**RESEARCH ARTICLE**

# Thermal adaptation in the honeybee (*Apis mellifera*) via changes to the structure of malate dehydrogenase

Thitipan Meemongkolkiat[1,2], Jane Allison[3], Frank Seebacher[4], Julianne Lim[2], Chanpen Chanchao[5] and Benjamin P. Oldroyd[2],*

## ABSTRACT

In honeybees there are three alleles of cytosolic malate dehydrogenase gene: F, M and S. Allele frequencies are correlated with environmental temperature, suggesting that the alleles have temperature-dependent fitness benefits. We determined the enzyme activity of each allele across a range of temperatures *in vitro*. The F and S alleles have higher activity and are less sensitive to high temperatures than the M allele, which loses activity after incubation at temperatures found in the thorax of foraging bees in hot climates. Next, we predicted the protein structure of each allele and used molecular dynamics simulations to investigate their molecular flexibility. The M allozyme is more flexible than the S and F allozymes at 50°C, suggesting a plausible explanation for its loss of activity at high temperatures, and has the greatest structural flexibility at 15°C, suggesting that it can retain some enzyme activity at cooler temperatures. MM bees recovered from 2 h of cold narcosis significantly better than all other genotypes. Combined, these results explain clinal variation in malate dehydrogenase allele frequencies in the honeybee at the molecular level.

KEY WORDS: Temperature clines, Temperature-dependent selection, Molecular dynamics, Temperature adaptation, Cold narcosis, Geographic distribution, Cline, Allozyme

## INTRODUCTION

Temperature is a major factor influencing the geographic distribution of species. At the molecular level, ectothermic species may adapt to the temperatures that they typically encounter by optimizing a trade-off between protein thermostability at high temperatures and enzyme activity at low temperatures (Dong and Somero, 2009; Jaenicke and Böhm, 1998). Temperature-dependent selection can also lead to clinal variation in protein variants (allozymes), such that a particular allozyme is more common in warmer regions of a species' range and less common in cooler areas (Rank and Dahlhoff, 2002).

Differences in enzyme activity may be accounted for by the structural stability and flexibility of the enzymes of organisms that are adapted to different thermal environments. Enzymes of cold-adapted organisms are generally less stable and more flexible than their homologues in warm climates (Fields, 2001; Olufsen et al., 2005). Enzymes of organisms from warm climates are often characterized by an increased number of salt bridges and cation-π interactions between amino acid residues, which may enhance the structural stability of the protein (Gromiha et al., 2002; Tang et al., 2014). In contrast, a smaller number of cation-π interactions may enhance the structural flexibility and therefore enzyme efficiency of organisms adapted to cold temperatures (Pucci and Rooman, 2017). There are counter-examples (Miyazaki et al., 2000; Wintrode and Arnold, 2001), however, and the relationship between structural stability, flexibility and activity remains controversial (Bjelic et al., 2008; Pabis et al., 2018). A recent proposal that molecular flexibility may be linked to enzyme activity via changes in the balance between the thermodynamic properties (such as activation enthalpies, entropies and heat capacities, all of which are temperature dependent) relevant to catalysis, rather than by directly influencing stability, suggests a way to reconcile these apparently contradictory observations (Arcus and Mulholland, 2020).

Malate dehydrogenase (MDH, EC 1.1.1.37) catalyses the interconversion of malate and oxaloacetate using NAD$^+$ as a coenzyme (Minárik et al., 2002). In eukaryotic cells, MDH occurs in two paralogues (Birktoft et al., 1989). The mitochondrial form is an enzyme of the citric acid cycle. The cytosolic form is an enzyme of many metabolic pathways including the malate/aspartate shuttle, amino acid synthesis and gluconeogenesis, among others (Nelson, 2005).

There are differences in the temperature sensitivity of orthologues of cytosolic MDH (cMDH) in several organisms (Cornuet et al., 1995; Dong and Somero, 2009; Hines et al., 1983), making it an interesting case study to investigate temperature-dependent selection on enzyme activity and stability. For example, in honeybees, *Apis mellifera*, there are three common forms of cMDH designated by their rates of migration during electrophoresis: fast F, medium M, and slow S (Badino et al., 1983). Temperature-dependent clines in the frequency of these alleles exist on four continents: in Europe (Italy) (Badino et al., 1983), North America (Nielsen et al., 1994), South America (Del Lama et al., 1990) and Australia (Hatty and Oldroyd, 1999). Independent clines suggest environment-dependent selection (Endler, 1986). The M allele appears to confer lower fitness in warm climates, but higher fitness in cold climates (Hatty and Oldroyd, 1999; Nielsen et al., 1994; Nunamaker et al., 1984). The reduced fitness of the M allele in warm climates may be explained by its reduced thermostability at higher temperatures (Cornuet et al., 1995).

Molecular modelling and studies of enzyme kinetics can link differences in structural stability and function to Darwinian fitness. For example, the northern-occurring limpet species *Lottia digitalis* and a southern-occurring species (*Lottia austrodigitalis*) have a

[1]Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand. [2]Macleay Building A12, The University of Sydney, Sydney, NSW 2006, Australia. [3]Digital Life Institute and Maurice Wilkins Centre for Molecular Biodiscovery, School of Biological Sciences, University of Auckland, Private Bag, 92019 Auckland, New Zealand. [4]Heyden Laurence Building, The University of Sydney, Sydney, NSW 2006, Australia. [5]Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.

*Author for correspondence (benjamin.oldroyd@sydney.edu.au)

T.M., 0000-0001-8699-4904; J.A., 0000-0002-5699-1726; F.S., 0000-0002-2281-9311; C.C., 0000-0002-2373-5590; B.P.O., 0000-0001-6831-3677

single amino acid substitution within their cMDH that probably accounts for adaptive differences in enzyme function (Dong and Somero, 2009). The substitution results in additional hydrogen bond interactions that are hypothesized to account for the higher resistance to heat denaturation and reduced rate of loss of activity at 42.5°C of the MDH from the warm-adapted southern species compared with its orthologue in the northern species. Information about the flexibility of enzymes at different temperatures can be gained with molecular dynamics (MD) simulation, a computational tool used to explore the structural fluctuations of molecules such as proteins. MD simulations have been used to link temperature-adaptive changes in kinetic properties, thermal stability and structural flexibility for orthologues of MDH from five mollusc genera adapted to different temperatures (Dong et al., 2018; Liao et al., 2019).

In order to be functional, the MDH protein must form a dimer. In heterozygous individuals, heterodimers ($A^1A^2$) form between the two different allozymes, and homodimers ($A^1A^1$ and $A^2A^2$) form between the same allozymes. Here, we examined whether the properties of the three structural variants of the cMDH protein of the honeybee can explain clinal variation in allele frequencies. First, we examined the catalytic efficiency of each allozyme in its homodimeric (SS, MM, FF) and heterodimeric (SM, SF, MF) forms in vitro over a tolerated range of temperatures (15–55°C) and at a temperature that cannot be tolerated (65°C). Temperatures of 15°C are likely to be experienced by worker bees in winter clusters in cool-temperate climates, and when individual bees get caught in the field while foraging at low temperatures. Thoracic temperatures exceeding 52°C occur in workers foraging at high ambient temperatures (Heinrich, 1993).

Second, we inferred the amino acid sequences of the three MDH alleles in silico by translating their DNA sequences. From the amino acid sequences, we inferred the three-dimensional structure of the proteins, and used these structures to identify amino acid residues and sub-structures that may contribute to heat or cold tolerance. We also used MD simulations to explore the structural flexibility of the proteins at different temperatures.

Finally, we examined the cold tolerance of bees of the six genotypes via an assay in which bees were subjected to cold narcosis for 2 h, after which their recovery was recorded at room temperature.

## MATERIALS AND METHODS
### Allozyme analysis and identification of bees of known genotype
Worker bees, Apis mellifera Linnaeus 1758 (n=100), were collected from each of 86 colonies in a MDH-polymorphic population in Thailand. The samples were immediately frozen on dry ice and stored at −80°C in the laboratory. The head of a subsample of bees (total 1032 bees) was used to determine its cMDH genotype (MM, MS, etc.) based on protein electrophoresis (Hebert and Beaton, 1989). The thorax of each genotyped bee was retained at −80°C for later use in enzyme activity assays and for DNA extraction.

