Enhancing the Efficacy of Cation-Independent Mannose 6-Phosphate Receptor Inhibitors by Intracellular Delivery

Vipul Agarwal, Priyanka Toshniwal,[†] Natalie E. Smith,[†] Nicole M. Smith, Binbin Li, Tristan D. Clemons, Lindsay T. Byrne, Foteini Kakulas, Fiona M. Wood, Mark Fear, Ben Corry* and K. Swaminathan Iyer*

Supporting Information

S1.1 Cell culture

Human primary dermal fibroblast cell cultures from normal skin were cultured in Dulbecco's Modified Eagle's Medium (DMEM/F12 - GlutaMAX; Invitrogen Gibco) supplemented with 10 % fetal bovine serum (FBS; Invitrogen Gibco) and 1% penicillin/streptomycin (Invitrogen Gibco). Analogues 1 and 2 were provided by Pharmaxis Ltd after stringent QC analysis. M6P and analogue 1 were dissolved in PBS filter **sterilised**, aliquoted and stored at -20 °C. Each aliquot was used for maximum 2 freeze thaw cycles. Analogue 2 was dissolved in DMSO and treated the similar way as other analogues. The cells were incubated at 37 °C in a humidified atmosphere of 5 % CO₂. All experiments were carried out with cells between passages 3-6.

S2.1 Cell Viability

Cell viability was determined using a LIVE/DEAD viability/cytotoxicity Kit (Invitrogen, UK) which measures the membrane integrity of cells, as per manufacturer's protocol. In brief, 20000 cells were seeded in each well in a 24 well plate and incubated with analogues at 10 μ M concentration in cell culture media (DMEM F-12 containing 10% FBS and 1% Penicillin/ Streptomycin) for 30 min (media was added in case of controls). Following which human recombinant TGF β_1 (2 ng/mL) (Cat.# 14-8348-62, eBioscience) was added to the wells (fresh media was added in controls) and plates were incubated for 24 h and 72 h in the humidified incubator at 37 °C with 5 % CO₂. At the stipulated time (24 h and 72 h), cells

were washed with PBS (3 times) and then stained with calcein (100 μ L, 1 μ M)/ ethidium bromide (100 μ L, 2 μ M) in PBS and incubated in the humidified incubator at 37 °C and 5 % CO₂ for 30 min. Images were captured using an Olympus IX71 inverted microscope with a 20 x objective with fixed exposure time. Both live and dead cells were counted using Image J software (NIH) with cell counter plug in. Data presented as mean ± standard error mean (n = 4).

S3.1 Cell Body Area

Cell size was measured using Image J software (NIH).¹ Minimum of 40 cells were randomly selected from the fluorescence images and their area was measured. Values reported as mean \pm standard error mean.

S4.1 Cell Proliferation

Cell proliferation was measured using the MTS assay (Cell Titer 96 [®]Aqueous, Promega, Madison, USA) as per the manufacturer's protocol. Briefly, 1500 cells were seeded in each well of a 96 well plate and treated with analogues at 10 μ M concentration in cell culture media (DMEM F-12 containing 10 % FBS and 1 % Penicillin/ Streptomycin) for 30 min (media was added in case of controls) prior to the addition of human recombinant TGF β_1 (2 ng/mL) to each well (fresh media was added in controls) and the plates (individual for each time point) were incubated for 72 h in the humidified incubator at 37 °C with 5 % CO₂. MTS solution (40 μ L) was added in each well the next day, and was considered as 0 h time point, and incubated for 3 h in the humidified incubator at 37 °C with 5 % CO₂. Following which 80 μ L from each well was transferred into a new 96 well plates and read under a plate reader at 490 nm excitation wavelength. Same protocol was followed at every time point for next 72 h. Data presented as mean ± standard error mean (n = 5).

S5.1 Gene Expression

Gene expression was measured using real time quantitative polymerase chain reaction. Cells (50000) were seeded in 24 well plates and incubated for 24 h in the humidified incubator at 37 °C and 5% CO₂. Next day culture media (10 % FBS) was replaced with starve media (0.1 % FBS) and incubated for further 24 h in the incubator to bring all the cells under same physiological cycle. Next day, cells were treated with required concentration of M6P analogue 30 min prior to TGF β_1 (2 ng/mL in starve media) stimulation and further incubated for 48 h in the humidified incubator at 37 °C and 5% CO₂. mRNA was extracted using RNeasy mini kit[®] according to manufacturers' protocol (Qiagen GmbH). For reverse transcription 1.5 µg of total mRNA was converted to cDNA using Superscript VILO (Cat.# 11754, Applied Biosystems) according to manufacturers' protocol. 150 ng of cDNA was analysed by ABI 7500 fast analysis real-time PCR system using TaqMan[®] master mix and collal probes (Hs01076777_m1, Life Technologies). GAPDH was used as a reference gene (Cat.# 4326317E, Life Technologies). Analysis was carried out using the instrument software. Data presented as mean \pm standard error mean (n = 3).

