Macrocycle peptides delineate locked-open inhibition mechanism for microorganism phosphoglycerate mutases

Hao Yu1, Patricia Dranchak2, Zhiru Li3, Ryan MacArthur2, Matthew S. Munson4, Nurjahan Mehzabeen5, Nathan J. Baird6, Kevin P. Battalie7, David Ross4, Scott Lovell5, Clotilde K.S. Carlow3, Hiroaki Suga1 & James Inglese2

Glycolytic interconversion of phosphoglycerate isomers is catalysed in numerous pathogenic microorganisms by a cofactor-independent mutase (iPGM) structurally distinct from the mammalian cofactor-dependent (dPGM) isozyme. The iPGM active site dynamically assembles through substrate-triggered movement of phosphatase and transferase domains creating a solvent inaccessible cavity. Here we identify alternate ligand binding regions using nematode iPGM to select and enrich lariat-like ligands from an mRNA-display macrocyclic peptide library containing $\approx 10^{12}$ members. Functional analysis of the ligands, named ipglycermides, demonstrates sub-nanomolar inhibition of iPGM with complete selectivity over dPGM. The crystal structure of an iPGM macrocyclic peptide complex illuminated an allosteric, locked-open inhibition mechanism placing the cyclic peptide at the bi-domain interface. This binding mode aligns the pendant lariat cysteine thiolate for coordination with the iPGM transition metal ion cluster. The extended charged, hydrophilic binding surface interaction rationalizes the persistent challenges these enzymes have presented to small-molecule screening efforts highlighting the important roles of macrocyclic peptides in expanding chemical diversity for ligand discovery.
Nematode worms are the most abundant animal on earth and are found in widely different environments. They can be free-living or parasitic, infecting plants, animals and humans. Parasitic nematode infection in humans can lead to a number of devastating diseases. Lymphatic filariasis and onchocerciasis are neglected tropical diseases caused by filarial nematode parasites that are transmitted to humans by insects. Collectively, they afflict 150 million people in over 80 countries and threaten the health of over 1.5 billion. These infections are responsible for extreme infirmity, social stigma and severe economic consequences. The lymphatic dwelling parasites such as Wuchereria bancrofti and Brugia malayi are the cause of lymphedema, hydrocele and in the most extreme cases, elephantiasis. Infection with Onchocerca volvulus results in severe dermatitis and blindness. The mainstay of filarial disease control for several years has been a limited number of drugs, predominantly ivermectin together with albendazole (where onchocerciasis is endemic) or diethylcarbamazine citrate (where onchocerciasis is not present). These compounds mainly target the larval stages and require annual or semi-annual administration. Furthermore, there are reports of drug resistance emerging.

Enzymes essential for nematode survival but absent from humans represent potential targets for intervention. Essential nematode genes have been identified using comparative genomic studies of the free-living nematode Caenorhabditis elegans based on an algorithm designed to evaluate criteria such as Homo sapiens homology and life stage expression profile. As a result several novel drug targets in filarial parasites have been proposed. Among the highest ranking is cofactor-independent phosphoglycerate mutase (iPGM) (EC 5.4.2.1). Silencing of ipgm in C. elegans leads to nematode death, demonstrating the importance of this enzyme in nematode viability and, therefore, its potential as an anthelmintic drug target.

Phosphoglycerate mutases (PGMs) catalyse the interconversion of 2- and 3-phosphoglycerate (2-PG; 3-PG) in the glycolytic and gluconeogenic pathways. Although these pathways are highly conserved among different organisms, two distinct PGM isoenzymes are known to exist, namely iPGM and cofactor-dependent phosphoglycerate mutase (dPGM). The enzymes have no amino-acid sequence similarity and differ in their mechanism of catalysis (Fig. 1a). iPGM is comprised of ~510 amino acids and facilitates the intramolecular transfer of the phosphoryl group on the monophosphoglycerates through a phosphoserine intermediate and is the sole PGM in nematode. In contrast, dPGM, the isoform found in humans is comprised of ~250 amino acids, and catalyses the intermolecular transfer of the phosphoryl group between the monophosphoglycerates and the cofactor (2,3-diphosphoglycerate) through a phosphorylhistidine intermediate. While the two forms of PGM are distinct isoenzymes, the amino-acid sequence of each isozyme family is conserved, when present, from bacteria to higher eukaryotes. The completely distinct structures and catalytic mechanism of iPGM and dPGM enzymes offer great promise for the discovery of inhibitors with high selectivity for the nematode enzymes. Furthermore, the high similarity in primary sequence and catalytic properties among the iPGMs suggests that a single inhibitor could be effective against a range of parasitic and microbial enzymes.

Here we report a series of cyclic peptides and analogues that exhibit potent and isoform-selective inhibition against iPGM orthologues. The parental peptide, we named ipglycermide, was discovered from a library containing over a trillion cyclic peptide members, each displayed on a cognate mRNA template. Ipglycermide has a unique lariat structure, where the ring peptide consists of eight amino-acid residues, one of which is D-tyrosine, closed by a thioether bond and a seven residue pendant extension which places l-cysteine (Cys) at the C terminus. Our 1.95 Å co-crystal structure reveals allosteric binding to a heretofore unknown site rich in charged side chains but devoid of a large hydrophobic surface. Ipglycermide binding constrains the dynamic movement between the phosphatase and transferase protein domains required for catalysis.

Our study suggests that the chemical diversity afforded by nucleic acid encoded cyclic peptides effectively maps putative, or even previously unknown, protein surfaces for drug discovery where small-molecule screening probes have failed. In addition, we show that macrocyclic peptides capture dynamic protein domains to allosterically impose constraints on functional motions. Overall, ipglycermide inhibition of iPGM will serve as an effective model for discovery and inhibition of the many challenging targets lacking deep hydrophobic binding pockets, and instead principally characterized by conformationally flexible hydrophilic surfaces.

Results

RaPID display affinity selection. A recently reported attempt to obtain small-molecule inhibitors against iPGM from a combined library of 380,000 compounds by Genzyme Corporation and the National Center for Drug Screening in Shanghai resulted in only two low-potency compounds, apparent metal ion chelators. Given the evident refractory nature of iPGM toward small-molecule inhibition outside of metal ion ligands, and the difficulty of identifying chemotypes of the alkaline phosphatase superfamily enzyme class from high-throughput screening (HTS) with sub-micromolar optimization potential, we approached inhibitor identification through the complementary method of affinity selection using an in vitro display system, referred to as RaPID (random nonstandard peptides integrated discovery).

The RaPID system (Fig. 1b) enabled us to exploit the diverse molecular topology of macrocyclic peptide populations numbering in a trillion unique members and enrich for and amplify low abundance, high-affinity ligands. We utilized a thioether-cyclic peptide library initiated with both L- or D-N-chloroacetyl tyrosine and performed ligand selection against iPGM protein from two nematode species, B. malayi and C. elegans, which were individually immobilized on magnetic beads via the His6 tag at the C terminus of these recombinant enzymes. The sequence alignments (Supplementary Fig. 1) from 69 RaPID-derived clones resulted in 11 independent sequence families, Bm-1–7 and Ce-1–4 obtained from panning B. malayi and C. elegans iPGM, respectively, and corresponding to macrocyclic peptides of a lariat structure with ring sizes ranging between 7–13 amino acids and C-terminal tails of 1–7 amino acids (Table 1).

