SUPPLEMENTARY INFORMATION

A covalent succinyl-cysteine-like intermediate in the enzyme-catalysed transformation of maleate to fumarate by maleate isomerase

Florian Fisch[1,2], Carlos Martinez Fleites[1], Marie Delenne[1], Nina Baudendistel[3], Bernhard Hauer[3], Johan P. Turkenburg[1], Sam Hart[1], Neil C. Bruce[2] and Gideon Grogan[1]*

[1]York Structural Biology Laboratory, Department of Chemistry, University of York, YO10 5YW York, United Kingdom

[2]Centre for Novel Agricultural Products, Department of Biology, University of York, YO10 5YW York, United Kingdom

[3]BASF SE, GVF/E, Carl-Bosch Strasse 38, 67056 Ludwigshafen, Germany

Correspondence should be addressed to N.C.B (ncb5@york.ac.uk) or G.G (grogan@ysbl.york.ac.uk)

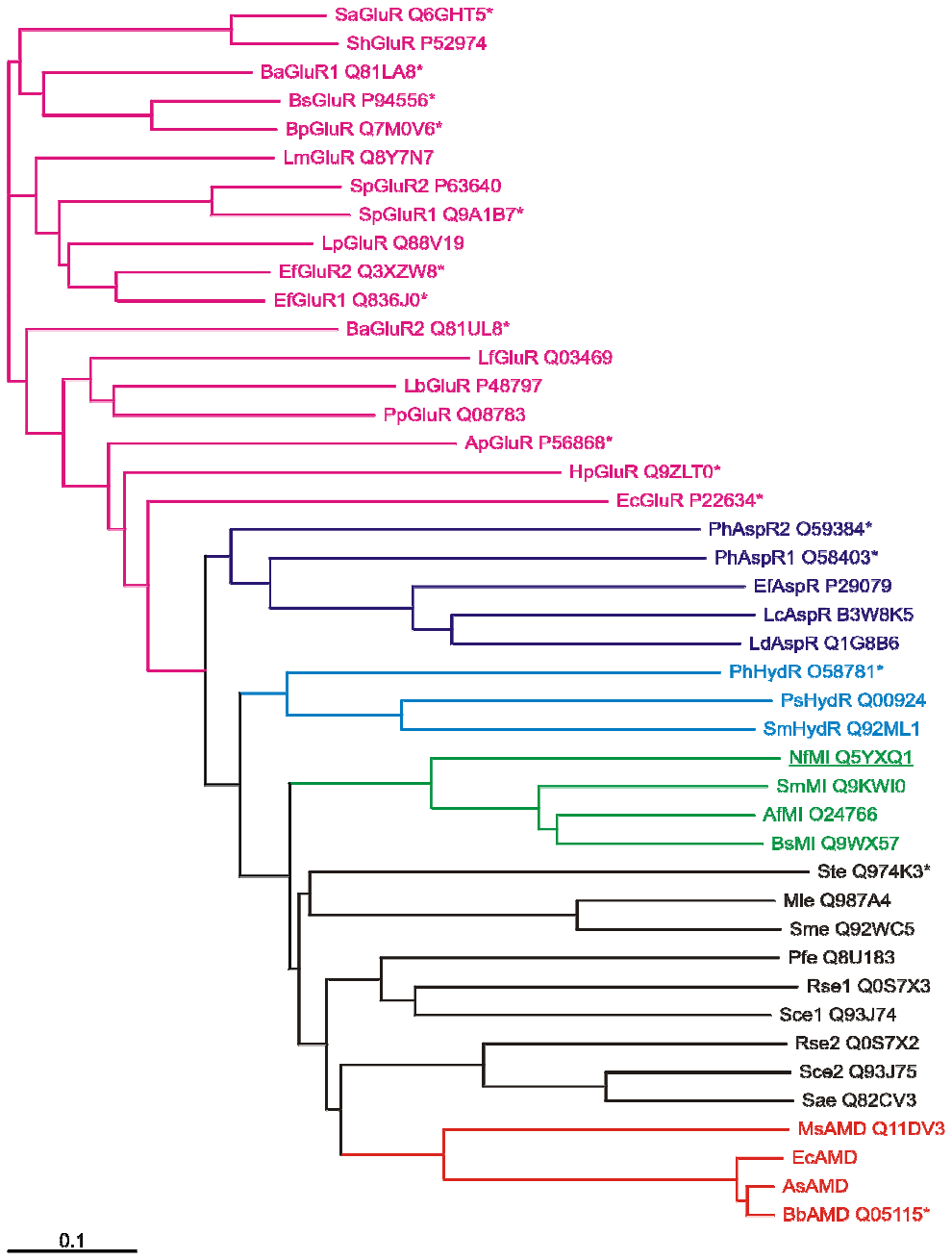

Figure S1. Phylogenetic tree of GluRs (magenta), AspRs (purple), HydRs (blue), AMDs (orange) and MI (green) with experimentally confirmed activities or known structures (*) and selected homologues that displayed no AMD activity in our assays. Sequences are labeled with the initials and UniProt accession numbers.

