Use of the interior cavity of the P22 capsid for site-specific initiation of atom-transfer radical polymerization with high-density cargo loading

Janice Lucon1,2, Shefah Qazi1,2, Masaki Uchida1,2, Gregory J. Bedwell3, Ben LaFrance1,2, Peter E. Prevelige Jr3 and Trevor Douglas1,2*

Virus-like particles (VLPs) have emerged as important and versatile architectures for chemical manipulation in the development of functional hybrid nanostructures. Here we demonstrate a successful site-selective initiation of atom-transfer radical polymerization reactions to form an addressable polymer constrained within the interior cavity of a VLP. Potentially, this protein–polymer hybrid of P22 and cross-linked poly(2-aminoethyl methacrylate) could be useful as a new high-density delivery vehicle for the encapsulation and delivery of small-molecule cargos. In particular, the encapsulated polymer can act as a scaffold for the attachment of small functional molecules, such as fluorescein dye or the magnetic resonance imaging (MRI) contrast agent Gd-diethylenetriaminepentacetate, through reactions with its pendant primary amine groups. Using this approach, a significant increase in the labelling density of the VLP, compared to that of previous modifications of VLPs, can be achieved. These results highlight the use of multimeric protein–polymer conjugates for their potential utility in the development of VLP-based MRI contrast agents with the possibility of loading other cargos.

We reported previously the use azide–alkyne ‘click’ chemistry to construct an anchored polymer network inside a small protein-cage architecture15–17. In this stepwise synthesis approach, polymer growth was directed to the protein-cage interior, which resulted in a protein-confined hyperbranched polymer. The protein shell acted as a barrier that limited the polymer size and left only the protein exterior exposed to the bulk solution. By labelling the resulting protein–polymer construct with a Gd-based magnetic resonance imaging (MRI) contrast agent, an enhanced magnetic resonance contrast agent was obtained, which highlights the utility of using the interior space to maximize cargo loading18. Although this method is effective, the stepwise nature of the polymerization reaction makes the process onerous for larger constrained polymer synthases.

A preferable alternative route to achieve an anchored addressable polymer is to proceed via a continuous polymerization of simple monomers from an easily modified initiator. Of the several suitable continuous biomolecule-anchored polymerization methods, we chose to use atom-transfer radical polymerization (ATRP) as it is particularly suited to an improved formation of polymer inside a protein cage. This method is not only rapid, but also results in products with relatively low polydispersity in bulk solutions and is indiscriminate with respect to the range of monomers that can be used. Also, the simplicity of the ATRP initiator means that it can be attached readily to the protein cage in a site-specific manner, and thereby we can control the site of polymer initiation. Thus, by combining ATRP with a container-like protein, the formation of a polymer scaffold constrained to the interior of a VLP architecture can be afforded in a single short reaction.

Here we report the use of ATRP to make addressable polymer networks within the confines of the bacteriophage P22-based VLP (Fig. 1). The 2-aminoethyl methacrylate (AEMA) monomer was selected because the primary amine-rich polymer synthesized...
within the P22 capsid could be modified subsequently with the small molecules of interest, which resulted in very high-density loadings of the capsid. The use of the AEMA network as a scaffold was demonstrated through the attachment of either fluorescein isothiocyanate (FITC) or Gd-DTPA-NCS (DTPA = diethylenetriaminepentacetate). Using this method, we can achieve a substantial increase in the degree of labelling per VLP compared to that given in previous reports, which demonstrates the potential capacity of the capsid interior for directed cargo loading.

Results and discussion
In this work we utilized a VLP (derived from the bacteriophage P22) that consists of 420 subunits arranged on an icosahedral lattice with a resulting exterior diameter of 64 nm and an unoccupied internal cavity with a 54 nm diameter. Recombinant expression in Escherichia coli requires co-expression of the coat protein and scaffold protein for self-assembly. This VLP is capable of transformation into a series of distinct morphologies, which includes the procapsid (PC) morphology that contains the scaffold protein, an empty shell (ES) form in which the scaffold protein is removed, an expanded form (EX) and a wiffle ball (WB) structure in which all 12 pentamers are removed (Supplementary Fig. S1). The EX form most closely mimics the morphology found in the DNA-containing infectious virion and is the form used in this study.