It should be noted that the cyclic peptides as isolated by RaPID are tethered at their carboxyl terminus via puromycin to the encoding mRNA (Fig. 1b). Any effect of the tethered nucleic acid during cyclic peptide binding to their target, either to facilitate binding or block possible productive target-cyclic peptide interactions is an inherent property of mRNA-display technology. Significant binding contributions made via the nucleic acid will not be present in samples made by the solid-phase peptide synthesis step.

Functional evaluation of cyclic peptide PGM inhibitors. To efficiently profile the activity of the cyclic peptides derived from in vitro selection, several PGM enzymes from a range of species were evaluated, including the parasite target, B. malayi iPGM and filarial orthologues (Onchocerca volvulus, Dirofilaria immitis), the corresponding model organism C. elegans iPGM...
Incorporation of a cysteine during translation results in macrocyclization via thiolate nucleophilic attack on the chloroacetyl electrophile. After incubating charged initiator tRNA and 19 proteogenic amino acids (grey spheres), methionine is excluded as its tRNA is charged with the chloroacetyl tyrosine.

For the direct evaluation of cyclic peptides on the target enzyme, we used gradient elution moving boundary electrophoresis (GEMBE)\(^{19,20}\) to enact an electrophoretic separation of enzyme, we used gradient elution moving boundary electrophoresis (Supplementary Fig. 3; Supplementary Table 2).

The peptide sequences deduced from the corresponding tethered nucleic acid were chemically synthesized as cyclic peptides in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1).

For the direct evaluation of cyclic peptides on the target enzyme, we used gradient elution moving boundary electrophoresis (GEMBE)\(^{19,20}\) to enact an electrophoretic separation of 3-PG and 2-PG\(^{21}\) (Supplementary Figs 2b and 4). The method provides a direct, label-free measurement of the substrate and pyruvate kinase, respectively\(^{16}\). A kinetic absorbance output through phosphoenol pyruvate to pyruvate and ATP via enolase was achieved using lactate dehydrogenase-mediated changes in NADH concentration through pyruvate conversion to lactate was measured using lactate dehydrogenase-mediated changes in NADH concentration through pyruvate conversion to lactate (Supplementary Fig. 2a, step 3a)\(^{17,18}\). For the bioluminescent end point assay the ATP produced from the coupled enzymatic reactions is used by Firefly luciferase and luciferin to generate measurable light (Supplementary Fig. 2a, Step 3b; Supplementary Table 1).

The peptide sequences deduced from the corresponding tethered nucleic acid were chemically synthesized as cyclic peptides in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1).

Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1). Concentration response curves obtained for \(\text{Bm-1} - 7\) across the PGM panel primarily revealed a selective, modestly potent macrocyclic series. All except \(\text{Bm-2}\), (for which RaPID selection may have been influenced by the tethered nucleic acid not present in the resynthesized cyclic peptide) inhibited the iPGM orthologues in sufficient quantities for evaluation as inhibitors of the enzymatic activity of a panel of seven PGM orthologues and isozymes using the bioluminescence endpoint assay (Table 1).
Table 1 | PGM panel inhibitory activity of RaPID selected, chemically resynthesized peptides.

| Peptide ID | Sequence | Ring size/tail length | pIC50, apparent |
|------------|----------|-----------------------|----------------|
| Bm-1 | Ac-DYSPWNPAPWEIWKCCG-NH₂ | 12/2 | C. elegans IPGM 5.30, B. malayi IPGM 5.89, O. volvulus IPGM 5.88, D. immitis IPGM 5.99, E. coli IPGM 5.31, H. sapiens dPGM NA, E. coli dPGM NA |
| Bm-2 | Ac-DYDRTLTPWLKRAHCAG-NH₂ | 13/1 | NA, NA, NA, NA, NA, NA |
| Bm-3 | Ac-DYQNSWLYCCG-NH₂ | 12/2 | C. elegans IPGM 5.05, B. malayi IPGM 5.56, O. volvulus IPGM 5.62, D. immitis IPGM 5.73, E. coli IPGM 5.03, H. sapiens dPGM NA, E. coli dPGM NA |
| Bm-4 | Ac-DYLEWPNCNTCG-NH₂ | 7/4 | C. elegans IPGM 5.61, B. malayi IPGM 6.24, O. volvulus IPGM 6.10, D. immitis IPGM 6.21, E. coli IPGM 5.56, H. sapiens dPGM NA, E. coli dPGM NA |
| Bm-5 | Ac-DYLDPNCTSCG-NH₂ | 7/4 | C. elegans IPGM 5.60, B. malayi IPGM 6.16, O. volvulus IPGM 5.99, D. immitis IPGM 6.10, E. coli IPGM 5.54, H. sapiens dPGM NA, E. coli dPGM NA |
| Bm-6 | Ac-DYEWPNCNSTCG-NH₂ | 7/4 | C. elegans IPGM 5.58, B. malayi IPGM 6.16, O. volvulus IPGM 6.08, D. immitis IPGM 6.26, E. coli IPGM 5.55, H. sapiens dPGM NA, E. coli dPGM NA |
| Bm-7 | Ac-DYAVVPNCRCTG-NH₂ | 7/4 | C. elegans IPGM 5.13, B. malayi IPGM 5.28, O. volvulus IPGM 5.37, D. immitis IPGM 5.42, E. coli IPGM 5.12, H. sapiens dPGM NA, E. coli dPGM NA |

Peptide sequences Bm-1–Bm-7 identified from theDY library after 5 rounds of selection with B. malayi iPGM, Ce-1 and -2 identified from theDY library after 6 rounds of selection with C. elegans iPGM, and Ce-3 and -4 identified from theDY library after 6 rounds of selection with C. elegans iPGM. pIC50 = \(-\log IC_{50}\). All resynthesized peptides are made as C-terminal carboxamides. For DNA sequences of the RaPID selections see Supplementary Fig. 1. Enzyme concentrations used in assays were IPGM: C and B, 5 nM; Ov, 20 nM; Di and Ec 10 nM. (dPGM: Hs 5 nM and Ec 4 nM). Additional data including error for pIC50 (the negative logarithm of the IC50) values can be found in Supplementary Table 3. Bold residues demarcate thioether linkage.

truncation analogues (Bm-4a – Bm-4d, Supplemenary Table 3). In addition, to probe the importance of the free sulphydryl group of the tail cysteine, Cys10 was replaced with serine. Interestingly, the Cys10Ser replacement and elimination of all but the terminal Gly11 resulted in inactive macrocycles (Supplementary Table 3).

Our initial RaPID experiments resulted in the Bm-series cyclic peptides, which have relatively modest potencies. In addition B. malayi iPGM had proven refractory to crystallographic structure determination (vide infra). However, successful crystals were obtained for C. elegans iPGM and motivate RaPID targeting of this orthologue towards the deduction of design rules guiding macrocycle-iPGM interactions. A second set of affinity selection experiments used the C. elegans model organism iPGM for panning with either an N-chloroacetyl D- or L-tyrosine-initiated macrocyclic peptide library. Around 700000 2D- and 15-mer peptides, ipglycermides A (Ce-1) and B (Ce-2, Fig. 2a) or 14-mer peptides (Ce-3, Ce-4), respectively. Analagous medium and large ring systems to those obtained from the B. malayi iPGM panning (Table 1) were isolated, but displaying some level of orthologue selectivity (for example, compare Ce-1 with Ce-3 in Table 1).