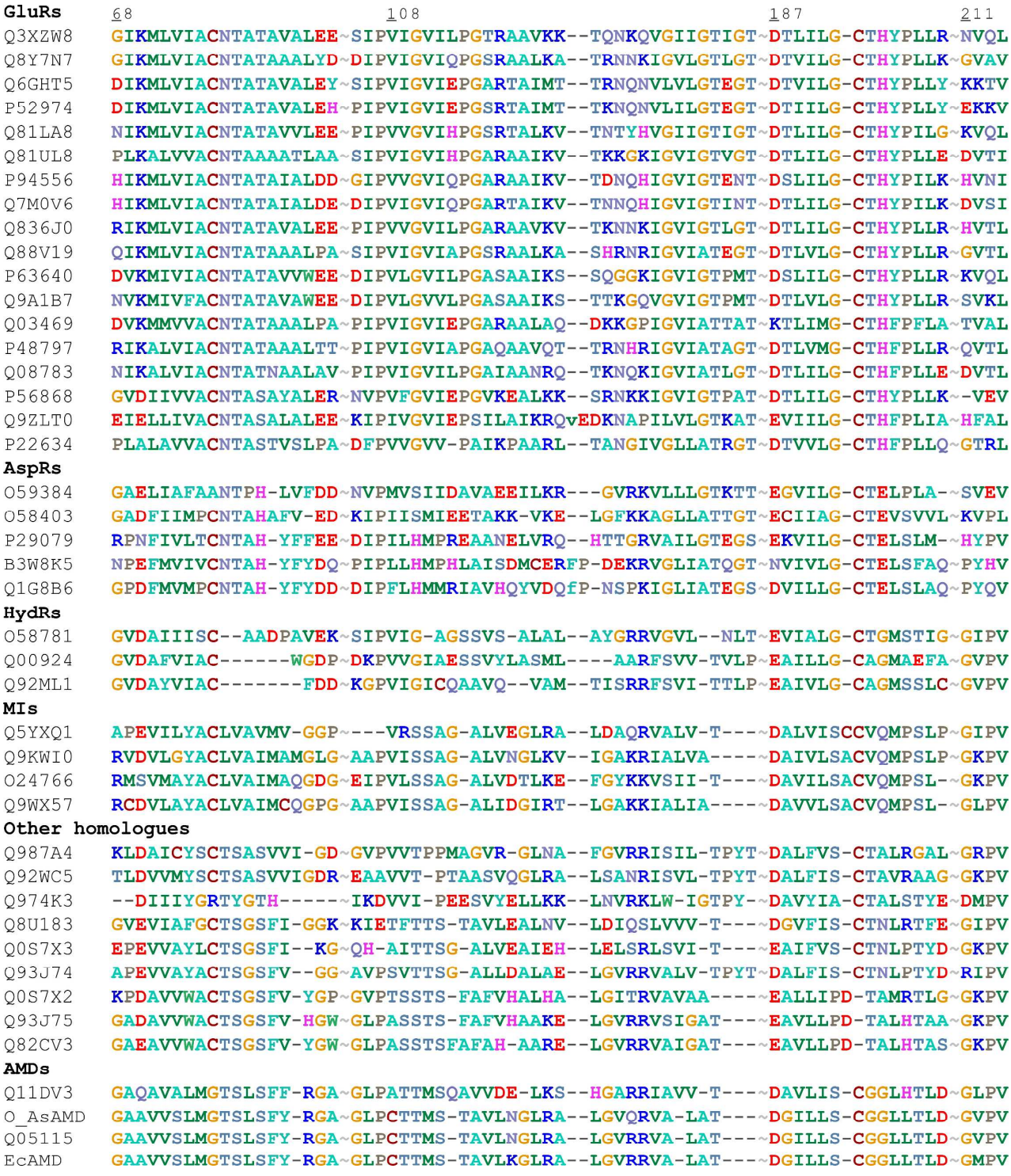

Figure S2. Sequence alignment of conserved regions used for the construction of the phylogenetic tree (numbered as in NfMI, UniProtID: Q5YXQ1), based on Pfam PF01177 entry with manual adjustments to account for the alignment of the catalytic residues. Sequences are labeled with UniProt accession numbers.

	WT	C194A mutant
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	50.1, 84.5, 238.1	53.0, 85.7, 239.9
Resolution (Å)	50.2-2.20	53.0-1.95
	(2.32-2.20)*	(2.06-1.95)*
R_{sym} or R_{merge}	0.083 (0.461)	0.099 (0.720)
$I / \sigma I$	9.6 (2.6)	9.1 (2.3)
Completeness (%)	99.6 (99.9)	99.4 (99.8)
Redundancy	4.3 (4.4)	5.9 (6.5)
Refinement		
Resolution (Å)	48.7-2.20	20.0-1.95
No. reflections	49570	75994
R_{work} / R_{free}	0.24 / 0.29	0.22 / 0.26
R_{pim}	0.044 (0.241)	0.044 (0.302)
No. atoms		
Protein	7240	7309
Ligand/ion	18	32
Water	119	635
B-factors		
Protein	54	34
Ligand/ion	51	31
Water	48	43
R.m.s. deviations		
Bond lengths (Å)	0.017	0.010
Bond angles (°)	1.7	1.3

*Values in parentheses are for highest-resolution shell.

Table S1. Data collection and refinement statistics for NfMI.