To attain the EX form, the scaffold protein is removed from P22 using successive guanidine \(\cdot\) HCl extractions followed by heating at 65 °C, which generates the capsid in its EX morphology.

A new P22 mutant was designed for use with ATRP that contained a single-point mutation in the wild-type coat protein (P22S39C) and introduced an addressable thiol suitable for the attachment of an initiator (1). This site was selected specifically such that the introduced thiol could be exposed exclusively to the interior according to the currently available P22 structural models (Fig. 2). Although the wild-type protein contains an intrinsic cysteine (C405) it was not removed because previous studies demonstrated that this site is not addressable. This new mutant was characterized and behaves in the same manner as the wild-type P22 capsid, going through the same series of morphological transformations. To obtain the EX morphology, the scaffold protein was removed from P22S39C to generate the ES, heated to 65 °C and subsequently analysed to ensure the formation of the EX morphology (Fig. 3). The characteristic shift of the particles to
a lower electrophoretic mobility was observed on heating, consistent with expansion of the capsid. Precipitation of the protein was observed only at temperatures greater than 80°C, which indicates that the protein architecture is relatively thermostable (Supplementary Fig. S2). The size of the VLP, by dynamic light scattering (DLS), increased as expected from 60±4 nm (PC morphology) to 71±5 nm, consistent with the known range of values for the P22 particle in either the EX or WB morphologies (Fig. 3d). By transmission electron microscopy (TEM) the overall structure of the VLP was retained and, after heating to 75°C, large pores became apparent in the structure (Fig. 3c), which is a characteristic of the WB morphology.

To make the P22 macroinitiator, the P22S39C mutant was labelled with initiator 1, an amide derivative of a previously reported ATRP initiator. This cysteine-reactive ATRP initiator (Fig. 1) was selected because of its efficient labelling, satisfactory initiation and demonstrated compatibility with biomolecules. Initiator 1, unlike the ester-containing form previously reported, is expected to be less susceptible to bond cleavage24,25. It was synthesized through the modification of established protocols and selectively reacted with P22S39C to make P22S39C-int with near-quantitative single labelling of the introduced cysteine, as observed by subunit mass spectrometry (Supplementary Fig. S3).

Using the P22S39C-int macroinitiator construct, cross-linked AEMA polymer strands were synthesized inside this protein cage to make P22-xAEMA under standard ATRP biomolecule conditions using a Cu(I)/bipyridine catalyst26. To explore the range of reaction conditions available to the P22S39C system for internally directed polymerization reactions, initially a selection of AEMA:bisacrylamide monomer to subunit ratios (3,000–26,000) and temperatures (23, 40 and 60°C), as well as catalyst-loading ratios, were investigated. Samples were monitored by gel electrophoresis and DLS to determine reaction completion and suitable conditions (an example comparison is shown in Supplementary Fig. S4). The purified protein–polymer hybrid constructs were stable after polymerization and only at the highest temperature and loading conditions did the diameter of the cage increase significantly, which indicates that under most of the tested conditions the
VLP effectively constrains the polymer growth to the interior of the capsid. From these test reactions it was apparent that the monomer loading had a greater impact on the extent of polymerization than the temperature of the reaction, and that the reactions were effectively complete after less than three hours.

To verify further that our selected P22 mutant was confining the polymer, we compared the behaviour of P22K118C to a second mutant of P22 coat protein (P22S39C), described previously, and which has a reactive cysteine site predicted to be partially exposed to the exterior (Supplementary Fig. S5). Both mutants were labelled with an ATRP initiator (Supplementary Figs S6 and S7) and analysed under simple polymerization conditions. We used the AEMA without addition of bisacrylamide at a loading of 26,000 monomers/subunit of the partially exterior-exposed P22K118C to make P22K118C-AEMA, which resulted in dramatically different material properties compared to those of the P22S39C treated under the same conditions (P22S39C-AEMA) and emphasizes the importance of site specificity for polymer initiation and growth in these protein architectures. When the initiator-labelled P22K118C was treated to make P22K118C-AEMA, the construct exhibited an increase in diameter of about 20 nm, from 60 ± 4 nm, as measured by DLS, which suggests the growth of polymer on the exterior of the cage (Supplementary Fig. S8), but the P22S39C-based sample diameter remained constant before (67 ± 4 nm) and after (67 ± 4 nm) polymerization. In addition, unlike P22S39C-AEMA, the P22K118C-AEMA hybrid was not very stable. Although no precipitation was observed under the reaction conditions, mass precipitation of the P22K118C-AEMA occurred on purification. This loss could be alleviated partially by increasing the ionic strength of the buffer (>250 mM NaCl).