Ipglycermide core displays iPGM orthologue selectivity. To dissect the contribution of the cyclic from the linear sequence of the Ce-2 lariat structure, and define the minimal sequence needed for activity we conducted a structure activity relationship study involving a C-terminal truncation/substitution series (Supplementary Fig. 6). Further, linear analogues were prepared to determine the effect on potency from constraining conformational states of the peptide. While removing the majority of the linear sequence of Ce-2 resulted in loss or greatly diminished iPGM inhibitory activity (Ce-2e–Ce-2g), truncated analogues Ce-2b–Ce-2d displayed a broadened range of activity among the iPGM orthologues (Table 2). Of particular note is the shift from subnanomolar to nanomolar activity of analogues Ce-2d for C. elegans and E. coli iPGM, while the activity for B. malayi, O. volvulus and D. immitis approach IC50 values in the micromolar range (Table 2c and Supplementary Fig. 7).

Ipglycermide subnanomolar affinity dependent on Cys14 thiol. Other than Ce-2 only Ce-2a, resulting from Gly14 truncation, retains subnanomolar potency against the iPGM orthologs (Table 2) pointing to Cys14 as a key determinant in high-affinity binding. Isosteric replacement of Cys14 with Ser in Ce-2 to generate Ce-2S caused an approximately 100-fold decrease in inhibitory activity (IC50 ~ 10 nM) for the C. elegans and E. coli iPGMs, and comparable to Ce-2d a separation in potency of 100-fold between the iPGMs of C. elegans and E. coli versus B. malayi, O. volvulus and D. immitis (Fig. 2c,e). These results indicate that the high-affinity binding of Ce-2 is dependent on its Cys14 thiol, possibly involving a sulfur-transition metal ion interaction at the catalytic centre of iPGM. Finally, while Cys14 contributes to important binding interactions, a peptide devoid of the macrocyclic core comprised solely of Ce-2 residues 9–14 alone (Ce-2tail) is inactive on all PGMs (Table 2; Supplementary Fig. 7b).
The large entropic contribution of macrocyclization of the peptide was demonstrated by comparing the IC_{50}s between Ce-2 and a linear form of Ce-2, Ce-L2, made by a Cys8Ser substitution, for iPGMs on which there was measurable activity of Ce-L2. The most reliable data from C. elegans, O. volvulus, and E. coli iPGM (Supplementary Fig. 7b; Table 2) allowed a calculation of between a 2,000- and 10,000-fold enhancement of activity attributable to reduction of random states by a calculation of between a 2,000- and 10,000-fold enhancement of activity attributable to reduction of random states by

To develop a structural model that delineates the molecular interactions mediating the pharmacologic–phylogenetic relationship between the macrocycles and iPGM orthologues, we attempted to co-crystallize Ce-2 with B. malayi and C. elegans iPGM, but failed to obtain crystals. Although soaking of pre-formed iPGM crystals with Ce-2 caused the crystals to shatter, we succeeded in obtaining two apo crystal forms, monoclinic P (iPGM-m) and orthorhombic P (iPGM-o) lattices (Supplementary Fig. 8a–c) of native C. elegans iPGM. These crystals provided the structure of a nematode iPGM (Supplementary Table 5; Supplementary Fig. 9). As anticipated from primary amino-acid sequence homology among iPGM orthologues, C. elegans iPGM is quite similar to other iPGMs of bacterial origin (Supplementary Fig. 10).

Structure of a nematode iPGM. The iPGMs are monomeric bi-domain enzymes where a phosphatase domain, structurally related to the alkaline phosphatase family of binuclear metalloenzymes, is connected by two hinge peptides to a phosphotransferase domain. X-ray crystal structures have been obtained for the enzymes derived from two trypanosomatids and several bacterial species.8,25–27.
domain since substrate binding produces a large conformational change in the phosphatase domain (Supplementary Fig. 11). Similar to other iPGMs, Mn\(^{2+}\) occupies one of the two phosphatase domain metal ion binding sites while, as in alkaline phosphatase, a Zn\(^{2+}\) ion was found in the second binding site of \(C.\) elegans iPGM. The identity of these transition metal ions was verified from phased anomalous difference maps calculated from data collected at wavelengths of 1.0000 and 1.9016 Å. These metal ions contact the histidine and aspartate triads as shown in Supplementary Fig. 12 at coordination distances listed in Supplementary Table 6, and the catalytic Ser86 nucleophile coordinates to the Zn\(^{2+}\) ion.

Co-crystal structure elucidates inhibitory mechanism. From the apo iPGM structure we observed that the N-terminal 18 amino acids, unique to \(C.\) elegans iPGM, were disordered. In a subsequent crystallization effort, using an 18 amino-acid N-terminal truncated form of \(C.\) elegans iPGM, we prepared a pre-formed Ce-2 complex purified by sizing chromatography, as well as a mixture with Ce-2d. The latter resulted in needle-like crystals diffraction to 1.95 Å (Supplementary Fig. 8d and Supplementary Table 5). The final model of the iPGM•Ce-2 complex (PDB: 5KGN) contained two molecules in the asymmetric unit (Supplementary Fig. 13a), exhibiting an inter-domain binding mode with the macrocycle cradled in a pocket shaped from the hinge peptides and adjacent phosphatase and transferase domain surfaces (Fig. 3a). The two subunits of the asymmetric unit are nearly identical with an RMSD of 0.20 Å between Cx atoms for 520 residues aligned. Therefore all subsequent analyses were carried out using subunit A of the model. The structure of iPGM•Ce-2d was compared with the aforementioned apo iPGM-m and iPGM-o (PDB: 5KGL) structures. Superposition yielded an RMSD deviation of 1.98 Å (503 residues) and 2.05 Å (502 residues), respectively. Although the RMSD are somewhat large, the overall structures are remarkably similar (Fig. 3b) with slight displacement of secondary structure elements due to the high flexibility of iPGM. The cavity that accommodates the binding of the Ce-2d cyclic peptide is very similar amongst all of the structures with no marked conformational changes observed to accommodate binding of the peptide. Rather, it appears that the peptide adopts an optimal fit within this cavity as might be expected from the affinity selection approach used to discover the parent macrocycle. This region of iPGM forms a somewhat negative asymetrically charged pocket that accommodates the polar residues of Ce-2d as shown in Fig. 3c. The total surface area of the Ce-2d peptide is 432.1 Å\(^2\) as calculated using Areaimol\(^{28}\), which provides information regarding total area, contact area and the solvent exposed area of a surface. A relatively small region of the total Ce-2d peptide surface makes direct contact to iPGM (127.1 Å\(^2\)) with the remaining area (305 Å\(^2\)) exposed to solvent as it is positioned within an open pocket between the phosphatase and transferase domains. On the basis of molecular weight (1501.6) and its 27 ring atoms Ce-2d can be categorized as a large macrocycle\(^{29}\). With 29% of its surface buried, Ce-2d solvent exposes slightly more surface area than comparably sized macrocycles. The electron density for the cyclic peptide was prominent for all of the residues except for the terminal tyrosine side chain which was somewhat disordered, while the C-terminal 4 residues form a short \(\alpha\)-helix (Supplementary Fig. 13b,c). From the CPK representation of Ce-2d (Fig. 3d,e) the orientation of the tyrosine side chains Tyr3, Tyr7 and Tyr11 can be seen enfolding the cyclic peptide (Fig. 3d) while tyrosine side chains Tyr1 and Tyr9 engage in an edge-to-face interaction (Fig. 3e). In Ce-1, His7 replaces Tyr7 maintaining similar activity (Table 1). The three extra-cyclic C-terminal residues, Tyr9, Leu10 and Tyr11 are wrapped close to the core macrocycle, with the carboxamide of Tyr11 visible on the exterior surface of this compact structure and pointing toward the metal ion active site (Figs 3e and 4a). Notably, the Ce-2d C-terminal amide of the Tyr11 residue is 6.5 and 8.4 Å from the Mn\(^{2+}\) and Zn\(^{2+}\) ions, respectively (Fig. 4c). Thus, it is feasible that the longer C terminus of Ce-2 (and Ce-1) could potentially extend from this cavity, positioning Cys14 within coordination distance to either metal ion.