### Purification, enzyme activity and thermostability of MDH
Separation of the mitochondrial and cytosolic forms of MDH was achieved via protein electrophoresis (Retamal et al., 1999). Thoracic muscle was dissected from two workers of each of the six possible genotypes, homogenized in buffer (KCl 140 mmol l$^{-1}$, HEPES 20 mmol l$^{-1}$, MgCl$_2$ 5 mmol l$^{-1}$, EGTA 2 mmol l$^{-1}$, ATP 1 mmol l$^{-1}$ and BSA 1%, pH 7.0) and then centrifuged at

9000 rpm for 5 min. The pellet was discarded and the supernatant centrifuged again at 9000 rpm for 10 min. Aliquots (50 µl) of the supernatant were then transferred to Mini-PROTEAN® TGX™ 7.5% precast polyacrylamide gels and electrophoresed at 140 V for 90 min in 1× Tris/glycine buffer (25 mmol l$^{-1}$ Tris, 192 mmol l$^{-1}$ glycine, pH 8.3). Gels were histologically stained (Cornuet et al., 1995) to identify bands associated with cMDH. cMDH bands were cut from the gel and placed in a microcentrifuge tube that was perforated at the base. The perforated tube was placed inside a second tube and centrifuged at 10,000 rpm for 2 min at 4°C. The resulting gel fragments were shaken gently in 180 µl of homogenization buffer (see above) at 4°C overnight. The tubes were then centrifuged at 10,000 rpm for 5 min at 4°C. An 8 µl aliquot of the supernatant was re-electrophoresed on a cellulose acetate gel to visually confirm that only the cMDH form was present. The mitochondrial isoform of MDH is eliminated by this method (Fig. S1). The final supernatant was used immediately to determine the biochemical characteristics of the six dimers.

Enzyme activity was measured with an Ultrospec 2100 pro UV/visible spectrophotometer (GE Healthcare) equipped with a temperature-controlled cuvette holder connected to a bath circulator (Thermoline Scientific) in 5°C steps from 15 to 60°C for all genotypes. Activity was determined from the reduction in absorbance of NADH at 340 nm (Childress and Somero, 1979) as units g$^{-1}$ of extracted tissue. The assay medium was 100 mmol l$^{-1}$ Tris-HCl buffer (pH 8.1), 20 mmol l$^{-1}$ MgCl$_2$, 0.4 mmol l$^{-1}$ oxaloacetate and 0.15 mmol l$^{-1}$ NADH (Childress and Somero, 1979), and was not limiting to reaction rates. There were five replicates of a pool of three bees for each temperature–genotype combination.

To determine the thermostability of the different isoforms, we prepared purified samples via electrophoresis as above. Aliquots (140 µl) of at least 1 mg ml$^{-1}$ of the six dimers were incubated at 35 and 55°C. Samples (15 µl) were transferred to ice at t=0, 1, 5 and 10 min. The activity of the treated enzymes was then quantified at 30°C (Childress and Somero, 1979). The relative activity was defined as the ratio between the mean activity at time 0 and the mean activity at time t across six replicates (five for MF and SF).

Wilk's lambda was used to evaluate the effects of genotype and temperature (with temperature as a repeated measure) on enzyme activity. The equality of variances for within-subject effects was satisfactory (Levene's test P=0.05–0.99) for the adjusted means.

### Recovery from cold narcosis
We sampled 20 random workers from 10 colonies of the same strain (commercial bees derived from Apis mellifera ligustica Spinola 1806) and genotyped the workers to assess the most likely cMDH genotype of the queen heading the colonies, as above. From these results, we selected four colonies that in combination would provide bees of all genotypes, with at least two colonies producing individuals of each genotype.

Combs of emerging brood were removed from the selected colonies and placed in an incubator overnight at 34.5°C in individual cages. After 36 h, 100 bees were placed in individually labelled 1.5 ml tubes and the tubes placed in an ice–water slurry, inducing immobility in all bees.

After 2 h, bees were removed from the ice slurry and placed on the bench at room temperature. A minimum of two observers recorded the time at which the first movement occurred. Bees that did not move after 60 min were classified as not recovering. Following this assay, all bees were frozen at −20°C before genotyping as above. Importantly, recovery time was recorded by observers unaware of

the bee's genotype. This process was repeated 6 times, producing a total of 532 successfully assayed and genotyped bees.

The time to recovery was analysed as a mixed model ANOVA with genotype as a fixed effect and colony as a random effect. Variances were homogeneous among means ($P=0.37$). To analyse survival, we used a binomial GLM to examine the effect of genotype on recovery with a logit link function.

## cMDH sequence analysis

We designed four sets of primers that covered the complete MDH gene as follows: set 1 forward: TGGAGCTGCAGGACAAATTG, reverse: TGAGCGCGATTTTGATCCAA; set 2 forward: CTCA-AGCAGCATTATCAGCAAA, reverse: CTTCTGCACGTTCCT-CTTCT; set 3 forward: ACAGCAATCATCCTAACTGACC, reverse: TGAGCGCGATTTTGATCCAA; set 4 forward: TTTCT-GGTTGGAGCAATGCC, reverse: TTGCCAGAACAAGTAATT-GGC. The expected product sizes were 736, 666, 1211 and 976 bp, respectively.

Total DNA was extracted from the thorax of workers with genotype FF, MM and SS using the DNeasy® Blood and Tissue Kit (69506, Qiagen). The PCR reactions comprised 1 µl of each of the 10 µmol l$^{-1}$ forward and reverse primer, 1.5 µl of 1.5 mmol l$^{-1}$ MgCl$_2$, 2.5 µl of 10× PCR buffer, 2 µl of 2 mmol l$^{-1}$ dNTP, 1.7 µl of DNA template, 0.4 µl of Taq-ti polymerase and 14.9 µl of H$_2$O. The PCR conditions were 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 56–57°C

for 1 min and 72°C for 1 min, with a final extension at 72°C for 7 min. Products were electrophoresed in 1% agarose gels and visualized under UV light. The amplified DNA products were purified and direct sequenced.

The 334 amino acid sequences of the three alleles were inferred from the DNA sequences (Fig. 1). DNA and amino acid sequences of genes have been uploaded to the GenBank database with the following accession numbers: cMDH S allele, MK967647; M allele, MK967648; F allele, MK967649.

## Molecular modelling

The three-dimensional structures of the F, M and S allozymes were modelled based on the template structure of porcine cMDH (*Sus scrofa*) (5MDH, ChainA) proteins in the PDB database (Chapman et al., 1999) using Swiss-Model (Waterhouse et al., 2018). Swiss-Model showed 63.53% amino acid sequence identity between porcine MDH and each allozyme of *A. mellifera.* Model quality was evaluated using the Swiss-Model QMEAN Z-score (Benkert et al., 2011). The Swiss-Model QMEAN Z-score provides an estimate of the 'degree of nativeness' of the structural features observed in the model on a global scale. QMEAN Z-scores around zero indicate good agreement between the model structure and experimental structures of similar size, whereas scores of −4.0 or below indicate models with low quality. Also, PROCHECK was used to independently check the percentage of the assessment of the



**Fig. 1. Amino acid sequence alignment of three cytosolic malate dehydrogenase (cMDH) alleles (F, M and S) from *Apis mellifera* and pig (*Sus scrofa*).** The inferred secondary structure was calculated using Swiss-Model software based on the porcine cMDH [PDB, 5mdhA is indicated above each segment of sequence by α (α-helices) or β (β-strands) followed by a number and/or letter]. The amino acid substitutions among the three honeybee alleles are indicated in red. The residues involved in substrate binding, cofactor binding and subunit–subunit interactions are shown in yellow, blue and green, respectively.

overall quality of the structure as compared with well-refined structures of the same resolution (Laskowski et al., 1993). The reliability of the predicted quaternary structure was evaluated with the Swiss-Model Quaternary Structure Quality Estimate (QSQE) score. The QSQE score ranges from 0 to 1, and is used to sort templates in each interface similarity level to a cluster. In general, a QSQE above 0.7 indicates that the modelled structure is reliable (Bertoni et al., 2017). The three-dimensional structures were visualized using Swiss-PDB Viewer (Guex and Peitsch, 1997). The QMEAN for F, M and S monomer was −0.89, −0.98 and −0.98, respectively, relative to −0.42 of template Porcine 5MDH-A. The PROCHECK for F, M and S monomer was 91%, 92% and 91%, respectively, relative to 98% of the template. The QSQE score of F, M and S was 0.97%, 0.98% and 0.97%, respectively, relative to 99% of the template.

### Intermolecular bonding interaction

The distances between donor (N) and acceptor (O) atoms were calculated and visualized for each cMDH allozyme structure using Swiss-PDB Viewer 4.1 (Guex and Peitsch, 1997). Hydrogen bonds were assumed to form within a distance of 3.5 Å, and the hydrogen donor–acceptor angle threshold for hydrogen bond formation was ≤30 deg.

Cation-π interactions were inferred to occur when an ε-amino nitrogen atom of lysine (NZ) and the guanidino nitrogen atom of arginine (NH1, NH2) were within 0.8 nm of any of the carbon atoms in the ring of an aromatic amino acid (tyrosine, tryptophan, phenylalanine) (Gromiha et al., 2002), as determined using the CaPTURE program (Gallivan and Dougherty, 1999).

The number of salt bridges was determined using VMD (Humphrey et al., 1996). Oppositely charged atoms (carboxylate oxygen atoms in aspartate and glutamate, and ε-amino, guanidino and the (de)protonated imidazole nitrogen atoms in lysine, arginine and histidine, respectively) <0.4 nm apart were considered to form salt bridges (Kumar and Nussinov, 2002).