S6.1 Protein Expression

Protein expression was measured using western blotting. Cells (1 x 10⁵) were seeded in each well of a 6 well plate and incubated for 24 h in the humidified incubator at 37 °C and 5% CO₂. Next day culture media (10 % FBS) was replaced with starve media (0.1 % FBS) and incubated for further 24 h in the incubator to bring all the cells under same physiological cycle. Next day, cells were treated with required concentration of M6P analogue 30 min prior to TGF β_1 (2 ng/mL in starve media) stimulation and further incubated for 72 h in the humidified incubator at 37 °C and 5% CO₂. Whole cell lysates were prepared from treated cells. 35 µg of protein was denatured and subjected to SDS-PAGE, and transferred to

nitrocellulose membrane (Cat.# 10600007, Amersham, General Healthcare Lifesciences) by standard transfer. Post transfer, membrane was blocked with 5% skim milk/0.1% TBST for 30 min at room temperature, then incubated overnight with rabbit anti-human collagen I antibody (1:2000 in 5% skim milk/0.1% TBST, Cat. # NB600-408, Novus) at 4 °C. Membranes were washed with 0.1% TBST and incubated with peroxide conjugated mouse anti-rabbit (1:5000 in 5% skim milk/0.1% TBST, Cat.# NA934VS, GE Healthcare Lifesciences) for 1 h in 5% skim milk/0.1% TBST at room temperature. Immunoreactivity was detected using the chemiluminescent HRP substrate (Cat.# WBKLS0100, Millipore IMMobilon) and the signal was captured with the Chemidoc (BioRad, Model #731BR02144) and analysed using ImageJ software (NIH). To confirm equality of protein loading, all membranes were stripped and reanalysed for β -actin expression using 1° antibody (1:50000 in 5% skim milk/0.1% TBST, Cat.# NA9310V, GE Healthcare Lifesciences). Data presented as mean ± standard error mean (n = 3).

S7.1 Statistics

The results for cell viability, cell body area, cell proliferation, gene and protein expression are expressed as mean \pm standard error mean (SEM) and analysed for analysis of variance (ANOVA). Significance was evaluated using Bonferroni and Turkey's post-hoc analysis and set at 95% confidence (p < 0.05).

S8.1 Detailed synthetic pathway of the analogues have been reported previously (refer to references 2, 3 and 4).²⁻⁴ Both the analogues were synthesised and provided by

Pharmaxis Pvt Ltd after thorough characterisation. Brief description of the synthetic pathway is as follows:

Briefly, 1-methyl mannoside was reduced and reacted with benzyl chloride to protect the reactive hydroxyl groups to give 1-methyl-2,3,4,6-tetrabenzyl-mannoside which was subsequently reacted with acetic anhydride to yield 1,6-diacetyl-2,3,4-tribenzyl-mannoside. Subsequent reaction with 1-trimethylsilyl-2,4-dimethyl phenol gave 1-(2,4-dimethylphenyl)-6-acetyl-2,3,4-tribenzyl-mannoside allowing the introduction of an aryl group at the anomeric position of the molecule. Subsequent deacetylation by reaction with sodium methoxide followed by oxidation with DMSO and oxalyl chloride gave 1-(2,4-dimethylphenyl)-6-formyl-2,3,4-tribenzyl-mannoside. A diisopropyl phosphonate group was introduced at the 6-position by reaction with tetraisopropyl methylene diphosphonate and this was followed by deprotection of the benzyl ethers by palladium-catalyzed hydrogenation to give 1-(2,4-dimethylphenyl)-6-diisopropyl phosphono-mannoside. Hydrolysis of the diisopropyl phosphonate ester by reaction with chlorotrimethyl silane and sodium iodide yielded **analogue 1**. Further, bis-pivaloxymethyl (bis-POM) ester derivatisation of analogue 1 gave **analogue 2**.