The Ce-2d macrocycle forms direct hydrogen bonds with \(C.\) elegans iPGM as well as water-mediated contacts as depicted in Fig. 4a,b and detailed in Supplementary Table 6. Two of these key H-bonds are made with C-terminal tail residues, Tyr9 and the carboxamide of Tyr11. Others include, iPGM Arg289 which forms two H-bonds with Ce-2d ring system residues, one directly with Asp2 and one water-mediated through the Tyr3 hydroxyl (Fig. 4b), while a bifurcated Gly5 carbonyl H-bond occurs with iPGM via Gln101 and Asp102 (Fig. 4a). Hydrophobic

---

**Table 2 | Inhibitory properties of ipglyceride B (Ce-2) analogues.**

| Peptide ID | Sequence                  | pIC\(_{50}\) |
|------------|---------------------------|-------------|
|            |                           | C. elegans iPGM | B. malayi iPGM | O. volvulus iPGM | D. immitis iPGM | E. coli iPGM | H. sapiens dPGM | E. coli dPGM |
| Ce-2\(^{\ast}\) | Ac-DYYPDYCYLYGTG-\(\mathrm{NH}_2\) | 9.60         | 9.02          | 9.18          | 9.36          | 9.26          | NA            | NA            |
| Ce-2S      | Ac-DYDPDYCYLYGTS-\(\mathrm{NH}_2\) | 8.08         | 6.28          | 5.99          | 6.01          | 8.00          | NA            | NA            |
| Ce-2a\(^{\ast}\) | Ac-DYDPDYCYLYGTC-\(\mathrm{NH}_2\) | 9.76         | 9.49          | 9.68          | 9.91          | 9.58          | NA            | NA            |
| Ce-2b      | Ac-DYDPDYCYLYGTG-\(\mathrm{NH}_2\) | 7.80         | 6.16          | 5.61          | 5.96          | 7.98          | NA            | NA            |
| Ce-2c      | Ac-DYDPDYCYLYG-\(\mathrm{NH}_2\) | 7.64         | 5.93          | 5.41          | 5.70          | 7.98          | NA            | NA            |
| Ce-2d\(^{\ast}\) | Ac-DYDPDYCYLY-\(\mathrm{NH}_2\) | 9.03         | 7.29          | 7.13          | 6.20          | 7.98          | NA            | NA            |
| Ce-2e      | Ac-DYDPDYCYN-\(\mathrm{NH}_2\) | 7.16         | 5.53          | 4.72          | 5.32          | 7.27          | NA            | NA            |
| Ce-2f      | Ac-DYDPDYCY-NH\(_2\) | 4.29         | NA            | NA            | NA            | 4.59          | NA            | NA            |
| Ce-2g      | Ac-DYDPDYCY-\(\mathrm{NH}_2\) | NA           | NA            | NA            | NA            | NA            | NA            | NA            |
| Ce-L2      | Ac-DYDPDYSYLYGTG-\(\mathrm{NH}_2\) | 5.94         | 4.89\(^{\dagger}\) | 5.18\(^{\dagger}\) | 5.18\(^{\dagger}\) | 5.88          | NA            | NA            |
| Ce-L2d     | Ac-DYDPDYSYLY-\(\mathrm{NH}_2\) | NA           | NA            | NA            | NA            | NA            | NA            | NA            |
| Ce-2tail   | Ac-\(\gamma\)LYGTG-\(\mathrm{NH}_2\) | NA           | NA            | NA            | NA            | NA            | NA            | NA            |

\(^{\ast}\)All values determined using end point assay and PGM concentrations listed in Table 1 unless indicated as above, in which case PGM concentration is 0.5 nM.

\(^{\dagger}\)Estimated from incomplete concentration response curves, Supplementary Fig. 7b. Bold residues demarcate thioether linkage.
interactions are also observed. For example, Leu10 of the macrocycle sits in a small pocket formed by iPGM Ile103, Leu78 and Leu82 (Supplementary Fig. 14), while Ile99 of the enzyme is within 3.5 Å or less of cyclic peptide residues Leu10 of the macrocycle (shaded orange in the protein sequence alignment shown in Fig. 5b). In the cavity defined by these points in both structures (Fig. 4d,e). The model places the macrocycle at a site non-overlapping with phosphoglycerate supporting an allosteric binding mode for the ipglycermides. This result is consistent with the independence of Ce-2d and Ce-2 IC50 on 3-PG substrate concentration (Supplementary Fig. 15).

To gain insight into the mechanism underlying the iPGM orthologue selectivity of the Ce-2 series observed in Fig. 5a, we began by defining a binding cavity from residues within 5 Å of the Ce-2d macrocycle (shaded orange in the protein sequence alignment shown in Fig. 5b). In the cavity defined by these amino acids we projected Ce-2d as a worm representation α-chain scaled by B-factor (gold) with several side chains, Tyr11, Pro4, Tyr3, Tyr9 and the thioether linkage shown in green (Fig. 5c), from which several salient observations can be made. As previously discussed, truncation of Ce-2 beyond Cys14 results in a >10-fold potency decrease (Ce-2b, c) until Tyr11 becomes the C-terminal residue, at which point potency for C. elegans and E. coli iPGM is recovered (Ce-2d), but only marginally so for the B. malayi, O. volvulus and D. immitis orthologues. The improvement in potency is likely the result of a new H-bond made possible by the C-terminal Tyr11 amide of Ce-2d with the highly conserved Glu87 of the phosphatase domain (Fig. 5c). Subsequent removal of Tyr11 (Ce-2e) results in nearly a 100-fold potency decrease, probably a consequence of the Tyr11 amide—Glu87 H-bond forfeiture. Continued truncation leads to virtual inactivation of the macrocycle (Ce-2f,g). A possible explanation for the dramatic separation of Ce-2d inhibitory potency between C. elegans and E. coli versus B. malayi, O. volvulus and D. immitis iPGM orthologues may be in part mediated by Ala334 within hinge 2 of C. elegans and E. coli iPGM, but replaced by a glutamate in the B. malayi, O. volvulus and D. immitis iPGM orthologues. Ala334 is <2.5 Å from Ce-2d Pro4 and Tyr3, thus the larger volume occupied by a glutamic acid residue may create a steric clash only partially compensated for by the Tyr11 amide H-bond. Sequence differences outside the binding cavity between C. elegans and E. coli versus the B. malayi, O. volvulus and D. immitis iPGMs highlighted in yellow in Supplementary Fig. 10 could also contribute to orthologue selectivity.