### Molecular dynamics simulations

GROMACS software version 2016.1 (Abraham et al., 2015) was used to prepare and run the MD simulations and analyse the resultant trajectories. The protein force field was CHARMM36 (Huang and MacKerell, 2013) with the transferable intermolecular potential 3P (TIP3P) water model (Mark and Nilsson, 2001).

Each three-dimensional structure of the dimeric protein was initially placed in a virtual cubic box filled with water molecules. The size of the water box was sufficient to provide a layer of water 1.2 nm thick between the protein and the edge of the box. Each system was energy minimized using the steepest descent method for 5000 steps. The systems were then pre-heated in the NVT [i.e. amount of substance (N), volume (V) and temperature (T) are conserved] ensemble and equilibrated under isotropic NpT conditions for 125 ps. Each dimer was then simulated in triplicate for 100 ns at 15, 30 or 50°C, with coordinates collected every 0.2 ps. Coulomb and van der Waals interactions were computed within a cut-off distance of 1.2 nm; outside of this distance, the long-range electrostatic interactions were treated using the particle mesh Ewald method (Kolafa and Perram, 1992). The lengths of all bonds involving hydrogen atoms were constrained using LINCS (Hess et al., 1997) with an order of four to allow an integration time step of 2 fs. The temperature (NVT) and the pressure (NpT) were maintained using the Berendsen thermostat and barostat (Berendsen et al., 1984) with coupling times of $\tau_T$=0.1 ps and $\tau_p$=2 ps, respectively. Three independent replicate simulations of each

system were run. VMD was used to visualize the simulation trajectories (Humphrey et al., 1996) and to create snapshot images.

The total number of hydrogen bonds between all possible donor and acceptor atoms occupied for more than 80% of the simulation in each structure sampled was calculated using VMD (Humphrey et al., 1996). The OH and NH groups are regarded as donors, and the O and N groups as acceptors. To be considered a hydrogen bond, the distance between donor and acceptor had to be within 3.5 Å and the hydrogen donor–acceptor angle was ≤30 deg.

To quantify and evaluate the structural flexibility of the cMDH proteins from different alleles, we carried out MD simulations and analysed them in terms of the atom-positional root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF). The RMSD and RMSF were calculated using the gmx rmsd and rmsf tool. Differences between the RMSD and RMSF calculated from simulations of different dimers or run at different temperatures can reveal their sensitivity to temperature. The RMSD is a measure of the movement of the position of the selected set of atoms from their initial position over time. Here, the RMSD was computed for the Cα atoms of the protein so that only major backbone motions were considered. Averages were calculated for the equilibrated portion of each simulation, from ∼25 ns onwards, and further averaged across the three replicate simulations under each set of conditions. Many different conformations can have a given RMSD value with respect to the initial structure. Higher RMSD values therefore indicate a broader equilibrium ensemble of structures, which in turn suggests that the protein has greater flexibility.

The RMSF is a measure of the degree of fluctuation of each atom or residue, computed by averaging the difference between the position of each atom at each point in time and its mean position. Here, it was calculated for all atoms and summarized for each residue of the protein, as well as averaged across all atoms in the protein. The per-residue RMSF analysis allows characterization of the movements of each amino acid residue, which is helpful in identifying regions where large structural displacements key to function occur.

The significance of any differences in average RMSD and RMSF values between dimers was assessed by using one-way analysis of variance followed by Tukey's HSD *post hoc* tests for differences between means.

## RESULTS

### Enzyme activity and thermal stability

Enzyme activity of all dimers increased with temperature (Fig. 2; Table S1). Maximum activity was recorded at 55°C. Above 55°C, activity gradually decreased. For the three homodimers there was a strong effect of temperature ($F_{7,15}$=7.15, $P$<0.001) and genotype ($F_{2,21}$=20.84, $P$<0.001) on enzyme activity, and a significant genotype by temperature interaction ($F_{14,30}$=3.43, $P$=0.002), a consequence of lower activity of the MM homodimer relative to the FF and particularly SS homodimers at temperatures above 50°C (Fig. 2). Both temperature ($F_{9,4}$=1619.0, $P$<0.001) and genotype ($F_{2,12}$=9.44, $P$=0.003) also had a strong effect on the activity of the heterodimers, but there was no significant interaction ($F_{18,8}$=1.54, $P$=0.27). The average activity of SM and FM was significantly less than that of SF (orthogonal contrast, $F_{1,21}$=23.58, $P$<0.001). Overall, the results indicate that a dimer containing one or more M allozyme has lower activity than dimers containing only F and S allozymes at 55°C in both homodimers and heterodimers.

The activity of all dimers declined only slightly with increasing incubation time at 35°C (Fig. 3). After 10 min incubation, all homodimers retained ca. 90% of their original activity (Fig. 3), with
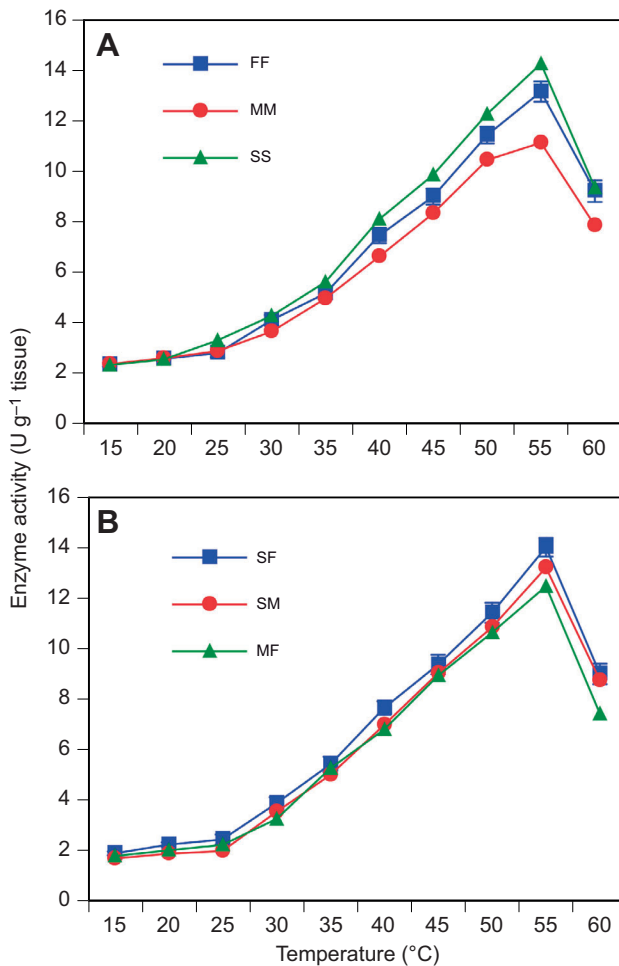
Fig. 2. Mean enzyme activity for five replicate samples of each of the six dimers of cMDH from *A. mellifera* at different temperatures. (A) Homodimers, (B) heterodimers. Error bars are s.e.m., but these are smaller than the symbols in most cases.

no significant differences between genotypes ($F_{5,27}=1.88$, $P=0.13$). With a pre-assay incubation temperature of 55°C, SS and FF had a greater residual activity after 10 min than MM, SM and MF (orthogonal contrast of the mean of SS and FF versus the mean of MM, SM and MF, $F_{1,27}=48.5$, $P<0.001$). Overall, these results show that dimers that include the M allozyme have less activity than dimers that do not, after incubation at 55°C (Fig. 3).
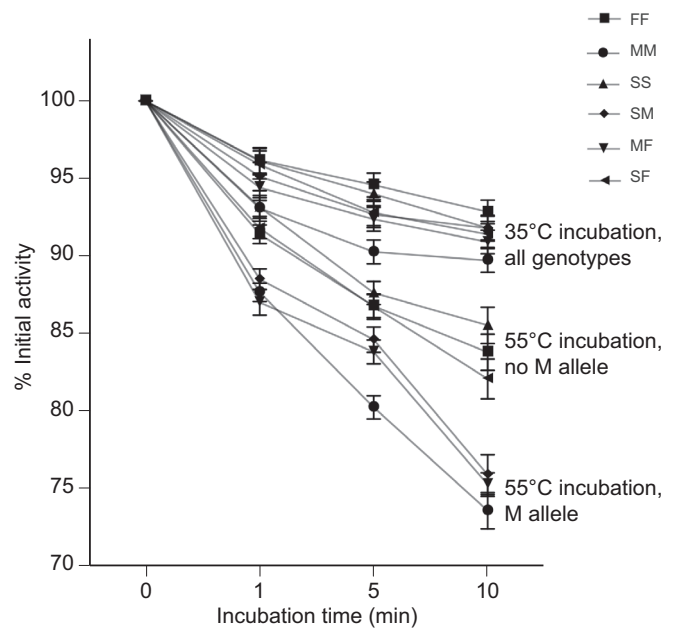


Fig. 3. Thermal stability of the six dimers of cMDH from *A. mellifera*. The residual enzyme activity of the heat-treated enzyme after incubation at different temperatures (35 and 55°C) is shown for different incubation periods. There were $n=6$ replicate assays per genotype except for MF and SF ($n=5$). Error bars are s.e.m.