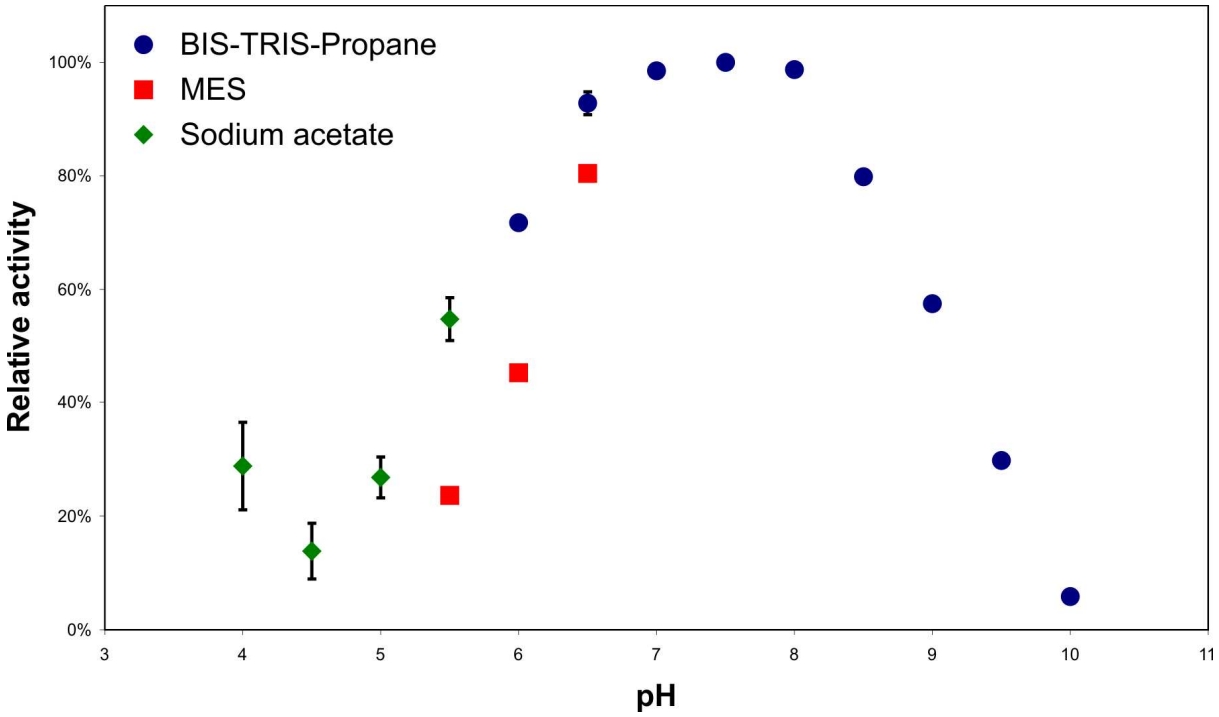

Figure S3. Dependence of NfMI activity on pH. Spectrophotometric assay using 20 µg mL^{-1} of purified NfMI with 10 mM maleate using various buffers. The reaction was performed for 20 min at 37 ℃. All experiments were performed in triplicate and the error bars represent standard deviation.