Discussion

The glycolytic phosphoglycerate mutase, iPGM is an attractive target for the development of broad spectrum antiparasitic and antibacterial agents owing to its essential metabolic function and evolutionary divergence from the human enzyme. Infectious organisms potentially susceptible to an iPGM inhibitor are responsible for diseases ranging from African trypanosomiasis (sleeping sickness), lymphatic filariasis (elephantiasis tropica), onchocerciasis (river blindness), Staphylococcus aureus toxic shock syndrome and Bacillus anthracis intoxication. Difficulty in achieving small-molecule inhibitors of even modest potency toward iPGM may be rationalized by our current understanding of the catalytic mechanism in which a fully formed, solvent-inaccessible active site assembles dynamically upon phosphoglycerate binding. Identification of only low-potency metal ion chelators from small-molecule HTS is not unexpected given the absence of a hydrophobic druggable pocket and dependence of enzyme activity on transition metal ions. Resembling a protein–protein interaction, the charged,
dynamic domain motions observed in iPGM catalysis necessitate the exploration of new chemical space that can present protein surface complementarity to achieve an inhibition mechanism13.

In an effort to elucidate additional ligand interfaces of iPGMs for the study of phosphoglycerate mutase enzymology and inhibition thereof, we applied an affinity selection approach to address the limited protein surface topology of small-molecule chemical libraries. As a complementary methodology to HTS-based ligand discovery31, in vitro affinity display technologies generating peptide32,33, DNA or RNA34 ligands can expose currently 'undruggable' target space by defining new molecular topologies to inform the design of complementary ligand scaffolds in synthetic small molecules. To this end, we pursued macrocyclic polypeptide templates incorporating rigid stereochemical complexity extensively used by nature to interact across extended protein binding surfaces29,35. Macromolecular diversity of the cyclic peptide library used in this study exceeds by several million-fold that achieved thus far from small-molecule HTS. The self-encoded nature of this cyclic peptide repertoire enables the amplification of low-abundance, high-affinity ligand subpopulations selected using an immobilized iPGM subtype. Candidate ligand sequences obtained from iterative rounds of enrichment were resynthesized by solid-phase peptide synthesis (SPPS) yielding potent and selective iPGM inhibitors (Fig. 1). Our co-crystal structure reveals that Ce-2 stabilizes a locked-open structure, which precludes the dynamic rearrangement of domains required for catalysis. Such a dynamic-constraint inhibition mechanism is an important benefit of cyclic peptide libraries owing to their larger size, which can span greater distances between flexible protein domains.

SPPS of active cyclic peptide compounds facilitates rapid examination of the molecular details of the inhibitory mechanism and immediately presents a synthetic path towards generating an improved second-generation inhibitor. Progressive C-terminal truncation of ipglycermide B (Ce-2), or introduction of a Cys14Ser point mutation, revealed that the Cys14 sulfhydryl engendered the pan-orthologue subnanomolar potency to the macrocycle. This observation is consistent with a cysteiny1 thiolate functioning as a potential catalytic-site metal ion ligand, consistent with a borderline soft Lewis acid Zn2+ at the iPGM active site suggested by our crystallographic findings. Loss of the metal ion-anchoring sulfhydryl side chain as a consequence of truncation or Cys14Ser substitution results in potency discrimination among the iPGM orthologues (Supplementary Fig. 7a) corresponding to their phylogenetic relationship. Taken together these results suggest a two-site allosteric binding mechanism whereby the cyclic sequence and pendant C-terminal cysteine bind at distinct iPGM regions. The binding orientation of the macrocycle positions the Cys14 sulfhydryl within coordinating distance to the metal ion site (Fig. 6). An interaction between Cys14 and the Zn2+ ion would likely require a conformational adjustment in the enzyme and possibly explain the mechanism by which Ce-2 stabilizes a locked-open structure.
Interestingly, the naturally occurring mono- and bicyclic depsipeptide histone deacetylase inhibitors harbour latent thiolates either as a thioester (largazole), or as a disulfide forming a second ring system (romidepsin, thailandepsin and spiruchosatin) which, upon hydrolysis or reduction, respectively, liberates an active site $\text{Zn}^{2+}$-coordinating thiolate in mammalian cells\(^6\). In striking parallel to our present observation of Cys14 dependence on potency and Hill slope gradient, Wang, et al.\(^{3,7}\) demonstrate a similar correlation between the potency and Hill slopes of the reduced and oxidized forms of the bicyclic depsipeptide thailandepsin. The macrocyclic library therefore appears to encode sufficient diversity to enable a synthetic selection capable of capturing similar mechanistic and pharmacologic behaviours as the natural products that have evolved to potently and selectively target zinc-dependent histone deacetylases.

The discovery of ipglycermides represents potent isozyme-selective iPGM inhibitors to enable mechanistic and structural studies of glycolytic mutases from microorganisms. With antibody-like affinity and selectivity, though lacking in vivo efficacy (Supplementary Fig. 16), the ipglycermides

---

**Figure 5** | Structural basis underlying the pharmacologic-phylogenetic Ce-2 macrocycle series-iPGM orthologue relationship. (a) Relationship between Ce-2 macrocycle truncation series IC\(_{50}\) and iPGM orthologues. Analogues with no detectable inhibitory activity are indicated as inactive. Data represent mean ± s.d. values of the log normal distributed IC\(_{50}\) determined for the given peptide for experiments with $\geq \text{4}$ replicates; otherwise error bar is determined from the nonlinear fit of the standard Hill equation to the aggregated data from the replicates. Values are from Supplementary Table 3 converted from pIC\(_{50}\) where IC\(_{50}\) = 10$^\text{pIC}_{50}$. (b) Select amino-acid sequence alignment of iPGM orthologues in this study (see Supplementary Fig. 10 for full alignment). iPGM residues within 5 Å of the Ce-2d macrocycle are coloured orange. Residues identical between C. elegans and E. coli iPGM are coloured yellow; grey indicated hinge regions; green and blue are amino acids that ligand metal ions. (c) Cavity formed from C. elegans iPGM residues (light blue chain under C. elegans and transparent spheres) formed from residues within 5 Å of the Ce-2d macrocycle shown as a worm $\alpha$-chain (gold) representation scaled by B-factor with select side chains (Tyr3, Pro4, thioether linkage, and C-terminal Tyr11 amide) shown. The iPGM Ala334 residue is shown as a CPK space fill. Electrostatic surface of the binding cavity is also shown. (Tyr3, Pro4, thioether linkage, and C-terminal Tyr11 amide) shown. The iPGM Ala334 residue is shown as a CPK space fill. Electrostatic surface of the binding cavity is also shown.

---

**Figure 6** | Modelling of C-terminal residues of Ce-2 onto Ce-2d. (a) The Ce-2d macrocycle is shown as worm $\alpha$-chain (gold) representation scaled by B-factor within a cavity of C. elegans iPGM residues (transparent spheres) formed from residues within 5Å of cyclic peptide. The C-terminal residues, -Gly12-Thr13-Cys14-Gly15 of Ce-2 were modelled onto the Ce-2d macrocycle and are shown as cyan sticks extending from Ce-2d. Electrostatic surface of the binding cavity is also shown. (b) Ce-2 van der Waals radii shown using a CPK model. The Cys14 sulphhydryl is shown in yellow.

---

\(\text{IC}_{50}\) (nM)

| IC\(_{50}\) (nM) |
|----------------|
| 10,000         |
| 1,000          |
| 100            |
| 0.1            |
| 0.01           |

---

\(\text{HVTFFFN GGVEKQFANEERC ADHGNAEKMMAP - DGSKHTAHTCNLVPFTCS}\)
exemplify the fertile yet uncultivated landscape between small molecules and protein biologics. The macrocyclic peptides and accompanying crystallographic information presented here for an enzyme previously considered ‘undruggable’ reveal an important binding mode and inhibition mechanism that may be applicable to other difficult targets and extendable to modulation of protein–protein interactions. The interfacial contacts between iglucydeme and iPGM may inspire peptide mimetic analogues, particularly as significant regions of the macrocyle act as a scaffold to position a subset of key residues (Fig. 4). We anticipate the molecular tools and atomic-level structural guidance provided should stimulate further progress in this area.