## Recovery from cold narcosis

Genotype had a highly significant effect on the time to recovery from cold narcosis ($F_{5,11.0}=13.2$, $P<0.001$). Bees of the MM genotype recovered significantly faster than all other genotypes ($P<0.001$; Fig. 4A). Colony was unimportant to recovery time ($F_{3,9.3}=0.53$, $P=0.67$) and there was no interaction between the main effects ($F_{9,370}=0.53$, $P=0.85$). Similarly, genotype had a significant effect on the proportion of bees that recovered (Wald $\chi^2=11.46$, d.f.=5, $P=0.04$; Fig. 4B). Bees of the MM genotype were significantly ($P=0.007$) more likely to recover than bees of other genotypes.

## Molecular sequence, structure and interactions

*Apis mellifera* MDH alleles have high DNA and inferred amino acid sequence identities (98.8–99.9%). To evaluate how the amino acid substitutions might contribute to their stability, we used the amino acid sequences to build three-dimensional homology models using porcine MDH structures as a template (Fig. 5). The amino acid
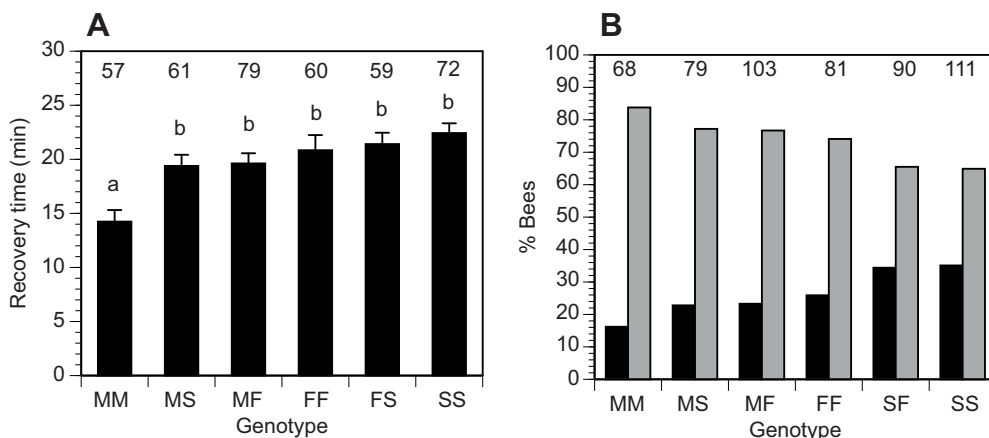


Fig. 4. Effect of cold narcosis on bees of different MDH genotype. (A) Mean recovery time and (B) percentage of bees that did or did not recover after 2 h of cold narcosis. Bees of the MM genotype recovered significantly more quickly than all others. In A, error bars are s.e.m. Means with a different letter are significantly different at the 5% level based on Tukey's HSD test. In B, black bars indicate the number of bees that died, grey bars indicate the number that recovered. Numbers above the bars are the sample size for that genotype.
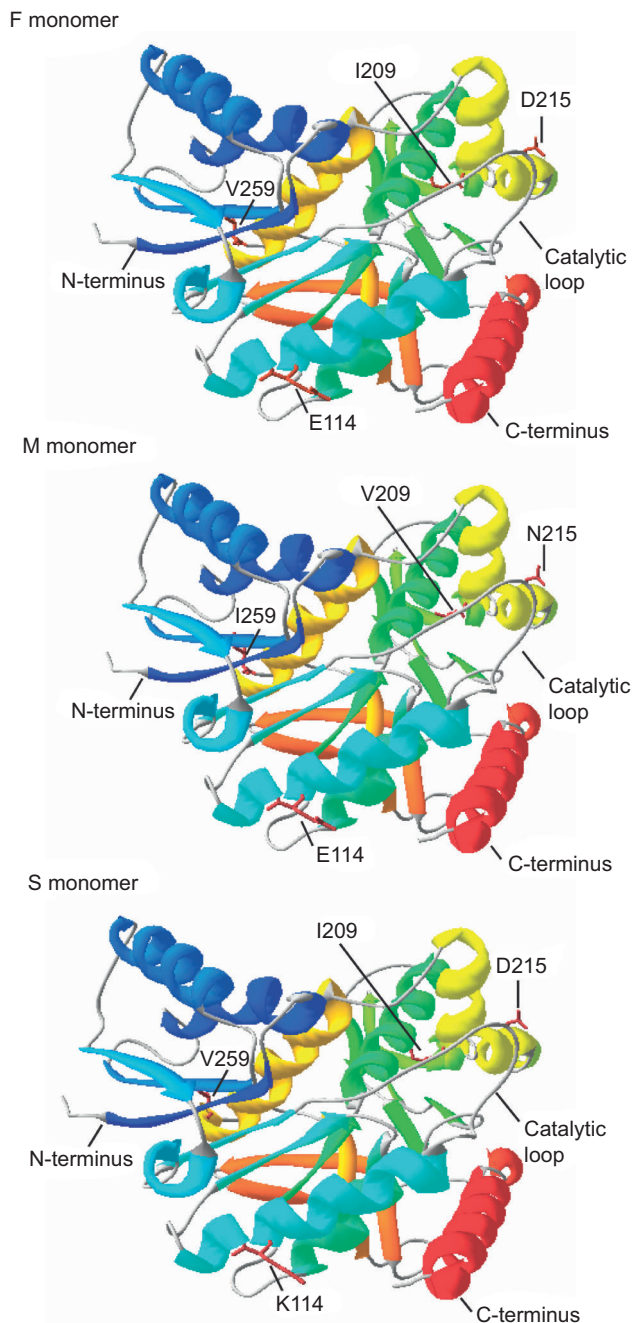
**Fig. 5. Three-dimensional model structure of each cMDH allozyme from *A. mellifera*.** The differences in amino acid residues across the three allozymes are indicated.

substitutions among the three honeybee alleles occur in the αDE-helix, between α1G′ and α3G, and in β-strand βJ (Figs 1 and 5). The key active site and cofactor binding residues are conserved between pig and bee (Fig. 1), suggesting that they are necessary to maintain the function of the protein (Chapman et al., 1999).

To determine the possible effect of these substitutions on the structural stability of each enzyme, we considered how many hydrogen bonds each residue forms in the model structures. The S allele lysine at position 114, in the αDE-helix, forms six hydrogen bonds with Lys-110, Val-111, Asp-117, Lys-118, Ala-115 and Gly-113 (Table S2, Fig. S2A). In contrast, the glutamate at position 114 of the M and F alleles form only five hydrogen bonds (Fig. S2A). At

position 209, in βJ, both Ile-209 of the F and S alleles and Val-209 of the M allele form six hydrogen bonds with other nearby amino acids (Table S2, Fig. S2B). At position 215, between α1G′ and α3G, both the F and S alleles have aspartic acid, which forms eight hydrogen bonds. In contrast, the M allele asparagine forms just seven hydrogen bonds (Table S2, Fig. S2C). At position 259, in α3G, the F and S alleles have valine and the M allele has isoleucine, but all form three hydrogen bonds with a nearby amino acid (Table S2, Fig. S2D). The majority of these hydrogen bonds involve the polypeptide backbone, with the exception of Asp-215 and Asn-215, where the side chain oxygen atoms of both variants also form a hydrogen bond with the backbone N of Trp-218, suggesting that they should be able to be formed by any sequence variant. The additional hydrogen bond interactions for the S and F alleles may imply greater stability of these structures.

Other types of non-covalent interactions that can contribute to protein stability are cation-π interactions and salt bridges (Tayubi and Sethumadhavan, 2010). The formation of cation-π interactions for four possible combinations of positively charged and aromatic amino acids (Arg-Tyr, Lys-Tyr, Lys-Phe, Lys-Trp) are shown in Table S3. The S allele has 11 cation-π interactions whereas the F and M alleles have 10. The number of salt bridges for each allele is given in Table S3. The S and F alleles have 15 salt bridges, one more than the M allele. In sum, the predicted allozyme structures suggest that the amino acid substitutions of the S and F alleles lead to additional hydrogen bonds and cation-π and salt bridge interactions, which may increase their structural stability relative to that for the M allele by providing structural scaffolding.

