 86–92
- 56 Tokuriki, N. and Tawfik, D. S. (2009) Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* **459**, 668–673
- 57 Zeldovich, K. B., Chen, P. and Shakhnovich, E. I. (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16152–16157
- 58 Tokuriki, N., Stricher, F., Serrano, L. and Tawfik, D. S. (2008) How protein stability and new functions trade off. *PLoS Comput. Biol.* **4**, e1000002
- 59 Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J. and Taddei, F. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3076–3081
- 60 Harris, R. M. and Waring, R. H. (2008) Sulfotransferase inhibition: potential impact of diet and environmental chemicals on steroid metabolism and drug detoxification. *Curr. Drug Metab.* **9**, 269–275
- 61 Hildebrandt, M. A., Carrington, D. P., Thomae, B. A., Eckloff, B. W., Schaid, D. J., Yee, V. C., Weinsilboum, R. M. and Wieben, E. D. (2007) Genetic diversity and function in the human cytosolic sulfotransferases. *Pharmacogenomics J.* **7**, 133–143

- 62 Ji, Y., Moon, I., Zlatkovic, J., Salavaggione, O. E., Thomae, B. A., Eckloff, B. W., Wieben, E. D., Schaid, D. J. and Weinshilboum, R. M. (2007) Human hydroxysteroid sulfotransferase SULT2B1 pharmacogenomics: gene sequence variation and functional genomics. *J. Pharmacol. Exp. Ther.* **322**, 529–540
- 63 Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P. and Zhang, Y. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* **293**, 853–857
- 64 Teng, S., Madej, T., Panchenko, A. and Alexov, E. (2009) Modeling effects of human single nucleotide polymorphisms on protein-protein interactions. *Biophys. J.* **96**, 2178–2188
- 65 Jurica, M. S., Mesecar, A., Heath, P. J., Shi, W., Nowak, T. and Stoddard, B. L. (1998) The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure* **6**, 195–210
- 66 Dombrackas, J. D., Santarsiero, B. D. and Mesecar, A. D. (2005) Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry* **44**, 9417–9429
- 67 Vetter, I. R., Baase, W. A., Heinz, D. W., Xiong, J. P., Snow, S. and Matthews, B. W. (1996) Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme. *Protein Sci.* **5**, 2399–2415
- 68 Lee, J., Lee, K. and Shin, S. (2000) Theoretical studies of the response of a protein structure to cavity-creating mutations. *Biophys. J.* **78**, 1665–1671
- 69 Buckle, A. M., Cramer, P. and Fersht, A. R. (1996) Structural and energetic responses to cavity-creating mutations in hydrophobic cores: observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry* **35**, 4298–4305
- 70 Deretic, D., Williams, A. H., Ransom, N., Morel, V., Hargrave, P. A. and Arendt, A. (2005) Rhodopsin C terminus, the site of mutations causing retinal disease, regulates trafficking by binding to ADP-ribosylation factor 4 (ARF4). *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3301–3306
- 71 Karkkainen, J., Forsstrom, T., Tornaues, J., Wahala, K., Kiuru, P., Honkanen, A., Stenman, U. H., Turpeinen, U. and Hesso, A. (2005) Potentially hallucinogenic 5-hydroxytryptamine receptor ligands bufotenine and dimethyltryptamine in blood and tissues. *Scand. J. Clin. Lab. Invest.* **65**, 189–199

Received 10 May 2009/18 August 2009; accepted 25 August 2009

Published as BJ Immediate Publication 25 August 2009, doi:10.1042/BJ20090723

SUPPLEMENTARY ONLINE DATA

A survey of proteins encoded by non-synonymous single nucleotide polymorphisms reveals a significant fraction with altered stability and activity

Abdellah ALLALI-HASSANI*¹, Gregory A. WASNEY*¹, Irene CHAU*¹, Bum Soo HONG*, Guillermo SENISTERRA*, Peter LOPPNAU*, Zhen SHI†, John MOULT†, Aled M. EDWARDS*, Cheryl H. ARROWSMITH*, Hee Won PARK*‡¹, Matthieu SCHAPIRA*‡¹ and Masoud VEDADI*²