Affinity selection and enrichment. Affinity selections were independently performed with V10 library against B. malayi iPGM (His6-tagged) and E. coli phosphoglycerate mutase (PGM) involved growth of cultures at 37 °C, 200 mM Tris-HCl, 200 mM NaCl and 50% glycerol, pH 7.5) and stored at −80 °C. Each PGM enzyme was cloned into pET21a(+). To have an AUG (ClAcL/DY) initiator codon followed by 4–12 NNK random acid residues, followed by a fixed UGC codon that assigns Cys. The theoretical diversity of the macrocycles based on the quantitation of the individual transformation steps (see below) is at least 1017. In vitro translation, a thioether bond formed spontaneously between the N-terminal Cys and the sulfhydryl group of the initiator L-Tyr residue and the C-terminal Cys residue.

Methods
Preparation of PGM enzymes. All PGM enzymes were cloned into pET21a(+) and the inserts fully sequenced to validate authenticity. Proteins were expressed in the E. coli strain C5866/T7 Express (Promega; Milan, Italy) in a pAK1 T7 polymerase (Novagen; Madison WI) and expressed and purified as previously described. Briefly, optimum conditions for production of soluble recombinant PGM involved growth of cultures inoculated with 0.1 mM IPTG overnight at 16 °C. The His-tagged proteins were purified on a 5 mL HiTrap chelating HP column (GE Healthcare; Pittsburg, PA) using an AKTA FPLC. After application of the sample, the column was washed with five volume column volumes of buffer A (20 mM NaPO4, 300 mM NaCl, 10 mM imidazole, pH 7.4) followed by 10 column volumes of buffer B (20 mM NaPO4, 300 mM NaCl, 400 mM imidazole, pH 7.4). Protein was then eluted using a linear gradient (8–100%) of buffer B equivalent to 40–400 mM imidazole. Fractions containing iPGM-His6x were pooled, dialysed against dialysis buffer (40 mM Tris-HCl, 200 mM NaCl and 50% glycerol, pH 7.5) and stored at −20 °C. Before use samples were prepared for analysis by size exclusion chromatography to separate active enzyme from aggregates formed during storage. Proteins were analysed by SDS-PAGE to confirm the predicted size and purity (Supplementary Fig. 3), and concentrations determined using the BCA protein assay. The sequences were expressed and purified as previously described (see Supplementary Information for additional details). Briefly, the chloroacetyl-phosphate-tRNAfMet (ClAc-puromycin, 250 µM) mRNA-puromycin conjugate and 30 µl (from round 2, 0.7 µl of 250 µM ClAc-L-Tyr-tRNAfMetCAU or ClAc-D-Tyr-tRNAfMetCAU were incubated at 37 °C for 30 min with an extra incubation at 25 °C. After an addition of 15 µl of 200 mM EDTA solution, the reaction solution was incubated at 37 °C for 2 h to facilitate macrocyclization and subject to incubation with 25-µl G-25 columns to remove salts. The desalted solution of peptide–mRNA (peptide–mRNA/cDNA from round 2, vide infra) was applied to Dynabeads His-tag Isolation & Pulldown magnetic beads (Invitrogen) to remove undesired sequences; this process was pre-clearance or negative selection and was repeated twice (six times from round 1). After pre-cleaning, the peptide–mRNA (peptide–mRNA/cDNA from round 2) solution was incubated with B. malayi iPGM or C. elegans iPGM-immobilized Dynabeads for 30 min at 4 °C to obtain iPGM-binders. This process is referred to as positive selection. The selected fused peptide–mRNA on the beads was reverse transcribed by M-MLV reverse transcriptase (Promega) and 30 µl of cDNA (from round 2, 0.7 µl of 250 µM ClAc-L-Tyr-tRNAfMetCAU or ClAc-D-Tyr-tRNAfMetCAU) was PCR amplified using primers (vide infra) and 3 µl of 200 mM EDTA solution. The PCR products were purified using a QIAquick PCR Purification Kit (Qiagen) and then verified by MALDI-TOF mass spectrometry, using a microflex or Autoflex instrument (Bruker Daltonics; Supplementary Fig. 17; Supplementary Table 4).

Gradient elution moving boundary electrophoresis (GEMBE). GEMBE was used for the direct monitoring of enzyme activity via label-free measurement of the substrate and product, 2-PG and 3-PG. Enzyme reaction mixtures were prepared in the GEMBE apparatus 48 with cyclic peptide concentrations between 195 pM and 2.5 µM for compounds Ce-2 and Ce-2d (plus no inhibitor controls). Separations were performed in 5 cm long, 15 µm inner diameter capillaries with 400 V cm−1 electric field strength and 12.5 Pa−1 pressure ramp rate. For each sample, the GEMBE separation was repeated multiple times to monitor the conversion of substrate to product. Electropherogram data (Supplementary Fig. 4) was fit to a sum of complementary error functions for quantitation. The reaction rate was determined by a linear fit to the per cent conversion versus reaction time data. Full details are available in Supplementary Information.

Macrocyclic peptide characterization across PGM orthologues. Macrocyclic peptide solutions were prepared in DMSO at a concentration beginning at 1 or 5 mM and titrated as an 11-point 1.3 or 1.6 point 1.2 series. For the 11-point titration series, compound dispense plates were prepared by using the NCATS compound management in 1356-well polypropylene deep well, v-bottom plates (Greiner Bio-One, #782270) in a single interleaved row-wise pattern per macrocyclic peptide resulting in a concentration range of 5 mM to 84.7 mM. For the 16-point titration series, compound dispense plates were prepared by hand down a single column per peptide of 254-well polypropylene deep well, v-bottom plates (Greiner Bio-One, #782270) in a single interleaved row-wise pattern per macrocyclic peptide resulting in a concentration range of 5 mM to 152.6 mM or 1 mM to 30.5 mM. Each macrocyclic peptide was characterized across five iPGM orthologues, two iPGM isozymes and the PK-FLuc
control in the Kinase-Glo Plus coupled enzyme assay described above. For the assay, 23 μL of the peptide titration series from either the 11-point or 16-point compound dispense plate were simultaneously transferred to 1536-well assay plates (Cat# 789802-F, Greiner Bio-One North America) using a 1536-pin tool (Wako) for a final concentration range of 19.2 μM-0.33 mM, 19.2 μM-0.58 mM, or 383 μM-1.17 mM, respectively.

Fitting Concentration response curves reported were generated using GraphPad Prism 5 employing the sigmoidal dose-response (variable slope) curve fitting function (equation (1)):

$$Y = Bottom + \frac{\text{Top} - Bottom}{1 + \left(\frac{IC_{50}}{IC_{50 \text{top}} - IC_{50 \text{bottom}}}\right)^n}$$

Where $IC_{50}$ is the apparent inhibition of the system with maximal response magnitude S, and $S_0$ is the inhibitor and enzyme concentrations, respectively, and $n$ is an offset term to accommodate variations in the normalized signal of the assay in the absence of inhibitor. The fit of the series of titrations allows $IC_{50}$ to be a shared variable for all titrations within the series. Ranges for $S$ and $E$ are described in Supplementary Information.