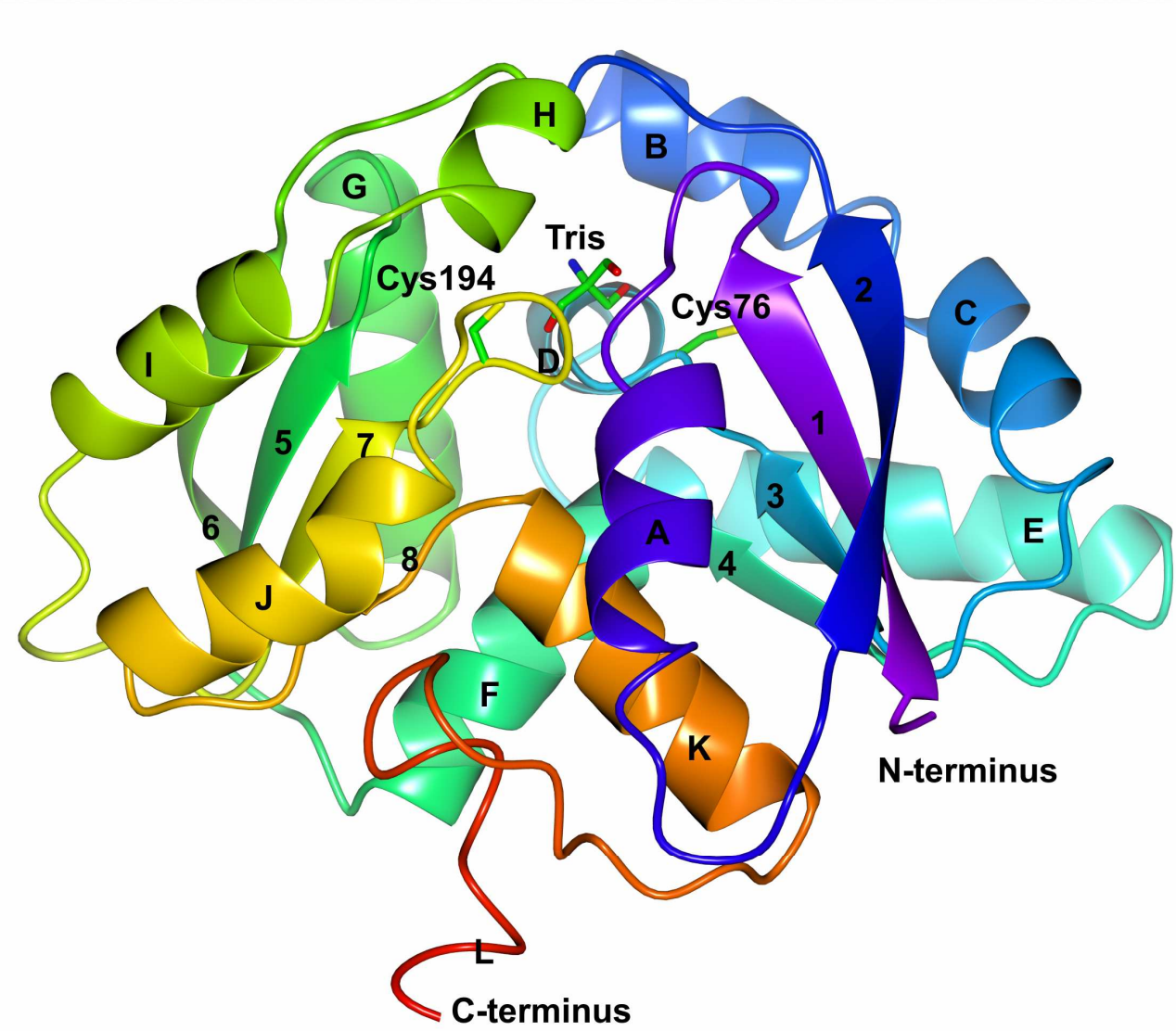

Figure S4. Three-dimensional structure of monomer of maleate isomerase from *Nocardia farcinica,* coloured violet (N-terminus) through to red (C-terminus) with sequentially numbered secondary elements (helices A-K; strands 1-8).

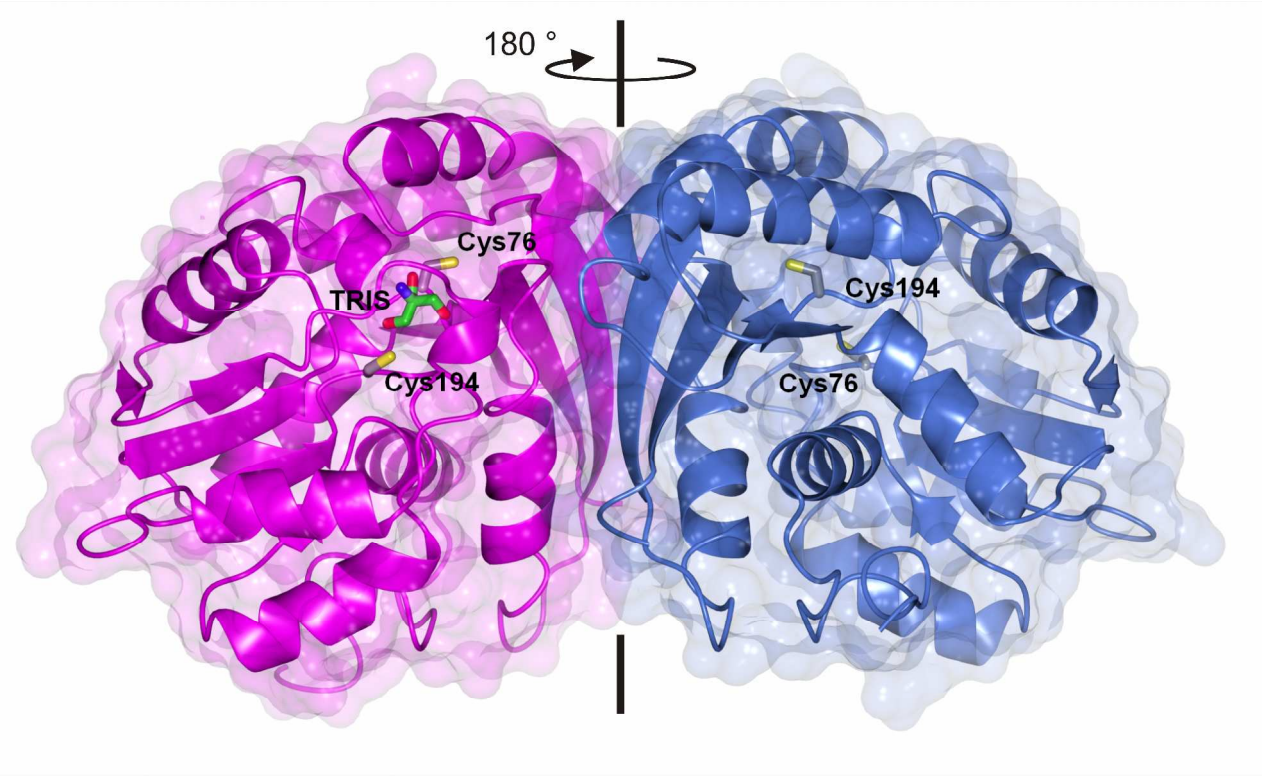

Figure S5. Dimer of NfMI observed in the crystal formed by molecules A and B. TRIS (green) is bound to the active site of molecule A.