Taken together, these data support a model in which the initiation site determines the overall access of the AEMA polymer to the exterior environment and when the polymer growth is exposed to the exterior of the cage, the protein–polymer composite destabilizes significantly.

We directed our efforts to the interior-facing P22S39C-int macro-initiator construct at a midrange monomer loading (6,000 monomers/subunit) with bisacrylamide present and moderate temperature conditions (23 °C) for further investigation with four experimental replicates. After three hours of reaction time, the synthesis was halted by exposing each sample to air and the protein–polymer conjugate (P22S39C-xAEMA) was purified away from the remaining AEMA and bisacrylamide monomers and the copper catalyst by pelleting the protein construct using ultracentrifugation, which easily separates large macromolecular complexes from small molecular species. The resulting construct exhibited a dramatic shift in electrophoretic mobility, by native agarose gel, which indicates the range of polymer chain lengths appended to the subunits (P22S39C-xAEMA) and emphasizes the importance of site specificity for polymer initiation and growth in these protein architectures. When the initiator-labelled P22K118C was treated to make P22K118C-AEMA, the construct exhibited an increase in diameter of about 20 nm, from 60 ± 4 nm, as measured by DLS, which suggests the growth of polymer on the exterior of the cage (Supplementary Fig. S8), but the P22S39C-based sample diameter remained constant before (67 ± 4 nm) and after (67 ± 4 nm) polymerization. In addition, unlike P22S39C-AEMA, the P22K118C-AEMA hybrid was not very stable. Although no precipitation was observed under the reaction conditions, mass precipitation of the P22K118C-AEMA occurred on purification. This loss could be alleviated partially by increasing the ionic strength of the buffer (>250 mM NaCl).

Figure 4 | Size and morphological characterization of the P22S39C-xAEMA composite and P22S39C-int. a. Representative SDS–PAGE gel of P22S39C-int and P22S39C-xAEMA. The polymerized sample shows some streaking to higher $M_w$, which indicates the range of polymer chain lengths appended to the subunits as compared to the starting P22. In addition, highly cross-linked material can be observed in the well at the top of the gel. b. TEM images of P22S39C-int (top) and P22S39C-xAEMA (bottom) illustrate that the P22S39C-xAEMA retains the size and shape homogeneity of the P22 capsid after the polymerization reaction. c. DLS of P22S39C-int (top) and P22S39C-xAEMA (bottom). The modified sample is monodisperse and has the same average diameter as that of the starting P22S39C-int.
was unchanged from that before (Fig. 4b). The average particle diameter was 51 ± 3 nm in the unpolymerized P22S39C-int and remained the same (53 ± 3 nm) after the reaction. In addition, the hydrodynamic diameter, as measured by DLS, remained unchanged on polymer formation (Fig. 4c). The particle diameters were 70 ± 10 nm and 71 ± 3 nm, respectively, for P22S39C-int and P22S39C-xAEMA, which further confirms that the polymer was confined to the interior of the protein cage.

Multiangle light scattering (MALS) was used to analyse further the P22S39C-xAEMA construct for the radius and molecular weight (Fig. 5). According to this method the radius of the particles remained nearly constant, measured as 29.3 nm (P22S39C-int) and 28.2 nm (P22S39C-xAEMA), but the molecular weight increased on polymerization. Using the standard refractive index increment (dn/dc) for protein (0.185) and the published value for AEMA (0.153) (ref. 27), the MALS data from the P22S39C-int and P22S39C-xAEMA samples were fitted to obtain molecular weights. The measured molecular weight for P22S39C-int was 18.9 ± 0.2 kDa, consistent with the predicted value for the EX morphology of P22, but the P22S39C-xAEMA had a combined molecular weight of 20.2 ± 0.6 kDa, with 18.7 ± 0.5 kDa contributed from the protein and 1.5 ± 0.4 kDa contributed from the polymer. This mass increase corresponds to the addition of 12,000 ± 3,000 AEMA monomers/VLP or 28 ± 7 AEMA monomers/subunit, on average.