### MD simulation of structural flexibility

For all simulations, the RMSD values needed approximately 25 ns to stabilize (Fig. 6). The average RMSD values reported in Fig. 7 were therefore calculated for the last 75 ns of each simulation. At 15°C, there was a highly significant effect of genotype on RSMD value ($F_{5,12}$=10.32, $P$=0.0001). The SS homodimer had significantly lower average RMSD values than all other genotypes, suggesting that SS dimers have a particularly rigid structure that may impede catalytic function at low temperatures. MM and MF had higher average RSMD values, indicating that they are more flexible and thus likely to retain some enzyme activity at cold temperatures. At simulated temperatures of 30°C there was no significant difference in the average RMSD values between genotypes ($F_{5,12}$=1.13, $P$=0.33; Fig. 7C). At 50°C, genotypic differences were again apparent ($F_{5,12}$=38.92, $P$<0.001; Fig. 7E); the MM and MF dimers had significantly higher average RMSD values than all others, exceeding 0.3 nm after 21 and 36 ns, respectively (Fig. 6E,F). In summary, dimers containing one or two M subunits tend to be the most flexible, and those containing one or two S subunits the least flexible. This may explain the resistance of S to prolonged exposure to high temperatures, and the tendency of dimers containing M to lose activity at high temperatures.

The pattern of RMSF values per residue indicates which residues are the most flexible and whether this changes at different temperatures. RMSF values more than 0.3 nm indicate regions of the protein that are particularly flexible. RMSF values per residue are plotted in Fig. 8. There was a significant effect of genotype on the average RMSF value at all temperatures (15°C $F_{5,12}$=14.27, $P$<0.001; 30°C $F_{5,12}$=10.48, $P$<0.001; 50°C, $F_{5,12}$=37.37, $P$<0.001; Fig. 7B,D,F). These results indicate an extremely rigid structure for SS at 15°C, higher flexibility for MF and MM than for other dimers at 30°C, and particularly flexible and therefore potentially unstable structures for MF and MM at 50°C. In summary, dimers containing
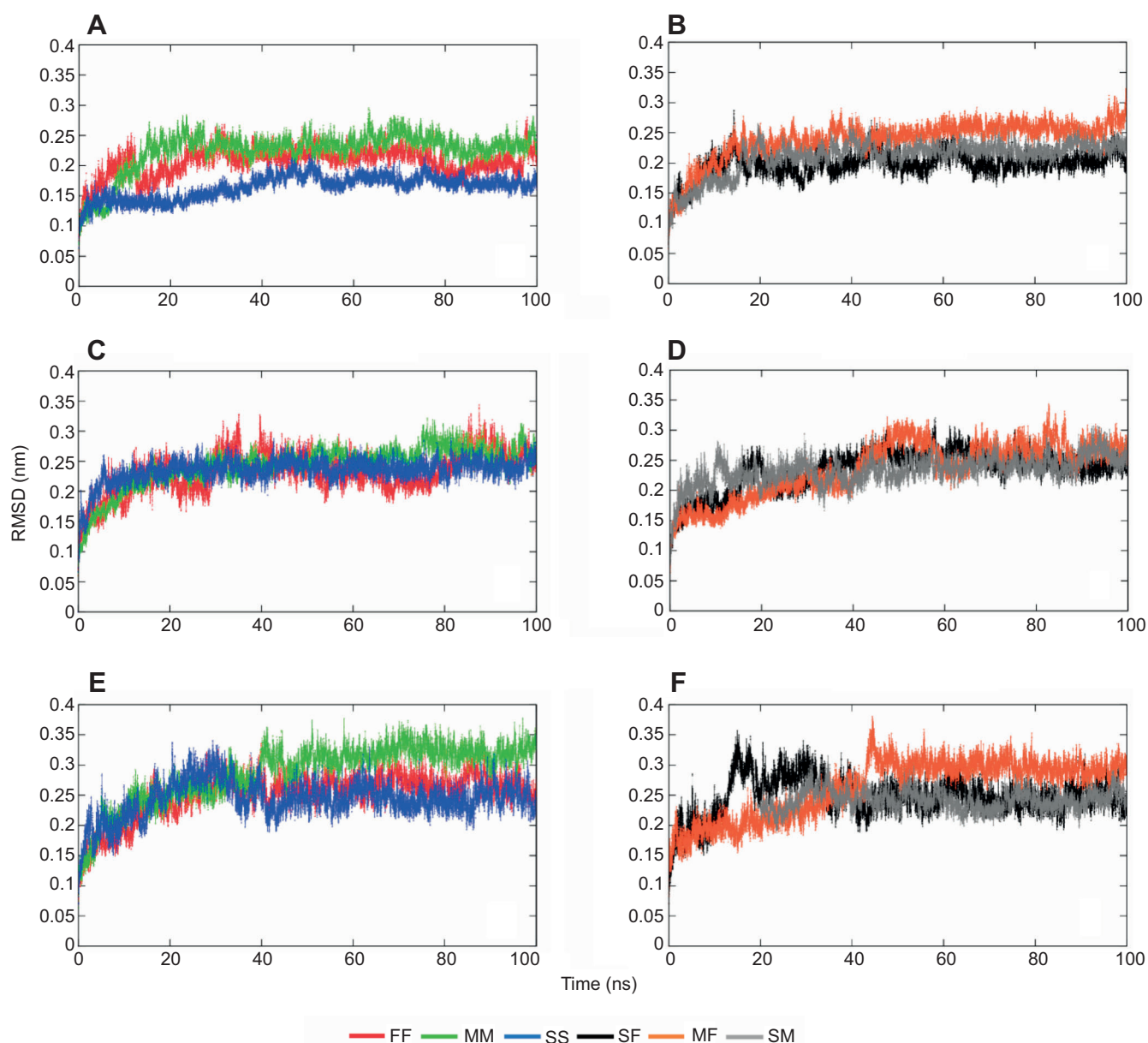
**Fig. 6. Root mean square deviation (RMSD) of the Cα atoms of cMDH from their initial position at different temperatures.** Simulations were run at 15°C (A,B), 30°C (C,D) and 50°C (E,F). Data are shown for the homodimers (A,C,E) and heterodimers (B,D,F) of cMDH.

one or two M subunits are the most flexible, and those containing one or two S subunits the least flexible. This may explain the resistance of S to prolonged exposure to high temperatures, and the tendency of dimers containing M to lose activity at high temperatures.

The pattern of RMSF values per residue indicates which residues are the most flexible and whether this changes with temperature (Fig. 8). Intriguingly, the highest RMSF values across all temperatures and dimers occurred at residues 90–100. This region is involved in the active site loop that folds onto the catalytic site during ligand–substrate binding (Chapman et al., 1999). Flexibility of these residues is therefore likely to be related to substrate binding and product release, and tuning of this flexibility could be critical to thermal adaptation (Gerstein and Chothia, 1991). In line with this hypothesis, these residues have the lowest flexibility in the SS dimer and in the S subunit of the SF and SM heterodimers at 15°C and, to a

lesser extent, at 30 and 50°C (Fig. 8). At 30°C, RMSF values were significantly higher for MM than for FF and SS, MF was more flexible than SF and SM, and the M monomer was the most flexible in the SM and MF heterodimers (Fig. 8). At 50°C, MM and the M monomer of the heterodimers again had the highest flexibility, while SS and FF retained values similar to those at 30°C (Fig. 8).