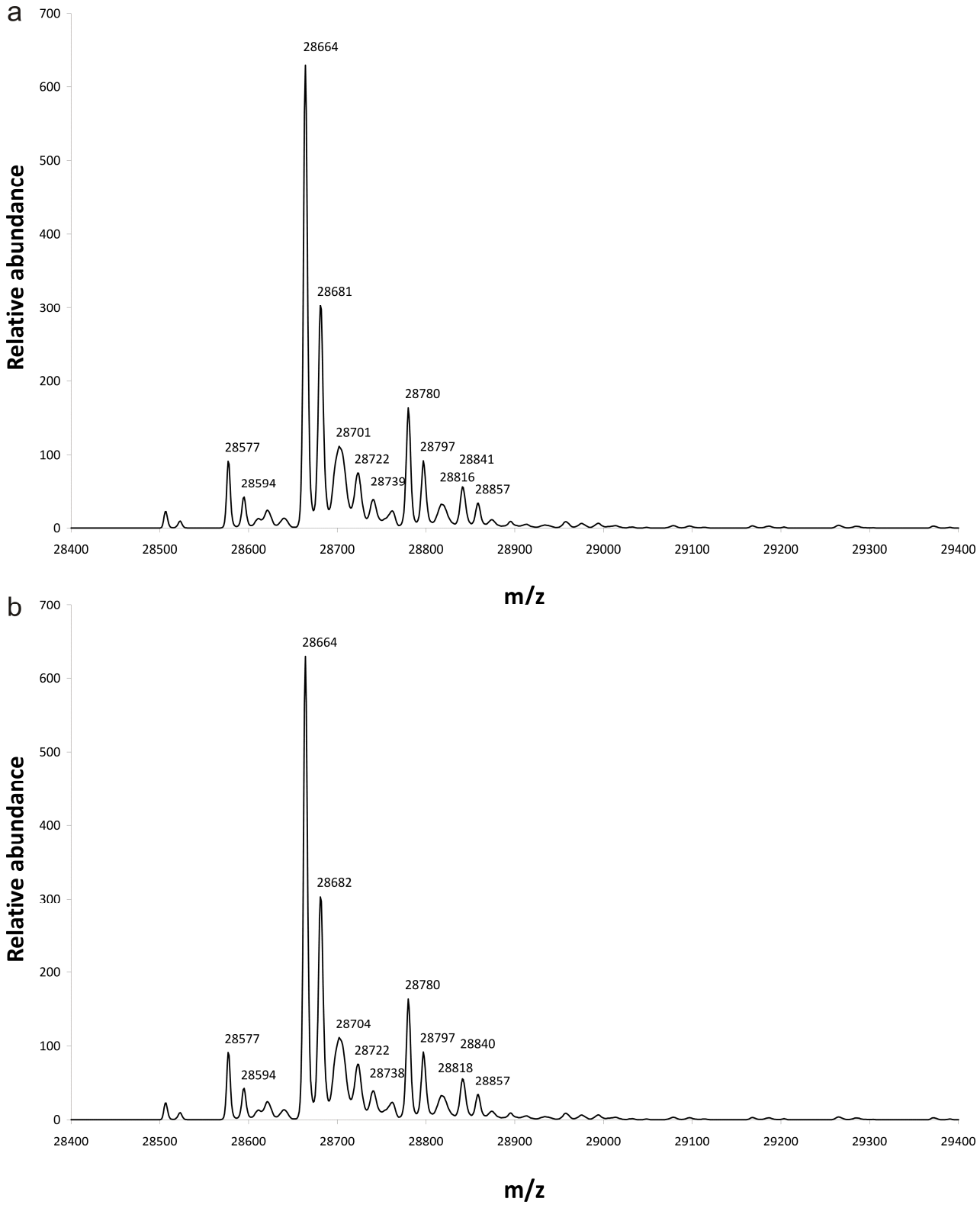

Figure S6 ESI-Mass spectra of ligand protein complexes. NfMI C194A was incubated overnight with 50 mM maleate (**a**) or 50 mM fumarate (**b**) at pH 6.5 and analyzed by ESI-TOF.

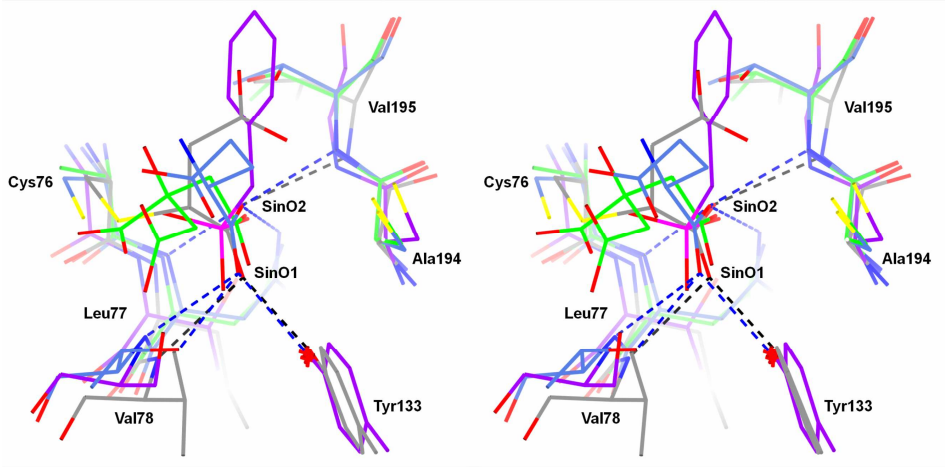

Figure S7. Stereoview of superposition of dioxyanion holes in Asp/Glu superfamily structures. Conserved positions of Val76Cα, Tyr133Oη/H_2O, Ala195Cα, Ala195Cβ and Val195N (numbering as in NfMI C194A mutant) of several Asp/Glu racemase superfamily members containing ligands were overlaid. NfMI C194A with succinyl-like intermediate (grey), *Bacillus subtilis* GluR containing D-Glu (blue, PDB: 1ZUW), *Pyrococcus horikoshii* AspR (green, PDB: 2DX7) containing citrate and *Bordetella bronchiseptica* AMD containing benzylphosphonate (purple, PDB: 3IP8) are shown. Six hydrogen bonds of GluR involved in enediolate stabilization are shown in blue broken line and the three conserved hydrogen bonds in NfMI are shown as broken black lines.