Error analysis. The concentration of cyclic peptide or cyclic peptide analogues resulting in an inhibition of 50% of the indicated PG activity test are reported as $IC_{50}$ values (Table 1 and Supplementary Table 3). The number of independent experiments is indicated in Supplementary Table 3. Inactive peptides were tested once. All experiments with reported s.d. for error bars in Fig. 2 were conducted with two technical replicates and are representative plots from at least 3 independent experiments. The data used to construct Fig. 5a was from Supplementary Table 3 converted from $IC_{50}$ to $IC_{20}$. Error bars represent the s.d. values of the log normal distributed $IC_{50}$ determined for the given peptide, such that $IC_{50 \text{win}} = 10^{-\left(pIC_{50} + 4.5\right)}$ and $IC_{50 \text{top}} = 10^{-\left(pIC_{50} - 4\right)}$.

Exclusion criteria. A data point would be eliminated from the curve fit if the value was determined to be an outlier based on the criteria described in Southall et al. Potential reasons for a data point to be eliminated from a curve fit would include, for example, known failure of compound transfer or under dispensing of assay reagent to the test well of the 1536-well assay plate. No data points needed to be excluded in the concentration response curves presented in this study.

Phylogenetic tree construction. The protein sequences for the seven C. elegans orthologues (see above) were aligned using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment analysis. The Pearson/Fasta alignment was uploaded to RAXML BlackBox (www.genome.jp/tools/raxml) for tree construction. Fifty gamma model of rate heterogeneity and the BLOSUM62 protein substitution matrix with a maximum likelihood search were applied for tree building. No outgroup was selected for tree rooting. A rapid algorithm bootstrapping analysis was performed with 1000 replicates.

Crystallization and data collection. Purified full length apo iPGM from C. elegans was subjected to refinement in the presence of inhibitor. The fit of the series of titrations allows $IC_{50}$ to be a shared variable for all titrations within the series. Ranges for $S$ and $E$ are described in Supplementary Information.

Structure solution and refinement. X-ray diffraction data were collected at a resolution of 2.00 Å and a wavelength of 1.50 Å at the Advanced Photon Source beamline 17-ID using a Dectris Pilatus 6M pixel array detector. Intensities were integrated using XDS48 via AutoPROC50 and the Laue class analysis and data scaling were performed with Aimless51 which suggested that there were two molecules in the asymmetric unit ($V = 2.7 \AA^3/\text{Da}$, % solvent = 54%) and ($V = 2.5 \AA^3/\text{Da}$, % solvent = 50%) for C. elegans iPGM-m and mmpm for iPGM-o. The Matthews coefficient $\text{V}_\text{cof}$ indicated that there were two molecules in the asymmetric unit ($V = 2.7 \AA^3/\text{Da}$, % solvent = 54%) and ($V = 2.5 \AA^3/\text{Da}$, % solvent = 50%) for C. elegans iPGM-m and mmpm for iPGM-o respectively. Structure solution for iPGM-m was conducted by molecular replacement with Balbex53 which generated a search model using a previously determined iPGM structure (PDB: 1098) (ref. 54). Searches were conducted in space groups P21 and P21 and the top solution was obtained in the latter space group which was used from this point forward. Initial refinement of the model with Refmac55 converged at $R_{\text{free}}$ of 34%/37%. For iPGM-o, molecular replacement was conducted using Phaser56 in all possible space groups with 222 point symmetry using PDB 2IEY57 as the search model. The top solution was obtained in the space group P212121. The models were improved by automated model building with Phenix58.

Structure solution for iPGM-o Ce-2d was conducted by molecular replacement using a single subunit of iPGM-o as the search model. Searches were conducted in space groups P21 and P21 (V = 2.3 Å3/Da, % solvent = 47%) for two molecules in the asymmetric unit and the top solution was obtained in P21, which was used from this point forward. Initial refinement of the model was carried out with Refmac55 and was improved by automated model building with Apr/warp58. Subsequent refinement and manual model building were carried out with Phenix and Coot59, respectively. Disordered side chains were truncated to the point for which electron density could be observed. Structure validation was conducted with Molprobity60 and figures were prepared using CCP4 prog package61. Superposition of iPGM structures was conducted using GESAMT62 via the CCP4 (ref. 63) interface. The following Ramachandran plot statistics were obtained for the final models. C. elegans iPGM-m (favoured: 97.6%, allowed: 1.8%, outliers: 0.6%), C. elegans iPGM-o (favoured: 96.6%, allowed: 3.1%, outliers: 0.3%) and C. elegans iPGM-Ce-2d (favoured: 98.0%, allowed: 1.7%, outliers: 0.3%). Relevant crystalographic data are provided in Supplementary Table 5.

Crystallographic analysis. The final model of iPGM-m consisted of two subunits with two Mn$^{2+}$ and Zn$^{2+}$ ions modelled within domain A of each subunit (Supplementary Fig. 9a) and the first 20 residues of the N terminus and last 13 residues of the C terminus were disordered and could not be modelled. The two subunits are nearly identical with an RMSD deviation of 0.58 Å between Cz atoms for 517 residues aligned using GESAMT62 (Supplementary Fig. 9b). Crystals of the orthorhombic form (C. elegans iPGM-o) were obtained after approximately 6 months and diffracted to higher resolution than iPGM-m. Similarly, the N- and C-terminal residues were disordered in the iPGM-o as well.

Metal ion sites. Large peaks of positive electron density (Fo-Fc) were observed in the metal binding sites of the phosophate domain following refinement (Supplementary Fig. 12a). This region is occupied by Asp 426, His 430 and His 485 (site1) and Asp 37, Ser 86, Asp 467 and His 468 (site2). On the basis of the coordination distances and electron density (difference and anomalous), Mn$^{2+}$ and Zn$^{2+}$ ions were assigned at these respective sites. Further details are provided in the Supplementary Information.

Data availability. Coordinates and structure factors have been deposited to the Protein Databank with the following accession codes: C. elegans apo iPGM-m (5KG), C. elegans apo iPGM-o (5KGL) and the complex C. elegans Met19 iPGM • Ce-2d (5KGK). The data that support the findings of this study are available from the corresponding author upon request.

References

1. Coghlan, A. Nematode Genome Evolution. (ed. Community, T.C.E.R.) (WormBook, 2005).
2. World Health Organization. The World Health Report: 2004: Changing Health (2009).
3. Churcher, T. S. et al. Identifying sub-optimal responses to ivermectin in the treatment of River Blindness. Proc. Natl Acad. Sci. USA 106, 16590-16595 (2009).
4. Osei-Atweneboana, M. Y. et al. Phenotypic evidence of emerging ivermectin resistance in Onchocerca volvulus. PLoS Negl. Trop. Dis. 5, e998 (2011).
5. Zeller, S. et al. Mining predicted essential genes of Brugia malayi for nematode drug targets. PLoS ONE 2, e1189 (2007).
6. Zhang, Y., Foster, J. M., Kumar, S., Fougere, M. & Carlow, C. K. Cofactor-independent phosphoglycerate mutase has an essential role in Caenorhabitis

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms14932 | ARTICLE
elegans and is conserved in parasitic nematodes. J. Biol. Chem. 279, 37185–37190 (2004).

7. Singh, P. K., Kushwaha, S., Mohd, S., Pathak, M. & Misra-Bhattacharya, S. In vitro gene silencing of independent phosphoglycerate mutase (iPGM) in the filarial parasite Brugia malayi. Infect. Dis. Povert. 2, 5 (2013).