Two other regions also showed increased flexibility relative to the remainder of the protein. Residues 203–207 reside in the non-catalytic domain. This region showed a similar degree of movement in all six dimers at 15°C and 30°C, with slightly higher RMSF values for the MM dimer and the M monomer of the SM and MF heterodimers. When the simulation temperature was 50°C, however, the RMSF values in this region of both the MM and FF dimers were significantly increased relative to those of other dimers (Fig. 8). The other region with high RMSF values occurs at residues 310–330, which is located near the C-terminus of the protein. Again, the MM
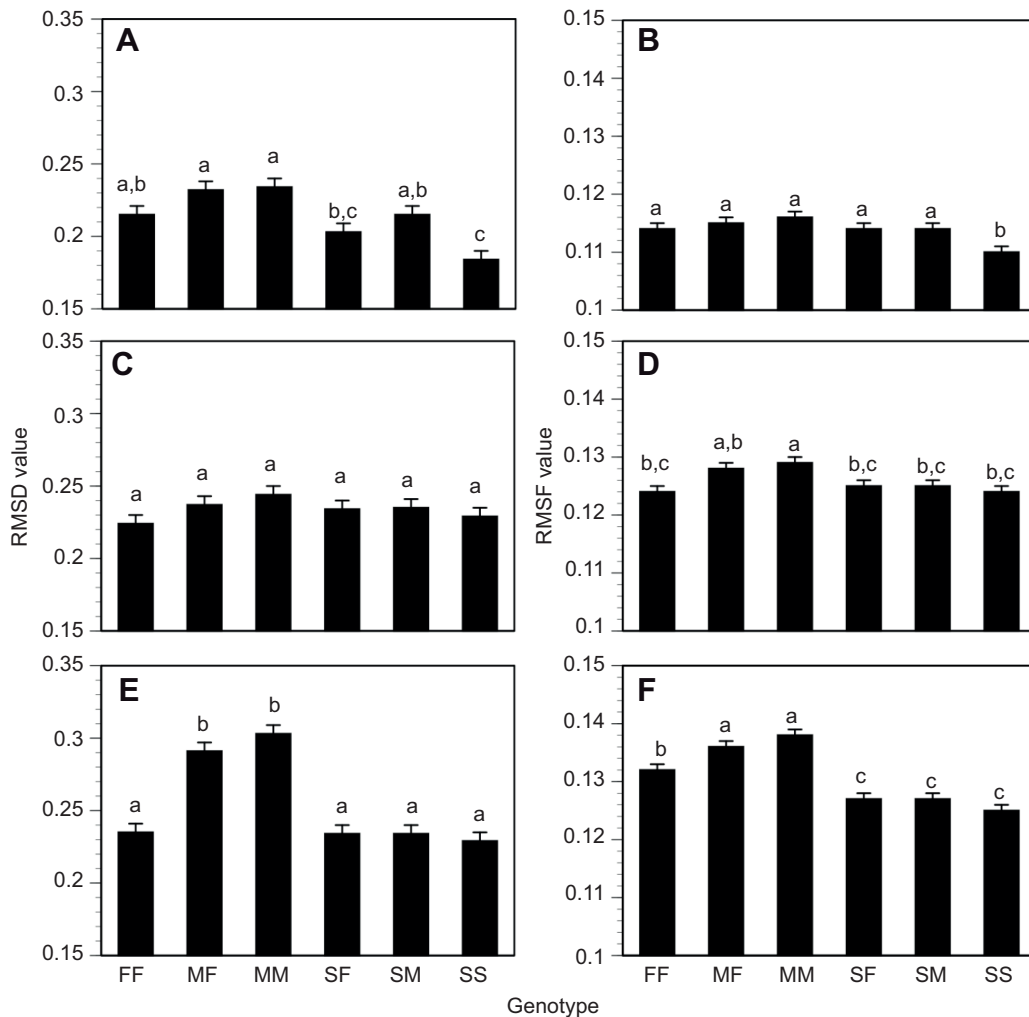
**Fig. 7. Mean equilibrium (approximately 25–100 ns) RMSD of Cα atoms from their initial position and mean equilibrium root mean square fluctuation (RMSF) of individual residues from their mean position for the six possible dimers of cMDH at different temperatures.** Values were averaged across the three simulations of each dimer at each temperature (A,B: 15°C; C,D: 30°C; and E,F: 50°C; $n$=3), and error bars are s.e.m. Bars with a different letter within a particular temperature are significantly different at the 5% level (Tukey's HSD test).

and MF dimers at 50°C showed significantly increased RMSF values at these residues compared with other dimers (Fig. 8).

We also monitored hydrogen bond formation during the MD simulations, focusing on the residues that differ between the dimers and are involved in hydrogen bonds that are occupied for more than 80% of the simulation (Table S4). The total number of hydrogen bonds of each dimer structure varied between allozymes, and was highest for SS, followed by FF, SM and SF, and was lowest in MF and MM (Table 1). This suggests that the sequence differences between alleles cause subtle conformational changes that affect the number of hydrogen bonds. The extra hydrogen bond interactions in dimers with an S allozyme explain their reduced flexibility, and may also contribute to the greater stability of these structures at high temperatures.

## DISCUSSION
It is often difficult to disentangle the effects of historical hybridization events and contemporary selection on allele frequencies. Therefore, compelling demonstrations of natural selection in the wild are rare, especially at the molecular level. Parallel clines in allele frequencies in independent populations

provide an exception to this generalization, and are regarded as clear-cut examples of natural selection in action (Endler, 1986). Enzyme molecules such as MDH must maintain a balance between stability and flexibility to ensure that the molecules do not unfold under typical cellular temperatures, while allowing the motions essential for catalysis at an appropriate rate to maintain metabolic flux (Feller, 2008; Somero, 2010). Enzyme activity may be adjusted quantitatively in response to temperature by changing enzyme concentration and rates of transcription (Crawford and Powers, 1989, 1992). Alternatively, there may be qualitative changes by expressing allozymes and isozymes with different thermal sensitivities (Fields and Somero, 1997). Our data indicate that enzyme activity (in U $g^{-1}$ tissue) changes in parallel with protein-specific (U $mg^{-1}$ protein) rates. It therefore appears that increases in enzyme concentration do not compensate for reduced enzyme-specific catalytic rates (Pierce and Crawford, 1997), and that qualitative changes explain differences in thermal sensitivity. Our data instead suggest that allelic substitution at the population level explains why similar temperature-dependent clines are maintained in honeybee populations worldwide. First, the M allozyme is less thermostable than the S and F allozymes. Second, MM dimers have
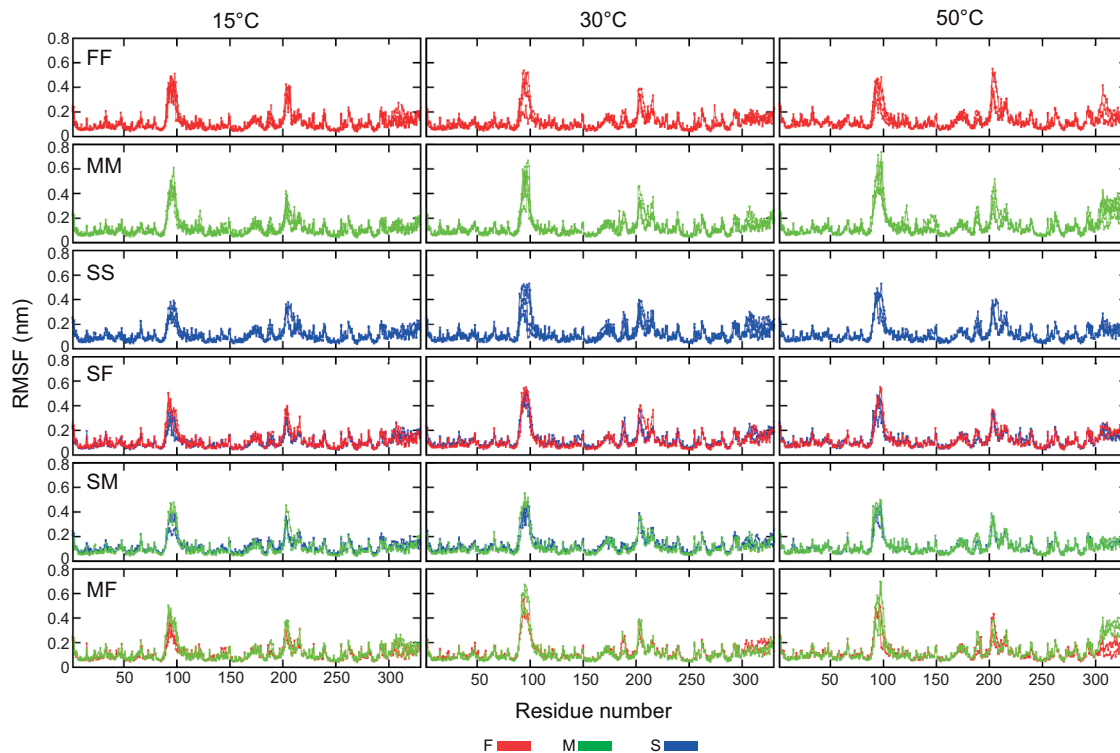
**Fig. 8. RMSF of individual residues of cMDH from their mean position at different temperatures.** Data are shown for homodimers and heterodimers during each of the three replicate simulations (15, 30 and 50°C) as labelled.

lower activity than SS and FF dimers at field-realistic temperatures between 35 and 55°C, possibly as a consequence of enzyme denaturation. These are significant selective disadvantages for the M allele, and probably explain its near absence in warm-climate populations. In contrast, bees carrying the M allele, in homozygous form, have significantly enhanced ability to recover from cold-induced narcosis. It is this capacity that is likely to enhance the fitness of the M allele in colder climates.

Our examination of protein structure suggests how the M allele enhances survival and reduces recovery times of MM bees from chilling relative to bees of other genotypes. Our predicted structures show that M has a reduced number of hydrogen bonds, salt bridges and cation-π interactions and that this weakens intermolecular electrostatic interactions. A similar reduction in intramolecular interactions has been noted for the low-temperature variant of cMDH from marine molluscs (Dong and Somero, 2009). Although both the mollusc and bee analyses were carried on homology-modelled structures, the high sequence similarity between cMDH

**Table 1. Mean (±s.e.m.) number of hydrogen bonds occupied for more than 80% of the 100 ns of MD simulation**

| Genotype | Mean no. of hydrogen bonds |
|---|---|
| FF | 228±0.082 |
| MM | 225±0.069 |
| SS | 235±0.076 |
| SF | 233±0.085 |
| SM | 232±0.043 |
| MF | 226±0.091 |

A hydrogen bond was defined as forming if the donor and acceptor atoms were within 3.5 Å of each other and the hydrogen donor–acceptor angle was ≤30 deg. Data were averaged across the three simulations of each *Apis mellifera* cMDH dimer at 30°C (*n*=3).

from the species studied and the template (porcine), and excellent structural quality scores provide confidence that the results are valid.