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
C193A	1.5 ± 0.6	3.1 ± 1.3	4.8×10^5
C193S	0.086 ± 0.001	0.36 ± 0.08	2.4×10^4
C193A/C194S	$2.5 \times 10^{-4} \pm 0.5 \times 10^{-4}$	3.5 ± 0.5	71
C193S/C194A	-*	N.D.	-
Y133F/C194S	$4.0 \times 10^{-4} \pm 0.0 \times 10^{-4}$	1.9 ± 0.1	211

Table S2 Summary of the kinetic parameters obtained for additional NfMI active site mutants. * Activity of C194A was beneath measurable limits of detection. N.D. not determined. Steady state kinetic constants were determined by an HPLC assay performed at the NfMI pH optimum of 7.5. The data correspond to mean ± maximal error of two independent constants determinations.

METHODS

Cloning and Expression

Genomic DNA from *Nocardia farcinica* IFM 10152 was a gift from Dr Jun Ishikawa of the National Institute of Infectious Diseases, Tokyo, Japan. The gene encoding NfMI was amplified from the genomic DNA using the following oligonucleotide primers: 5'-CCAGGGACCAGCAATGGGCATCCGCCGCATCGGGTTG-3' (Forward) and 5'-GAGGAGAAGGCGCGTCAGCTCGCGGTGACGGCGGAG-3' (Reverse) and the PCR product cloned into the pET-YSBLIC-3C vector using the published protocol **(1)**. The resulting plasmid was used to transform *E. coli* Rosetta2 (DE3) (Novagen) grown on Luria-Bertani (LB) agar containing 100 μg mL^{-1} kanamycin and 34 μg mL^{-1} chloramphenicol. For gene expression, transformants were used to inoculate 5 mL starter cultures, which were grown in LB medium containing 100 μg mL^{-1} kanamycin and 34 μg mL^{-1} chloramphenicol and shaken at 200 r.p.m. overnight at 37 ℃. One starter culture was then used to inoculate 500 mL of LB medium in a 2L Erlenmeyer flask, which was then grown at 37 ℃ in an orbital shaker until the cell suspension had achieved an OD_{600} of approximately 0.8. At this point 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added and the culture was incubated with shaking overnight at 30 ℃. Cells were then harvested by centrifugation at 2,000 × g for 10 min. The cell pellets were resuspended in a buffer (Buffer 'A') composed of 20 mM TRIS and 150 mM NaCl at pH 8.0 supplemented with 200 μM phenylmethylsulfonyl fluoride (PMSF), 5 mM dithiothreitol (DTT), 5 μg mL^{-1} hen egg white lysozyme, 5 μg mL^{-1} DNase I and 1 mM $MgCl_2$. The cell suspension was then lysed using an S-4000 Sonicator (Misonix) at 70% amplitude during for a total duration of 4 min in 30 s intervals interrupted by 7 s cooling periods on ice. The lysate was then cleared of cell debris by centrifugation at 45,000 × g for 20 min.

Purification and Crystallization of NfMI

The clarified cell lysate was filtered through a 0.22 µm membrane (Millipore) and loaded with a syringe onto a 5 mL HisTrap HP nickel column (GE Healthcare) that had been pre-equilibrated with buffer consisting of 20 mM TRIS and 150 mM NaCl (pH 8.0). The column was then washed with a linear concentration gradient of imidazole from 0 to 50 mM, for 100 mL, and NfMI then eluted by increasing the imidazole concentration from 50 to 300 mM for another 100 mL. Fractions determined to contain NfMI, as established by sodium dodecylsulfate polyacrylamide gele electrophoresis (SDS-PAGE) were pooled and concentrated to a total volume of 2 mL using Amicon Ultra15 centrifugal filters of molecular weight cut-off of 10 kDa (Millipore). The concentrated protein sample was then loaded onto into a HiLoad 16/60 Superdex 75 preparative grade gel filtration column (GE Healthcare) and eluted with buffer consisting of 20 mM TRIS and 150 mM NaCl (pH 8.0). Fractions displaying absorbance peaks at 280 nm were pooled and the total volume reduced using centrifugal filters until a final protein concentration of 15-22 mg mL^{-1} was achieved (as determined using the protein estimation reagent from BioRad). Aliquots of NfMI were flash cooled in liquid nitrogen and stored at -80 °C. The monodispersity and oligomerisation state of NfMI were analysed by dynamic light scattering on a DynaPro particle sizer (Wyatt Technology) at 18°C in order to obtain an indication of the oligomeric state of the protein in solution. The protein was diluted to 1 mg mL^{-1} in buffer and centrifuged for 10 min at 18,000 × g to remove large particles. The samples were measured until 15 statistically good data points were recorded (indicated by the software). With 14% polydispersity, the sample was considered monodisperse. The hydrodynamic radius was determined to be 3.2 nm which is equivalent of an approximate molecular mass of 51 kDa corresponding assuming a spherical protein. This would correspond to a dimer of NfMI with a molecular mass of 53 kDa.