*Structural Genomics Consortium, University of Toronto, 101 College Street, Room 839, MaRS Centre, South Tower, Toronto, ON, Canada M5G 1L7, †Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, MD 20850, U.S.A., and ‡Department of Pharmacology, University of Toronto, 101 College Street, Room 839, MaRS Centre, South Tower, Toronto, ON, Canada M5G 1L7

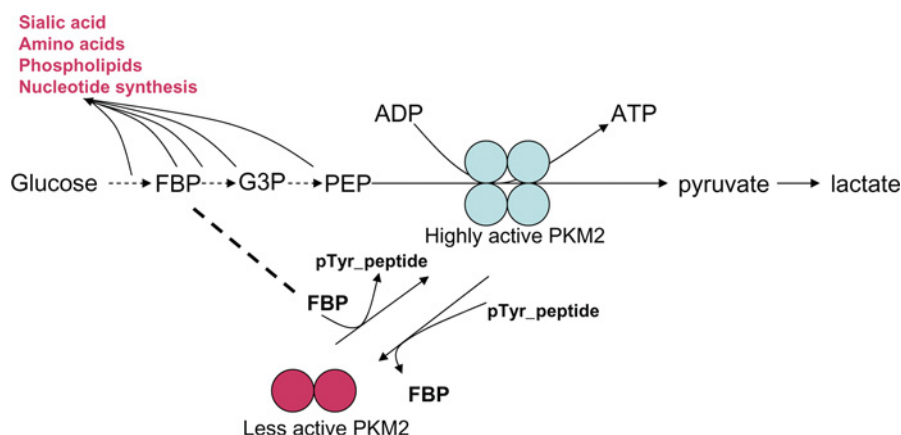


Figure S1 Activity and allosteric regulation of PKM2

See also [1–5]. G3P, glycerate 3-phosphate.

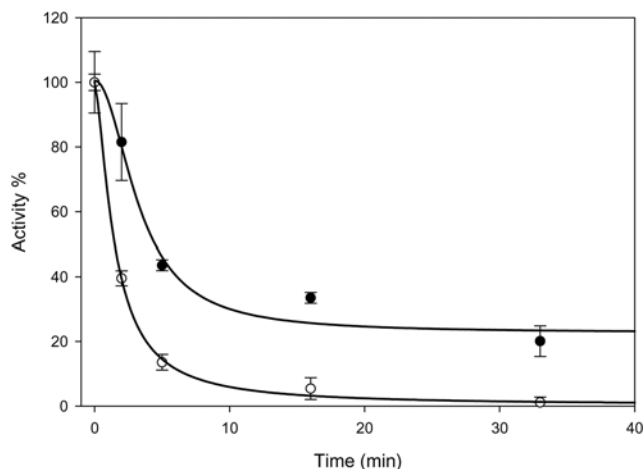


Figure S2 Effect of pre-incubation at 37 °C on the activity of wild-type PKM2 and its SNP

Activity of wild-type pyruvate kinase (●) and PKM2_E28K (○) were assayed as described in the Material and methods section of the main paper, after pre-incubation of the proteins at 37 °C for 2–33 min.

REFERENCES

- Mazurek, S., Boschek, C. B., Hugo, F. and Eigenbrodt, E. (2005) Pyruvate kinase type M2 and its role in tumor growth and spreading. *Semin. Cancer Biol.* **15**, 300–308
- Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M. and Cantley, L. C. (2008) Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature* **452**, 181–186
- Ferguson, E. C. and Rathmell, J. C. (2008) New roles for pyruvate kinase M2: working out the Warburg effect. *Trends Biochem. Sci.* **33**, 359–362
- Ashizawa, K., McPhie, P., Lin, K. H. and Cheng, S. Y. (1991) An *in vitro* novel mechanism of regulating the activity of pyruvate kinase M2 by thyroid hormone and fructose 1,6-bisphosphate. *Biochemistry* **30**, 7105–7111
- Mazurek, S., Zwerschke, W., Jansen-Durr, P. and Eigenbrodt, E. (2001) Metabolic cooperation between different oncogenes during cell transformation: interaction between activated ras and HPV-16 E7. *Oncogene* **20**, 6891–6898

Received 10 May 2009/18 August 2009; accepted 25 August 2009
Published as BJ Immediate Publication 25 August 2009, doi:10.1042/BJ20090723

¹ These authors contributed equally to this work.

² To whom correspondence should be addressed (email mvedadi@uhnres.utoronto.ca).