Three flexibility-increasing amino acid substitutions in the M allozyme are likely to enhance its catalytic efficiency at low temperatures. First, position 114 falls within the αDE-helix located at the edge of the catalytic loop near the residues of the active site. This α-helix needs to be flexible because it forms the catalytic vacuole within which the reduction of oxaloacetate takes place (Chapman et al., 1999; Gerstein and Chothia, 1991). Investigations of cMDH from marine molluscs have also shown that low-temperature variants also have high flexibility in this region (Dong and Somero, 2009; Dong et al., 2018). Second, in a clear example of convergent evolution, substitutions related to temperature sensitivity and adaptation have arisen multiple times at residue 114 in different species (Liao et al., 2017), suggesting it is a key residue for catalytically important flexibility. Third, the residue at 215 is between the αG′- and α1G′-helices. These helices are known as the 'major-mover' regions. Their mobility is essential for catalysis and they are hot spots for adaptation to cold environments (Fields et al., 2015). Such a key role of relatively few substitutions via a change in flexibility to regions involved in catalysis (but not necessarily the active site) has been previously observed in temperature-sensitive enzymes (Fields et al., 2015). In contrast to findings for mollusc cMDH enzymes (Dong and Somero, 2009; Dong et al., 2018), we did not detect variability in flexibility among allozymes for residues 230–240. This may be because these residues are close to the dimer interface, and we studied dimers rather than monomers.

In order to function, an enzyme requires a degree of flexibility, and thus an enzyme that is adapted to a cold temperature needs to be flexible at low temperatures (Feller, 2008; Olufsen et al., 2005). This may mean that it becomes overly flexible at higher temperatures.

Conversely, an enzyme adapted to function at a higher temperature may become overly rigid at lower temperatures. At 15°C, the largest average RMSF values, indicative of greater flexibility, were detected in the MM and MF dimers, followed by FF. This is in keeping with the lower number of hydrogen bonds, cation-π and salt bridge interactions of M and F allozymes relative to the S allozyme. The greater flexibility of the M allozyme at low temperatures may be evidence of adaptation to cold environments. At a moderate simulation temperature of 30°C, which is within the typical body temperature range of *A. mellifera* (30–35°C; Heinrich, 1993), all dimers exhibited relatively similar RMSF values, consistent with the similar thermostability of the proteins at 30°C, and showing no effects of temperature adaptation. However, when the simulation temperature was increased to 50°C, a significant increase in the average RMSF value was observed for the heat-labile MM dimer. In contrast, the more heat-stable FF and SS dimers showed average RMSF values similar to those exhibited at 30°C. This increased rigidity is in keeping with their higher incidence of intermolecular bonding. These results suggest that, in the case of *A. mellifera* cMDH, and in keeping with other cMDH enzymes (Dong et al., 2018; Liao et al., 2019), flexibility is inversely correlated with thermal stability.

The amino acids with the highest RMSF values were residues 90–100, which are involved in catalysis (Chapman et al., 1999). High RMSF values were also observed for these residues in MD simulations of monomeric *Lottia* spp. cMDH (Dong et al., 2018; Liao et al., 2017). In our simulations, the MM dimer retained flexibility of this region at lower temperatures, whereas the SS dimer was more rigid. In contrast, at 50°C, several groups of residues involved in catalysis were extremely flexible in the MM dimer, whereas the SS and FF dimers exhibited similar flexibility to that observed at 30°C. This is in keeping with the widely understood importance of the balance between protein flexibility and rigidity in ligand binding (as quantified by $K_m$ values), and observations that increased flexibility especially around the active site increases activity at low temperatures (Siddiqui and Cavicchioli, 2006).

Our study suggests that the selective pressures on cMDH of *A. mellifera* are complex. The M allele is clearly at a selective disadvantage in hot climates, but an advantage in cold climates, suggesting that the locus is subject to diverging selection. The S allele retains its structural integrity and has the highest thermal stability and the highest activity at high temperatures, whereas the F allele is somewhere between the other two. The formation of heterodimers further complicates the picture, and suggests a heterozygote advantage. Heterodimers of the M allozyme have higher thermal stability than the MM homodimer, and so this may reduce selection against M when M is at low frequency in hot climates. It is even possible that M heterozygotes have higher fitness than all other genotypes, in all but the hottest environments, because they can tolerate a broader temperature range. Note for example, that M heterozygotes had the second highest survival rate following cold narcosis after MM. Heterozygote advantage may help to maintain polymorphisms in populations such as those in Italy (Badino et al., 1983), California (Nielsen et al., 1994), Tasmania (Oldroyd et al., 1995a), Kangaroo Island South Australia (Oldroyd et al., 1992) and north west Victoria, Australia (Oldroyd et al., 1995b) that experience strong fluctuations in temperature.

In conclusion, our study shows how a fitness trade-off in cMDH variation is manifest at the level of protein structure, and clarifies why the M allele has a fitness advantage in cold climates and a disadvantage in hot climates. In this period of rapid global temperature change, we predict that MDH allele frequencies will change as honeybee populations experience new thermal environments.

## References
**Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. and Lindahl, E.** (2015). GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25. doi:10.1016/j.softx.2015.06.001

**Arcus, V. L. and Mulholland, A. J.** (2020). Temperature, dynamics, and enzyme-catalyzed reaction rates. *Annu. Rev. Biophys.* **49**, 163-180. doi:10.1146/annurev-biophys-121219-081520

**Badino, G., Celebrano, G. and Manino, A.** (1983). Population structure and Mdh- 1 locus variation in *Apis mellifera ligustica*. *J. Hered.* **74**, 443-446. doi:10.1093/oxfordjournals.jhered.a109834

**Benkert, P., Biasini, M. and Schwede, T.** (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343-350. doi:10.1093/bioinformatics/btq662

**Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Dinola, A. and Haak, J. R.** (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690. doi:10.1063/1.448118

**Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. and Schwede, T.** (2017). Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 10480. doi:10.1038/s41598-017-09654-8

**Birktoft, J. J., Fu, Z., Carnahan, G. E., Rhodes, G., Roderick, S. L. and Banaszak, L. J.** (1989). Comparison of the molecular structures of cytoplasmic and mitochondrial malate dehydrogenase. *Biochem. Soc. Trans.* **17**, 301-304. doi:10.1042/bst0170301

**Bjelic, S., Brandsdal, B. O. and Åqvist, J.** (2008). Cold adaptation of enzyme reaction rates. *Biochemistry* **47**, 10049-10057. doi:10.1021/bi801177k

**Chapman, A. D. M., Cortés, A., Dafforn, T. R., Clarke, A. R. and Brady, R. L.** (1999). Structural basis of substrate specificity in malate dehydrogenases: crystal structure of a ternary complex of porcine cytoplasmic malate dehydrogenase, α-Ketomalonate and TetrahydoNAD. *J. Mol. Biol.* **285**, 703-712. doi:10.1006/jmbi.1998.2357

**Childress, J. J. and Somero, G. N.** (1979). Depth-related enzymic activities in muscle, brain and heart of deep-living pelagic marine teleosts. *Mar. Biol.* **52**, 273-283. doi:10.1007/BF00398141

**Cornuet, J.-M., Oldroyd, B. P. and Crozier, R. H.** (1995). Unequal thermostability of allelic forms of malate dehydrogenase in honey bees. *J. Apic. Res.* **34**, 45-47. doi:10.1080/00218839.1995.11100885

**Crawford, D. L. and Powers, D. A.** (1989). Molecular basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish Fundulus heteroclitus. *Proc. Natl Acad. Sci. USA* **86**, 9365-9369. doi:10.1073/pnas.86.23.9365

**Crawford, D. L. and Powers, D. A.** (1992). Evolutionary adaptation to different thermal environments via transcriptional regulation. *Mol. Biol. Evol.* **9**, 806-813.

**Del Lama, M. A., Lobo, J. A., Soares, A. E. E. and Del Lama, S. N.** (1990). Genetic differentiation estimated by isozymic analysis of Africanized honeybee populations from Brazil and from Central America. *Apidologie* **21**, 271-280. doi:10.1051/apido:19900401

**Dong, Y. and Somero, G. N.** (2009). Temperature adaptation of cytosolic malate dehydrogenases of limpets (genus *Lottia*): differences in stability and function due to minor changes in sequence correlate with biogeographic and vertical distributions. *J. Exp. Biol.* **212**, 169-177. doi:10.1242/jeb.024505

**Dong, Y.-W., Liao, M.-L., Meng, X.-L. and Somero, G. N.** (2018). Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. *Proc. Natl. Acad. Sci. USA* **115**, 1274-1279. doi:10.1073/pnas.1718910115

**Endler, J. A.** (1986). *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

**Feller, G.** (2008). Enzyme function at low temperatures in psychrophiles. In *Protein Adaptation in Extremophiles* (ed. T. T. and K. S. Siddiqui), pp. 35-69. Hauppauge, NY: Nova science.