The internally directed P22S39C-xAEMA polymerization resulted in the introduction of a large number of addressable amines sequenced within the protein cage. To demonstrate that the introduced amines on the poly(AEMA) inside the P22S39C-xAEMA construct were addressable, FITC (2) was used as an amine-specific labelling agent. Both the P22S39C-xAEMA and P22S39C-int control were incubated with a 100-fold excess of FITC per subunit to give P22S39C-xAEMA-FITC and P22S39C-int-FITC, respectively. Excess FITC was removed by pelleting the protein twice using ultracentrifugation and resuspension before analysis. The difference in the degree of labelling between the P22S39C-int-FITC and P22S39C-xAEMA-FITC was observed by native agarose gel electrophoresis (Supplementary Fig. S10). When fluorescently labelled P22S39C-xAEMA construct was analysed using gel electrophoresis the fluorescent signal was observed to migrate with the protein. Under denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) conditions, the fluorescein migrated with the subunit and with the polymer-modified subunit, which indicates that the fluorescent dye was bound covalently to the construct and not just sorbed onto the protein–polymer composite (Supplementary Fig. S11). When analysed by native agarose gel electrophoresis, a net shift in the electrophoretic mobility of the P22S39C-xAEMA and P22S39C-xAEMA-FITC was observed, caused by the polymer and FITC altering the charge of the construct (Fig. 6a). The observed shift indicates that the polymer and dye were associated tightly with the protein cage. The P22S39C-xAEMA-FITC was labelled to the extent that the migration of a distribution of species was visible prior to staining under ambient light, but P22S39C-int-FITC was observed only weakly (Fig. 6b). In contrast, the emission signal of the P22S39C-xAEMA-FITC distribution was depressed in intensity compared to that of the much less heavily labelled P22S39C-xAEMA-FITC (Fig. 6c), which probably results from self-quenching of the fluorophore caused by the abundance of polymer-bound fluorescein in close proximity. This apparent loss of fluorescence intensity was confirmed by solution-phase analysis of the constructs, in which a greater than 95% reduction in fluorescence was observed in the P22S39C-xAEMA-FITC compared to P22S39C-int-FITC (Supplementary Fig. S12). Similar quenching was shown in systems with high fluorophore concentrations, such as micelles, where derivatives of fluorescein can lose more than 98% of their fluorescence signal because of proximity and dimerization effects28,29.

After demonstrating polymerization within the P22 capsid and the postsynthetic modification of the functional groups, we explored the P22S39C-xAEMA composite system as a potential T1 contrast agent through the attachment of Gd-DTPA-NCS (3) to the encapsulated polymer. Gd-DTPA-NCS was added, in a 100-fold excess per subunit, to the P22S39C-xAEMA or P22S39C-int and allowed to react overnight, followed by pelleting by ultracentrifugation and resuspension of the protein twice in each of four experimental replicates.

**Figure 5 | Molecular weight increase as a result of polymerization, monitored by MALS.** Using size-exclusion chromatography, both P22S39C-xAEMA and P22S39C-int exhibit the same elution time when monitored by absorbance (continuous line) at 280 nm, which indicates retention of the particle size. Fitting the MALS data across the elution profile reveals the molecular weight of the material (dashed line), which increases substantially for the P22S39C-xAEMA (top) compared to the P22S39C-int (bottom). a.u. = arbitrary units.