S9.1 Nuclear Magnetic Resonance

Analogues 1 and 2 were analysed by ¹H-NMR, ¹³C-NMR, ³¹P-NMR and 2D NMR (¹H-¹H COSY, ¹H-¹³C HMQC, ¹H-¹³C HMBC). All NMR experiments were performed at 298 K on a Bruker Avance 600 MHz spectrometer with Topspin 2.1 software. The samples were in a 5 mm NMR tube. Analogue 1 was run in D₂O while analogue 2 was run in DMSO-d6. Internal standards used were: H_3PO_4 for ³¹P-NMR, TMS or DSS for ¹³C-NMR and ¹H-NMR.

Analogue 1

¹H NMR 600 MHz (D₂O) δ ppm: 7.09 (s, 1H), 7.07-7.06 (m, 1H), 7.05 (s, 1H), 5.50 (s, 1H), 4.20-4.19 (q, *J* = 4.80 Hz, 1H), 4.04-4.02 (dd, *J* = 12.96 Hz, 1H), 3.68-3.64 (t, *J* = 19.20 Hz, 1H), 3.61-3.58 (m, 1H), 2.28 (s, 3H), 2.22 (s, 3H), 2.01-1.65 (m, 2H), 1.59-1.13 (m, 2H). ¹³C NMR 150 MHz (D₂O) δ ppm: 133.2, 132.3, 128.4, 151.8, 115.9, 127.9, 98.5, 74.6-74.5 (d, *J* = 16.75 Hz), 71.1, 70.7, 70.6, 26.6-26.5 (d, *J* = 2.93 Hz), 25.9-24.8 (d, *J* = 130.71 Hz), 20.2, 15.8. ³¹P NMR 243 MHz (D₂O) δ ppm: 22.4 (s). HRMS (ES: m/z): Exact mass calculated for C₁₅H₂₁Na₂O₈P [M + H]⁺: 407.08, Found: 407.0848.

Analogue 2

¹**H NMR 600 MHz (DMSO-d6) \delta ppm:** 6.953-6.950 (m, 1H), 6.92 (s, 1H), 6.904-6.902 (m, 1H), 5.56-5.51 (m, 4H), 5.313-5.310 (d, *J* = 1.68 Hz, 1H), 5.035-5.028 (s, 1H), 4.98-4.97 (m, 1H), 4.77-4.76 (m, 1H), 3.84 (s, 1H), 3.66-3.64 (m, 1H), 3.35-3.31 (m, 1H), 3.28-3.24 (td, *J* = 19.62 Hz, 1H), 2.19 (s, 3H), 2.11 (s, 3H), 1.96-1.79 (m, 2H), 1.53-1.44 (m, 2H), 1.136-1.128 (s, 18H). ¹³C NMR 150 MHz (DMSO-d6) δ ppm: 176.2, 131.4, 130.3, 127.0, 126.3, 151.8, 114.2, 98.2, 81.32-81.29 (m, 4C), 72.8-72.7 (d, *J* = 18.52 Hz), 70.7, 70.3, 69.9, 38.219-38.214 (d, 2C, *J* = 0.86 Hz), 26.51-26.49 (d, 6C, *J* = 3.06 Hz), 24.12-24.10 (d, *J* = 3.9 Hz), 22.6-21.7 (d, *J* = 140 Hz), 20.2, 15.8. ³¹P NMR 243 MHz (DMSO-d6) δ ppm: 34.2 (s). HRMS (ES: m/z): Exact mass calculated for C₁₅H₂₁Na₂O₈P [M + Na]⁺: 613.25, Found: 613.2390



Figure S1: ¹H NMR spectrum of analogue **1** in D₂O



Figure S2: ¹³C NMR spectrum of analogue 1 in D₂O



Figure S3: ³¹P NMR spectrum of analogue 1 in D2O



Figure S4: 2D COSY spectrum of analogue 1 in D₂O



Figure S5: 2D HMBC NMR spectrum of analogue 1 in D₂O



Figure S6: 2D HSQC NMR spectrum of analogue 1 in D₂O



Figure S7: ¹H NMR spectrum of analogue 2 in DMSO-d6



Figure S8: ¹³C NMR spectrum of analogue 2 in DMSO-d6



Figure S9: ³¹P NMR spectrum of analogue 2 in DMSO-d6



Figure S10: 2D COSY NMR spectrum of analogue 2 in DMSO-d6



Figure S11: 2D HSQC NMR spectrum of analogue 2 in DMSO-d6

S10.1 Simulation Systems:

Experimental Methods:

6 simulation systems of the M6P/IGFII receptor were investigated; these included the domain 3 and domain 5 dimers with a ligand in each binding site (2 binding sites per dimer). These ligands were M6P, analogues **1** and **2**. The coordinates for the domain 3 dimer with M6P bound (pdb accession code 1SYO)⁵ and the domain 5 monomer (pdb accession code 2KVB)⁶ were obtained from the protein database. In order to obtain the dimer for domain 5 and position M6P in the binding pocket, the domain 5 beta sheet regions were aligned with the corresponding beta sheet regions of domain 3. Analogue **1** and **2** were positioned in each binding pocket by aligning the mannose ring and phosphate group to the M6P coordinates obtained for domain 3.⁵ Each system was then solvated in a TIP3P water box of dimensions 65 x 60 x 114 Å and ionised with 150 mM KCl. All simulations were run with periodic boundary conditions, constant temperature (310 K) maintained using Langevin dynamics and constant pressure (1 atm) maintained with a Langevin piston, and the particle mesh Ewald method was used to compute full system electrostatics.⁷ The CHARMM 36 force field was used for protein, water and M6P parameters.⁸ The ion parameters were obtained from Joung and Cheatham.⁹ Missing parameters for analogues **1** and **2** were obtained using ab-initio techniques ¹⁰ with the program Gaussian 09.¹¹ All molecular dynamics simulations were run with the program NAMD¹² using rigid bonds to hydrogen and 2 fs time steps. Molecular graphics were generated using VMD.¹³

Water and ions were energy **minimised** for 5000 steps and equilibrated for 20 ps with the protein and substrate held fixed. A harmonic restraint with a force constant of 20 kcal/mol was then applied to the backbone atoms of the protein and on each ligand and the system was **minimised** for a further 10,000 steps prior to 500 ps of equilibration. This step was repeated with gradual reductions in the force constant with values of 10, 5, 2.5 and 1 kcal/mol. Finally, to replicate the influence of the surrounding protein domains on the individual domain being simulated, a harmonic restraint with a force constant of 0.1 kcal/mol was placed on all of the protein C α atoms which were located more than 10 Å from the binding pocket in order to ensure no loss of secondary structure throughout the simulation. The system was then **minimised** for a further 10,000 steps prior to 10 ns of equilibration. Subsequently 100 ns of equilibrium simulation were obtained for each of the 6 systems.

Data Analysis:

Cluster analysis was performed on the final 100 ns of equilibrium simulation for each ligand in order to determine the most occupied binding positions. Each ligand was clustered according to the RMSD of its coordinates with a cut-off of 3 Å. Subsequently the NAMDEnergy plugin of VMD was **utilised** to determine the interaction energy of each ligand with the protein for the entire time and for the most populated clusters.



Figure S12: Most populated binding positions (clusters) of each drug in the domain 3 and domain 5 binding domains as determined from 100 ns of molecular dynamics simulation. In each case the protein is shown in grey, and the most to the least populated clusters are for ligand 1 (initially bound to dimer 1): cyan, pink, mauve, white and green and for ligand 2 (initially bound to dimer 2): dark blue, red, orange, yellow and green. a) M6P in domain 3: Ligands 1 and 2 both remain in the vicinity of the binding site determined in the crystal structure. However as there are multiple hydrogen bond acceptors and donors, both ligands sample multiple positions as the mannose ring hydroxyl groups form hydrogen bonds with various residues and these change over the 100 ns simulation. b) Analogue 1 in domain 3: Ligand 1 has only one binding position (cyan) and ligand 2 has only two clusters indicating that in both cases it binds stably to the protein and remains in the binding pocket throughout the simulation. This binding appears to be **stabilised** by the interaction between the m-xylene rings of the 2 ligands. c) Analogue 2 Domain 3: Both ligand 1 and 2 are oriented such that the benzyl groups associate in the center of the **dimerisation** domain, hence while the ligands do have more than one binding position they remain in the vicinity of the predicted binding site. d) Domain 5 M6P: In this case the ligands occupy positions which are not asymmetrical, while ligand 1 favors a horizontal orientation with one major binding position (cyan) ligand 2 favors a vertical orientation where it has more flexibility to sample new positions. e) Analogue 1 in domain 5: Ligand 1 moves away from its initial position in the binding site and occupies a position where it is in close contact with ligand 2. f) Analogue 2 in domain 5: In

this case both ligands remain in the vicinity of the original binding position. However, while ligand 1 has one major cluster implying its motion is restricted, ligand 2 has more flexibility to move occupying multiple clusters.