Crystallization trials were performed with 96-well sitting drop vapor diffusion plates (MRC-Wilden) using the buffer screens PACT Premier (Molecular Dimension Ltd), Index (Hampton Research), CSS 1 and 2 (with MES pH 6.0 and TRIS pH 8.0 **(2)**, and PEG Ion/Ion 2 (without buffers, Hampton Research) using a reservoir volume of 60 µL. Drops of 150 nL protein solution and 150 nL reservoir solution respectively were dispensed using a Mosquito robot (TTP Latech). Positive hits were scaled up and optimized as 1 µL protein solution and 1 µL reservoir solution using the hanging drop vapour diffusion technique on siliconised glass cover slides over 24-well cell culture plates (Greiner) using 1 mL reservoirs and incubated at 25 °C. The best crystals were found in drops above wells containing 300 mM calcium acetate, 100 mM TRIS Cl (pH 7.5), 5 mM sodium maleate, 16% (w/v) PEG 3350 for the wild-type (WT) NfMI and 20 mM ammonium formate, 100 mM MES (pH 6.5) 20% (w/v) PEG 3350, 50 mM sodium maleate (premixed with the protein) for the C194A mutant. The crystals were flash-cooled in liquid nitrogen with 30% (w/v) PEG 3350 as cryoprotectant for WT NfMI. No cryoprotectant was required for crystals of the C194A mutant.

Data Collection, structure solution, building and refinement

Crystals displaying promising diffraction, as assessed using in-house X-ray detectors, were sent to the European Synchrotron Radiation Facility (ESRF) in Grenoble, and data collected on beamlines ID23-1 at 0.976 Å (WT) and ID14-4 at 0.977 Å (C194A) at a temperature of 100 K (**Table S1**). The synchrotron data were indexed and integrated with imosflm **(3)** and the intensities were scaled and reduced with SCALA **(4)**. The structure phases of WT were solved by molecular replacement using the coordinates of *Bordetella bronchiseptica* AMD (PDB: 3DG9) as a search model using Phaser **(5)**. The phases were

improved by NCS averaging and solvent flattening using DM **(6)**. The initial model was refined with Phenix **(7)** using simulated annealing to reduce model bias followed by several cycles of manual model building in COOT **(8)** and refinement with Phenix. The C194A structure was solved by molecular replacement using Phaser with the WT structure as the search model. The final models were refined using REFMAC **(9)** using TLS corrections for each protein chain. Water molecules were placed automatically by COOT and ligands were placed after refinement. In C194A the distance between C2 of succinyl and Sγ of Cys76 was close to 1.8 Å. The final structure was refined introducing a covalent link between the two atoms. Details of the data collection, processing and refinement are given in **Table S1**. Ramachandran statistics were produced with Procheck **(10)**, which indicated for the residues of WT structure that 89.8% were in favored regions, 10.0% in allowed regions, 0.1% in generously allowed regions and 0.0% in disallowed regions. In C194A 93.4% were in favored regions, 6.6% in allowed regions, 0.0% in generously allowed regions and 0.0% in disallowed regions. Coordinates and structure factors have been deposited with the Protein Data Bank with the accession numbers 3XEC (WT) and 3XED (C194A).

Cloning, Expression and Purification of NfMI mutants

Mutants were produced following the QuickChange protocol of Stratagene (http://www.stratagene.com/manuals/200518.pdf). The primers used were:

5'-GTGATCCTCTACGCCAGCCTGGTCGCGGTC-3'	for	C76S,
5'-CTGGTGATCTCCTGCGCGGTGCAGATGCCCTCG-3'	for	C194A,
5'-GGTGATCTCCTGCAGCGTGCAGATGCC-3'	for	C194S,
5'-CTGGTGACCCCGTTTATGCGCCCGCTCG-3'	for	Y133F,

5'-GCTGGTGATCTCCGCGTGCGTGCAGATGC-3' for C193A,

5'-CGCTGGTGATCTCCAGCTGCGTGCAGATG-3' for C193S,

5'-GTGATCTCCGCGAGCGTGCAGATGC-3' for C193A/C194S,

5'-CGCTGGTGATCTCCAGCGCGGTGCAGATG-3' for C193S/C194A

and the corresponding reverse complements.

Mutant strands were produced using 2 U Phusion High Fidelity DNA Polymerase (Finnzymes), 1x HF buffer, 0.22 mM of each dNTP, 0.22 µM of both forward and reverse primer (Sigma Genosys) and 3.4 ng μL^{-1} original plasmid DNA in a total reaction volume of 50 µL. The temperature program was 98 °C for 30 s followed by 16 cycles of 98 °C for 30 s, 55 °C for 60 s and 72 °C for 370 s. The original DNA was digested specifically with 10 U DpnI (NEB) for 3 h at 37 °C and 2 µL of the digest mix were used to directly transform aliquots of NovaBlue Singles competent cells of *E. coli* (Novagen). DNA sequencing was used to verify that the required mutations had been incorporated.