8. Jedrzejas, M. J., Chander, M., Setlow, P. & Krishnasamy, G. Structure and mechanism of action of a novel phosphoglycerate mutase from Bacillus stearothermophilus. EMBO J. 19, 1419–1431 (2000).

9. Jedrzejas, M. J., Chander, M., Setlow, P. & Krishnasamy, G. Mechanism of catalysis of the cofactor-independent phosphoglycerate mutase from Bacillus stearothermophilus. Crystal structure of the complex with 2-phosphoglycerate. J. Biol. Chem. 275, 23146–23153 (2000).

10. Rigden, D. J., Mello, L. V., Setlow, P. & Jedrzejas, M. J. Structure and mechanism of action of a cofactor-dependent phosphoglycerate mutase homolog from Bacillus stearothermophilus with broad specificity phosphatase activity. J. Mol. Biol. 315, 1129–1143 (2002).

11. Jedrzejas, M. J. Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. Prog. Biophys. Mol. Biol. 73, 263–287 (2000).

12. Crowther, G. J. et al. Cofactor-independent phosphoglycerate mutase from nematodes has limited druggability, as revealed by two high-throughput screens. PLoS Negl. Trop. Dis. 8, e2628 (2014).

13. Lazo, J. S. & Shabow, E. R. Drug undruggable molecular cancer targets. Annu. Rev. Pharmacol. Toxicol. 56, 23–40 (2016).

14. Narisawa, S. et al. Novel inhibitors of alkaline phosphatase suppress vascular smooth muscle cell calcification. J. Bone Miner. Res. 22, 1700–1710 (2007).

15. Chakhudguin, N. K. & Suga, H. Construction and screening of vast libraries of natural-product-like macrocyclic peptides using in vitro display technologies. Curr. Opin. Chem. Biol. 24, 131–138 (2015).

16. Feraudi, M., Gartner, C., Kolb, J. & Weicker, H. Bioluminescent and fluorometric techniques for determinations of 19 metabolites of ADP/ATP-sugars using capillary electrophoresis with indirect photometric detection. Anal. Chem. 83, 6316–6322 (2011).

17. White, M. F. & Fothergill-Gilmore, L. A. Development of a mutagenesis, fluorometric techniques for determinations of 19 metabolites of ADP/ATP-sugars using capillary electrophoresis with indirect photometric detection. Analyst 136, 1310–1317 (2011).

18. Hoque, M. A. et al. Design and synthesis of mono and bicyclic tetrapeptides thioester as potent inhibitor of histone deacetylases. Amino Acids 46, 2435–2444 (2014).

19. Wang, C. et al. ThaiLandepins: bacterial products with potent histone deacetylase inhibitory activities and broad-spectrum antiproliferative activities. J. Nat. Prod. 74, 2031–2038 (2011).

20. Raverdy, S., Zhang, Y., Foster, J. & Carlow, C. K. Molecular and biochemical characterization of nematode cofactor independent phosphoglycerate mutases. Mol. Biochem. Parasitol. 156, 210–216 (2007).

21. Strycharski, E. A., Henry, A. C. & Ross, D. Expanding the capabilities of microfluidic gradient elution moving boundary electrophoresis for complex samples. Anal. Chem. 83, 6316–6322 (2011).

22. Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. Nat. Protoc. 6, 779–790 (2011).

23. Yamagishi, Y. et al. Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. Chem. Biol. 18, 1562–1570 (2011).

24. Hayashi, Y., Morimoto, J. & Suga, H. In vitro selection of anti-Akt2 thioether-macrocylic peptides leading to isofrome-selective inhibitors. ACS Chem. Biol. 7, 607–613 (2012).

25. Hipolito, C. J., Tanaka, Y., Katoh, T., Nureki, O. & Suga, H. A macrocyclic peptide that serves as a co-crystallization ligand and inhibits the function of a MATE family transporter. Molecules 18, 10514–10530 (2013).

26. Morimoto, J., Hayashi, Y. & Suga, H. Discovery of macrocyclic peptides armed with a mechanism-based warhead: isofrome-selective inhibition of human deacetylase SIRT2. Angew. Chem. Int. Ed. Engl. 51, 3423–3427 (2012).

27. Yamagata, K. et al. Structural basis for potent inhibition of SIRT2 deacetylase by a macrocyclic peptide inducing dynamic structural change. Structure 22, 345–352 (2014).

28. Southall, N. T., Jadhav, A., Huang, R., Nguyen, T. & Wang, Y. in Handbook of Drug Screening 442–446 (CRC Press, 2009).

29. Arai, T., Sakata, A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014).

30. Kabsch, W. Automatic-Indexing of Rotation Diffraction Patterns. J. Appl. Crystallogr. 21, 67–71 (1988).

31. Kabsch, W. Xds. Acta Crystallogr D 66, 125–132 (2010).

32. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D 65, 293–302 (2011).

33. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr. D 67, 282–292 (2011).

34. Matthews, B. W. Solvent content of protein crystals. J. Mol. Biol. 33, 491–497 (1968).

35. Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. BALBES: a molecular-replacement pipeline. Acta Crystallogr. D 64, 125–132 (2008).

36. Rigden, D. J., Lamani, E., Mello, L. V., Littlejohn, J. E. & Jedrzejas, M. J. Insights into the catalytic mechanism of cofactor-independent phosphoglycerate mutase from X-ray crystallography, simulated dynamics and molecular modeling. J. Mol. Biol. J. Mol. Biol. 328, 909–920 (2003).

37. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240–255 (1997).

38. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr 40, 658–674 (2007).

39. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221 (2010).

40. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. 3, 1171–1179 (2008).

41. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 66, 486–501 (2010).
60. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

61. Potterton, L. et al. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D* **60**, 2288–2294 (2004).

62. Krissinel, E. Enhanced fold recognition using efficient short fragment clustering. *J. Mol. Biochem.* **1**, 76–85 (2012).

63. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).

Acknowledgements

This research was supported in part (Z1ATR000247-01) by the Intramural Research Program of the National Center for Advancing Translational Sciences, NIH (J.I.), and in part by the AMED Basic Science and Platform Technology Program for Innovative Biological Medicine (JST-CREST Molecular Technologies) to H.S. Use of the University of Kansas Protein Structure Laboratory was supported by a grant from the National Institute of General Medical Sciences (P30GM110761) from the National Institutes of Health. Use of the IMCA-CAT beamline 17-ID at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Hauptman-Woodward Medical Research Institute. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. We thank Drs W. Jack, L. McReynolds and B. Wright for critical reading of the manuscript. This paper is dedicated to Professor S.J. Benkovic on the occasion of his 50th year of research at the Pennsylvania State University.

Author contributions

J.I, H.S. and C.K.S.C. designed the study. Z.L. and R.M. expressed and purified PGMs. H.Y. carried out RaPID and SPPS. P.D. conducted PGM assays. Z.L. performed *C. elegans* culture experiments. S.L., N.M. and K.P.B. conducted x-ray crystallography. D.J.R. and M.S.M. carried out GMEBE analysis. R.M. and N.J.B. analysed the data. J.I., H.S, S.L. and C.K.S.C wrote the manuscript with input from others.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

How to cite this article: Yu, H. et al. Macrocycle peptides delineate locked-open inhibition mechanism for microorganism phosphoglycerate mutases. *Nat. Commun.* **8**, 14932 doi: 10.1038/ncomms14932 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.