**Figure 6 | Polymer formation and covalent modification with FITC was verified by native agarose gel electrophoresis.** a–c. Three different views of the same native agarose gel with P22S39C-int-FITC (lane 1), P22S39C-int (lane 2), P22S39C-xAEMA (lane 3) and P22S39C-xAEMA-FITC (lane 4). a. The gel was stained with Coomassie to detect the protein component. The P22S39C-xAEMA runs slightly out of the well in the opposite direction to the other proteins because of the net positive charge on the poly(AEMA) at the running-buffer pH. Labelling this construct with the negatively charged FITC reverses the effective net charge on the construct and results in a shift in the migration direction. b. The unstained gel under ambient light. The P22S39C-int-FITC sample is faintly visible and the P22S39C-xAEMA-FITC sample is clearly visible, which signifies the relative degree of FITC labelling in each sample. c. The unstained gel illuminated with a laser at 488 nm and detected at 520 nm to highlight FITC. The P22S39C-int–FITC is considerably brighter than the P22S39C-xAEMA–FITC, which indicates the relative degree of fluorescence quenching in each sample.
The resulting material (P22\textsubscript{39C}-xAEMA-Gd or P22\textsubscript{39C}-int-Gd) was analysed for both sulfur and gadolinium content by inductively couple plasma mass spectrometry (ICP-MS) to determine both the Gd and protein concentrations. To rule out the possibility of simple electrostatic interaction between the polymer and Gd-DTPA, controls with both P22\textsubscript{39C}-xAEMA and P22\textsubscript{39C}-int were incubated and isolated, under the same conditions as above, with Magnevist (Gd-DTPA), which lacks the amine-reactive isothiocyanate. When these constructs were analysed by native agarose gel electrophoresis (Supplementary Fig. S13), a net shift in the electrophoretic mobility of the P22\textsubscript{39C}-xEAMA-Gd was observed, caused by the polymer and Gd-DTPA-NCS altering the net charge of the construct, but the P22\textsubscript{39C}-xAEMA-Magnevist sample retained the same low electrophoretic mobility as that of the P22\textsubscript{39C}-xAEMA, which provides evidence that there were no significant electrostatic interactions between P22\textsubscript{39C}-xAEMA and Gd-DTPA.

The Gd-DTPA loading per P22 was determined quantitatively for each of the constructs from the ICP-MS data. The P22\textsubscript{39C}-xAEMA-Gd contained 28 times more Gd/VLP than the P22\textsubscript{39C}-int-Gd control (320 Gd/VLP, <1 Gd/subunit), where endogenous lysines were modified, which indicated that it is largely the polymer that is being labelled rather than the protein shell. The low reactivity of the endogenous P22 lysines was reported previously, with only a few of the 20 lysines per subunit observed to be reactive.\(^{40}\) This minimal background reactivity is advantageous because it means that in the P22\textsubscript{39C}-xAEMA sample the vast majority (>95\%) of the addressable sites are located on the encapsulated polymer. If necessary, endogenous lysines could be blocked chemically prior to polymerization.\(^{41,42}\) The P22\textsubscript{39C}-xAEMA-Gd loading per cage was 9,100 ± 800 Gd/VLP (22 ± 2 Gd/subunit), which corresponds to an internal concentration of 150 mM Gd within the VLP. This is significantly more Gd, both on a per-cage and per-subunit basis, than that in previous reports using VLPs, with values that ranged from less than 1.0/subunit to 6.6/subunit (60/VLP to 650/VLP), and highlights the advantage of using the full capacity of the interior volume of VLPs.\(^{6,16,31–37}\) In addition, both the P22\textsubscript{39C}-xAEMA and P22\textsubscript{39C}-int incubated with Magnevist (Gd-DTPA) contained Gd levels below the lower limit of quantification, which demonstrates that the Gd detected is attached covalently to the protein–polymer construct rather than associated via electrostatic interactions under the labelling reaction conditions.

To verify that the high loading observed was reasonable and occurred homogeneously across the population of P22 capsids, the particles were analysed by analytical ultracentrifugation to investigate differences in sedimentation velocity (Fig. 7). The measured sedimentation value of the P22\textsubscript{39C}-xAEMA (167 S) falls within the range observed for P22\textsubscript{39C}-int (142 S) and the scaffold protein-filled PC morphology (191 S), which indicates that the polymer content does not exceed the packing observed in the naturally occurring self-assembled system. Covalent modification of polymer with Gd-DTPA-NCS to make P22\textsubscript{39C}-xAEMA-Gd resulted in a shift to a higher S value (227 S), consistent with the incorporation of additional mass via Gd-DTPA-NCS labelling of the polymer. These data indicate that there is a homogeneous shift in the entire population to higher S values with each modification.