WT NfMI and its mutants were expressed in parallel as described above in 50 mL cultures, harvested by centrifugation and resuspended in 2.5 mL BugBuster (Novagen) containing 200 µM PMSF, 5 µg mL^{-1} DNase I, 5 µg mL^{-1} hen egg white lysozyme, 1 mM $MgCl_2$ and 5 mM DTT, then incubated for 15 min at 4 °C with shaking. The lysates were diluted with 14 mL Buffer 'A' and added to 1 mL of fresh His-select resin (Sigma). Fresh resin was used in each case for the purification of WT NfMI and each of its mutants. The resin samples with bound protein were washed in three consecutive steps with 15 mL buffer containing 0, 10, 20 mM imidazole and finally eluted with 15 mL 100 mM imidazole and cleared through 22 µm syringe filter (Millipore). Imidazole was removed by

concentrating the sample to 500 µL using 10 kDa molecular weight cut-off centrifuge filters (Millipore) and diluting in 15 mL buffer without imidazole four times; diluting first with 5 mM DTT, then twice without reducing agent. Last, TCEP was added to a final concentration of 5 mM. After the final concentration step, the concentrations of the protein samples were adjusted to 1 mg mL^{-1}.

Kinetic Analysis of NfMI and mutants

Samples of 1, 3, 5, 10, 30, 50, 100, 150 and 200 µM maleic acid were incubated with purified enzyme at concentrations of 14 ng mL^{-1} for WT and C193A, 56 ng mL^{-1} for Y133F, 560 ng mL^{-1} for C193S and 3800 ng mL^{-1} for all other mutants. Each reaction was buffered with 50 mM HEPES (pH 7.5) at 30 °C, in a total volume of 1 mL. Reactions were initiated by the addition of enzyme, and samples were then incubated without shaking at 30ºC. 200 µL samples were taken during the linear phase of reaction, at 6 and 12 min for reactions containing WT, C193A, C193S and Y133F and at 10 and 70 min for the other mutants. Samples were quenched by adding them each to a mixture of 10 µL concentrated hydrochloric acid plus 10 µL methanol. Samples were then analysed by HPLC: 10 µL of the mixture were injected onto a reverse phase Onyx Monolithic C18 column (100 mm x 4.8 mm, Phenomenex) and eluted with a solvent mixture of 95:5 (v/v) 25 mM KH_2PO_4 pH 2.5 / methanol at a flow rate of 1 mL min^{-1}. In the case of each experiment, the average of the rate of fumarate formation from two separate runs was plotted against substrate concentration and a simple square hyperbola was fitted ($v = k_{cat}\cdot[S]/(K_M+[S]$ or, when considerations of substrate inhibition were appropriate $v = k_{cat}\cdot[S]/(K_M+[S]\cdot(1+[S]/K_i))$).

Effect of pH on NfMI Activity

To determine the pH optimum of NfMI the amount of fumarate produced from 10 mM maleate after 20 min was measured in three different buffers with pH ranging from 4 to 10. The maximal activity was observed at pH 7.5 being approximately 1 pH point lower than those measured for other known maleate isomerases. At pH 4.5 and 11.0 the activity was almost completely abolished (Figure S3).

Sequence Alignment and Phylogenetic Tree

A selection of Asp/Glu racemase superfamily sequences was created including sequences with published three dimensional structures, published activities and all sequences tested by us. The sequence alignment from Pfam (PF01177) was used and the misalignment of the first catalytic cysteine with the glycine of AMDs was corrected by hand. The phylogenetic tree was produced using the ClustalW2 tool at EBI (http://www.ebi.ac.uk/Tools/clustalw2/) using the Neighbor-joining method.

Electrospray Ionisation Mass Spectrometry (ESI-MS) of Enzyme-Ligand Adducts

A sample of 50 µL of C194A protein solution was mixed with 6 µL 500 mM maleate or fumarate (adjusted to pH 7 withsodium hydroxide), 56 µL of a 20 mM solution of ammonium formate in 100 mM MES (pH 6.5) and incubated overnight at room temperature. The buffer of NfMI was replaced by several rounds of dilution in 20 mM ammonium acetate (pH 6.5) and subsequent concentration with Amicon Ultra 15 centrifuge filters with a 10 kDa molecular weight cut off (Millipore). The enzyme was finally concentrated to 3 mg mL^{-1} and diluted into a acetonitrile / water / formic acid

50:50:0.1% (v/v) solution and analyzed on an ABI Qstar ESI-TOF mass spectrometer (Applied Biosystems).

REFERENCES

1. Bonsor, D. et al. *Org. Biomol. Chem.* **4**, 1252-60 (2006).
2. Brzozowski, A.M. & Walton, J. *Journal of Applied Crystallography* **34**, 97-101 (2001).
3. Leslie, A.G.W. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **No. 26**(1992).
4. COLLABORATIVE COMPUTATIONAL PROJECT, N. *Acta Cryst.* **D50**, 760-763 (1994).
5. A. J. McCoy, R.W.G.-K., P. D. Adams, M. D. Winn, L.C. Storoni and R.J. Read. . *J. Appl. Cryst.* **40**, 658-674 (2007).
6. Cowtan, K.D. & Zhang, K.Y.J. *Progress in Biophysics & Molecular Biology* **72**, 245-270 (1999).
7. P.D. Adams, R.W.G.-K., L.-W. Hung, T.R. Ioerger, A.J. McCoy, N.W. Moriarty, R.J. Read, J.C. Sacchettini, N.K. Sauter and T.C. Terwilliger. *Acta Cryst.* **D58**, 1948-1954 (2002).
8. Emsley, P. & Cowtan, K. *Acta Cryst.* **D60**, 2126-2132 (2004).
9. Winn, M., Isupov, M. & G.N.Murshudov. *Acta Cryst.* **D57**, 122-133 (2000).
10. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & MThornton, J. *J. Appl. Cryst.* **26**, 283-291 (1993).