**Fields, P. A.** (2001). Review: protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 417-431. doi:10.1016/S1095-6433(00)00359-7

**Fields, P. and Somero, G.** (1997). Amino acid sequence differences cannot fully explain interspecific variation in thermal sensitivities of gobiid fish A4-lactate dehydrogenases (A4-LDHs). *J. Exp. Biol.* **200**, 1839-1850.

**Fields, P. A., Dong, Y., Meng, X. and Somero, G. N.** (2015). Adaptations of protein structure and function to temperature: there is more than one way to 'skin a cat'. *J. Exp. Biol.* **218**, 1801-1811. doi:10.1242/jeb.114298

**Gallivan, J. P. and Dougherty, D. A.** (1999). Cation-π interactions in structural biology. *Proc. Natl Acad. Sci. USA* **96**, 9459-9464. doi:10.1073/pnas.96.17.9459

**Gerstein, M. and Chothia, C.** (1991). Analysis of protein loop closure: two types of hinges produce one motion in lactate dehydrogenase. *J. Mol. Biol.* **220**, 133-149. doi:10.1016/0022-2836(91)90387-L

**Gromiha, M. M., Thomas, S. and Santhosh, C.** (2002). Role of cation-π interactions to the stability of thermophilic proteins. *Prep. Biochem. Biotechnol.* **32**, 355-362. doi:10.1081/PB-120015459

**Guex, N. and Peitsch, M. C.** (1997). Swiss-model and the Swiss-pdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723. doi:10.1002/elps.1150181505

**Hatty, S. and Oldroyd, B. P.** (1999). Evidence for temperature-dependent selection for malate dehydrogenase allele frequencies in honeybee populations. *J. Hered.* **90**, 565-568. doi:10.1093/jhered/90.5.565

**Hebert, P. D. N. and Beaton, M. J.** (1989). *Methodologies for Allozyme Analysis Using Cellulose Acetate Electrophoresis*. Beaumont: Helena Laboratories.

**Heinrich, B.** (1993). Hot-headed honeybees. In *The Hot-Blooded Insects: Strategies and Mechanisms of Thermoregulation*, pp. 292-322. Springer. doi:10.1007/978-3-662-10340-1

**Hess, B., Bekker, H., Berendsen, H. J. C. and Fraaije, J. G. E. M.** (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

**Hines, S. A., Philipp, D. P., Childers, W. F. and Whitt, G. S.** (1983). Thermal kinetic differences between allelic isozymes of malate dehydrogenase (*Mdh-B* locus) of largemouth bass, *Micropterus salmoides*. *Biochem. Genet.* **21**, 1143-1151. doi:10.1007/BF00488466

**Huang, J. and MacKerell, A. D., Jr**. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135-2145. doi:10.1002/jcc.23354

**Humphrey, W., Dalke, A. and Schulten, K.** (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38. doi:10.1016/0263-7855(96)00018-5

**Jaenicke, R. and Böhm, G.** (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738-748. doi:10.1016/S0959-440X(98)80094-8

**Kolafa, J. and Perram, J. W.** (1992). Cutoff errors in the ewald summation formulae for point charge systems. *Mol. Simulation* **9**, 351-368. doi:10.1080/08927029208049126

**Kumar, S. and Nussinov, R.** (2002). Close-range electrostatic interactions in proteins. *Chembiochem* **3**, 604-617. doi:10.1002/1439-7633(20020703)3:7<604::AID-CBIC604>3.0.CO;2-X

**Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M.** (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291. doi:10.1107/S0021889892009944

**Liao, M.-L., Zhang, S., Zhang, G.-Y., Chu, Y.-M., Somero, G. N. and Dong, Y.-W.** (2017). Heat-resistant cytosolic malate dehydrogenases (cMDHs) of thermophilic intertidal snails (genus *Echinolittorina*): protein underpinnings of tolerance to body temperatures reaching 55°C. *J. Exp. Biol.* **220**, 2066-2075. doi:10.1242/jeb.156935

**Liao, M.-L., Somero, G. N. and Dong, Y.-W.** (2019). Comparing mutagenesis and simulations as tools for identifying functionally important sequence changes for protein thermal adaptation. *Proc. Natl Acad. Sci. USA* **116**, 679. doi:10.1073/pnas.1817455116

**Mark, P. and Nilsson, L.** (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **105**, 9954-9960. doi:10.1021/jp003020w

**Minárik, P., Tomášková, N., Kollárová, M. and Antalik, M.** (2002). Malate dehydrogenases - structure and function. *Gen. Physiol. Biophys.* **21**, 257-265.

**Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N. and Arnold, F. H.** (2000). Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J. Mol. Biol.* **297**, 1015-1026. doi:10.1006/jmbi.2000.3612

**Nelson, D. L.** (2005). *Lehninger Principles of Biochemistry*, 4th edn. New York: W.H. Freeman.

**Nielsen, D., Page, R. E. and Crosland, M. W. J.** (1994). Clinal variation and selection of MDH allozymes in honey bee populations. *Experientia* **50**, 867-871. doi:10.1007/BF01956474

**Nunamaker, R. A., Wilson, W. T. and Haley, B. E.** (1984). Electrophoretic detection of Africanized honey bees *Apis mellifera scutellata* in Guatemala and Mexico based on malate dehydrogenase allozyme patterns. *J. Kans. Entomol. Soc.* **57**, 622-631.

**Oldroyd, B. P., Sheppard, W. S. and Stelzer, J. A.** (1992). Genetic characterization of the bees of Kangaroo Island, South Australia. *J. Apic. Res.* **31**, 141-148. doi:10.1080/00218839.1992.11101276

**Oldroyd, B. P., Cornuet, J.-M., Rowe, D., Rinderer, T. E. and Crozier, R. H.** (1995a). Racial admixture of *Apis mellifera* in Tasmania, Australia: similarities and differences with natural hybrid zones in Europe. *Heredity* **74**, 315-325. doi:10.1038/hdy.1995.46

**Oldroyd, B., Smolenski, A., Lawler, S., Estoup, A. and Crozier, R.** (1995b). Colony aggregations in *Apis mellifera L*. *Apidologie* **26**, 119-130. doi:10.1051/apido:19950205

**Olufsen, M., Smalås, A. O., Moe, E. and Brandsdal, B. O.** (2005). Increased flexibility as a strategy for cold adaptation: a comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase. *J. Biol. Chem.* **280**, 18042-18048. doi:10.1074/jbc.M500948200

**Pabis, A., Risso, V. A., Sanchez-Ruiz, J. M. and Kamerlin, S. C. L.** (2018). Cooperativity and flexibility in enzyme evolution. *Curr. Opin. Struct. Biol.* **48**, 83-92. doi:10.1016/j.sbi.2017.10.020

**Pierce, V. A. and Crawford, D. L.** (1997). Phylogenetic analysis of glycolytic enzyme expression. *Science* **276**, 256-259. doi:10.1126/science.276.5310.256

**Pucci, F. and Rooman, M.** (2017). Physical and molecular bases of protein thermal stability and cold adaptation. *Curr. Opin. Struct. Biol.* **42**, 117-128. doi:10.1016/j.sbi.2016.12.007

**Rank, N. E. and Dahlhoff, E. P.** (2002). Allele frequency shifts in response to climate change and physiological consequences of allozyme variation in a montane insect. *Evolution* **56**, 2278-2289. doi:10.1111/j.0014-3820.2002.tb00151.x

**Retamal, C. A., Thiebaut, P. and Alves, E. W.** (1999). Protein purification from polyacrylamide gels by sonication extraction. *Anal. Biochem.* **268**, 15-20. doi:10.1006/abio.1998.2977

**Siddiqui, K. S. and Cavicchioli, R.** (2006). Cold-adapted enzymes. *Annu. Rev. Biochem.* **75**, 403-433. doi:10.1146/annurev.biochem.75.103004.142723

**Somero, G. N.** (2010). The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *J. Exp. Biol.* **213**, 912-920. doi:10.1242/jeb.037473

**Tang, M. A. K., Motoshima, H. and Watanabe, K.** (2014). Cold adaptation: structural and functional characterizations of psychrophilic and mesophilic acetate kinase. *Protein J.* **33**, 313-322. doi:10.1007/s10930-014-9562-1

**Tayubi, I. A. and Sethumadhavan, R.** (2010). Nature of cation-π interactions and their role in structural stability of immunoglobulin proteins. *Biochemistry* **75**, 912-918. doi:10.1134/s000629791007014x

**Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L. et al.** (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296-W303. doi:10.1093/nar/gky427

**Wintrode, P. L. and Arnold, F. H.** (2001). Temperature adaptation of enzymes: lessons from laboratory evolution. In *Advances in Protein Chemistry*, Vol. 55 (ed. F. H. Arnold), pp. 161-225. Academic Press.