Figure S13: Snapshots from 100 ns molecular dynamics simulations representative of the most heavily occupied binding positions for each ligand in the domain 3 and domain 5 dimers. The two protein subunits are shown in pink and grey respectively and residues which form common hydrogen bonds are labelled. a) M6P and b) Analogue 1 in domain 3: Ligand 1 and 2 form the strongest energetic interactions with Lys350, Lys358, Ser386 and Arg391. c) Analogue 2 in domain 3: Ligand 1 and 2 form the strongest interactions with Gln348, Arg391 and Glu416. d) M6P in domain 5: Ligand 1 and 2 are oriented differently to each other in the binding pocket. M6P forms its strongest interactions with Arg687. e) Analogue 1 in domain 5: Ligand 1 and 2 form the strongest favorably with Asn680 andArg687. f) Analogue 2 in domain 5: Ligand 1 and 2 form the strongest and Tyr714.



Figure S14: Average interaction energies obtained for the most occupied positions of each ligand in domains 3 and 5. R391 in domain 3 is equivalent to R687 in domain 5. M6P (a) and analogue **1** (b) in domain 3 both show similar interactions for each ligand and the protein subunit it is most closely associated with (for example Protein 1 with Ligand 1 as compared to Ligand 2 with Protein 2). This is because the binding mode is similar for each ligand in its respective binding site (Figure S2). Similarly, in each case the ligands also interact closely with Lys350 which is actually located on the alternate protein subunit (for example Ligand 1 to Protein 2). When compared to (a) and (b) analogue **2** (c) in domain 3 shows significant decreases in the interaction energies which is reflected by its lower overall interaction energy. M6P (d) and analogue **1** (e) in domain 5 have two major interactions the first one to Arg687 and the second to Glu709. Overall the interaction energies for domain 5 are much lower than those observed for domain 3. Analogue **2** (f) in domain 5 shows a similar pattern with a further reduction in the observed interaction energies.



Figure S15: Cell Proliferation assay showing cell growth over the period of 72 h post incubation with analogues both in the presence and absence of TGF β_1 . First and second column in each condition is representing 24 h and 72 h respectively. Significant proliferation was observed in all groups despite the reduction in proliferation in TGF β_1 treated groups. Data presented as Mean ± SEM. Significance was set at * p < 0.05 using one way ANOVA with Bonferroni post-hoc analysis.



Figure S16: HDF cell morphology post calcein AM staining imaged using fluorescent microscopy. Cells were treated for 72h, stained and imaged: a) untreated (control), b) TGF β_1 (2 ng/mL) treatment, c) M6P (10 μ M) + TGF β_1 (2 ng/mL), d) Analogue **1** (10 μ M) + TGF β_1 (2 ng/mL) and e) Analogue **2** (10 μ M) + TGF β_1 (2 ng/mL). Scale bar 2 μ m.



Figure S17: Collagen I gene time response curve post $TGF\beta_1$ (2 ng/mL) stimulation. Collagen I gene expression was significantly upregulated at 24 and 48 h post $TGF\beta_1$ stimulation compared to non-treated control. The 48 h stimulation was selected for all further experiments as it yielded significantly higher response. Data presented as average \pm SEM (n = 3). Significance was set at * p < 0.05 using bonferroni post-hoc test in one way ANOVA analysis.

References:

- 1. T. J. Collins, *BioTechniques*, 2007, **43**, S25-S30.
- 2. *Australia Pat.*, WO2010125445 A1, 2013.
- 3. US Pat., US7,648,966 B2, 2010.
- H. Schilter, C. Z. Cantemir-Stone, V. Leksa, A. Ohradanova-Repic, A. D. Findlay, M. Deodhar, H. Stockinger, X. Song, M. Molloy and C. B. Marsh, *Immunol. Lett.*, 2015, 165, 90-101.
- 5. L. J. Olson, N. M. Dahms and J.-J. P. Kim, J. Biol. Chem., 2004, 279, 34000-34009.
- L. J. Olson, F. C. Peterson, A. Castonguay, R. N. Bohnsack, M. Kudo, R. R. Gotschall, W. M. Canfield, B. F. Volkman and N. M. Dahms, *Proc. Natl. Acad. Sci.*, 2010, 107, 12493-12498.
- 7. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, J. Chem. Phys., 1995, **103**, 8577-8593.
- A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J Phys Chem B*, 1998, **102**, 3586-3616.
- 9. I. S. Joung and T. E. Cheatham, 3rd, J Phys Chem B, 2008, 112, 9020-9041.
- 10. N. Foloppe and A. D. MacKerell, *J Comput Chem*, 2000, **21**, 86-104.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L.

Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. M. Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, 2009, **Revision D.01**.

- 12. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J Comput Chem*, 2005, **26**, 1781-1802.
- 13. W. Humphrey, A. Dalke and K. Schulten, *J Mol Graph Model*, 1996, 14, 33-38.