In addition to the ionic relaxivity enhancement, each P22\textsubscript{39C}-xAEMA-Gd carries 9,100 ± 800 Gd/VLP, which leads to a per particle relaxivity of 200,000 mM\(^{-1}\) s\(^{-1}\). The particle relaxivity of P22\textsubscript{39C}-xAEMA-Gd dramatically exceeds that of the P22\textsubscript{39C}-int-Gd control (7,500 mM\(^{-1}\) s\(^{-1}\)) and those in previous reports of VLP nanoparticle-based MRI contrast agents, with values that fall in the range of 10\(^2\)–10\(^3\) mM\(^{-1}\) s\(^{-1}\) (refs 6,16,31–37). The particle relaxivity is also higher than those observed for many other macromolecular assemblies, such as micelles, liposomes and polymers, which places this construct comfortably in the upper end of observed macromolecular relaxivities for its size.\(^{38,42}\)

**Conclusions**

In summary, the application of ATRP for site-directed polymer formation inside a VLP results in an anchored network that is unparalleled for labelling purposes and utilizes the previously largely untapped interior volume of the VLP. The P22\textsubscript{39C}-based macrominitiator effectively directs polymer growth to the VLP interior, which results in a confined polymer growth because the protein shell acts as a barrier to unconstrained polymer growth, to leave only the protein shell exposed to the bulk solution. By selecting an appropriate macronitinitiator and monomer, this new multimeric protein–polymer composite acts as a scaffold for the attachment of small molecules of interest, such as the fluorophore FITC or paramagnetic MRI contrast agents (here, Gd-DTPA-NCS). The introduced polymer scaffold results in a significantly increased number of labels per cage compared to those of other VLP-based systems. The improvement in labelling is important for the delivery of contrast agent on a per-particle basis as it allows for a high concentration delivery of contrast agent or other cargo molecules of interest. This material exhibits an order-of-magnitude improvement
in relativity per particle over that of existing VLP systems. The use of this material as a targeted MRI contrast agent is particularly interesting because of the high relativity and is an application we are exploring further. As a consequence of the simplicity, modular nature and loading level of the ATRP-based approach taken to make these P22-polymer internal composites, currently this same method is being employed to make a range of novel VLP-polymer composites with biomedical and catalytic applications.

Methods

All materials were analytical grade and purchased from either Sigma-Aldrich or Fisher Scientific and used as received unless otherwise noted. Dichloromethane was distilled over calcium hydride prior to use. All water was deionized using a Nanopure water-purification system. DLS measurements were taken on a 90Plus particle size analyser (Brookhaven) and the data were processed by volume distribution.

Synthesis of 2-bromo-2-methylpropionyl bromide

2-Bromo-2-methylpropionyl bromide (200 μl, 1.6 mmol) was added dropwise. The reaction was allowed to warm to room temperature and was stirred for three hours at room temperature. The solution was subsequently distilled over calcium hydride prior to use. All water was deionized using a Nanopure water-purification system. DLS measurements were taken on a 90Plus particle size analyser (Brookhaven) and the data were processed by volume distribution.

Analysis of 2-bromo-2-methylpropionyl bromide

@ 1.5.27, 289.0188, 291.0167 (MH+ = d = 5, 5.5, 2H, NCH2) 3.72 (dd, J = 5.0, 1.5, 2H, NCH2) 6.71 (2H, CH=C=CH2) 6.97 (3H, NCH2) 3.25 (3H, CH3) 3.74 (NCH, 39.84 (NCH3), 62.51 (C), 134.59(CH2), 170.96 (CO), 173.74 (CO). Electrospray ionization mass spectrometry: m/z calculated 289.0188, 291.0167 (MH+), found m/z 289.0192, 291.0179.

P22-int macroinitiator formation conditions.


Received 12 August 2011; accepted 23 July 2012; published online 26 August 2012

References


**Acknowledgements**

This research was supported in part by grants from the National Institutes of Health (R01-EB01207), the National Science Foundation (CBET-0709358) and a National Science Foundation Graduate Research Fellowship (J.L.). P.E.P. and G.J.B. were supported by the US Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering (DE-FG02-08ER46537).

**Author contributions**

J.L. designed and carried out the experiments. S.Q. characterized the samples by NMR and analysed the relaxivity data. M.U. and B.L.F. assisted in the initial characterization of the S93C mutant. G.J.B. characterized the samples by analytical ultracentrifugation. M.U. and T.D. assisted in the experimental design. J.L. and T.D. co-wrote the manuscript. P.E.P. and T.D. coordinated the project. All authors discussed the results.

**Additional information**

Supplementary information and chemical compound information are available in the online version of the paper. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.D.

**Competing financial interests**

The authors declare no competing financial interests.