



King's Research Portal

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Weber, C. R., & Al-Badri, M. A. (2019). Super-exchange mechanism and quantum many body excitations in the archetypal hemocyanin/tyrosinase di-Cu oxo-bridge. *Nature Communication Physics*. Advance online publication.

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Superexchange mechanism and quantum many body excitations in the archetypal di-Cu oxo-bridge

Mohamed Ali al-Badri,^{1,†} Edward Linscott,² Antoine Georges,^{3,4,5,6} Daniel J. Cole,⁷ and Cédric Weber^{1,‡}

¹*King's College London, Theory and Simulation of Condensed Matter (TSCM), Strand, London WC2R 2LS, United Kingdom*

²*Theory and Simulation of Materials (THEOS), École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*

³*Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France*

⁴*Center for Computational Quantum Physics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*

⁵*CPHT, Ecole Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France*

⁶*DQMP, Université de Genève, 24 Quai Ernest Ansermet, CH-1211 Genève, Suisse*

⁷*School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom*

† mohamed.al-badri@kcl.ac.uk, ‡ cedric.weber@kcl.ac.uk

The hemocyanin protein binds and transports molecular oxygen via two copper atoms at its core. The singlet state of the Cu_2O_2 core is thought to be stabilised by a superexchange pathway, but detailed in situ computational analysis is complicated by the multi-reference character of the electronic ground state. Here, electronic correlation effects in the functional

site of hemocyanin are investigated using a novel approach, treating the localised copper 3d electrons with cluster dynamical mean field theory. This enables us to account for dynamical and multi-reference quantum mechanics, capturing valence and spin fluctuations of the 3d electrons. Our approach explains the stabilisation of the experimentally observed di-Cu singlet for the butterflyed Cu_2O_2 core, with localised charge and incoherent scattering processes across the oxo-bridge that prevent long-lived charge excitations. This suggests that the magnetic structure of hemocyanin is largely influenced by the many-body corrections.

Copper-based metalloproteins play a major role in biology as electron or dioxygen (O_2) transporters. Hemocyanin (Hc) is one of three oxygen transporting proteins found in nature, alongside the iron-based hemerythrin and hemoglobin, and is common to a number of invertebrates, such as molluscs and arthropods. Deoxy-Hc employs two half-spin copper (I) cations, each coordinated by the imidazole rings of three histidine residues, to reversibly bind O_2 . Furthermore, various type 3 copper-based systems possess catalytic properties. Hc can decompose hydrogen peroxide into water and oxygen¹ and synthetic analogues have been shown to reversibly cleave the dioxygen bond² — a mechanism that enables tyrosinase and catechol oxidases to oxidise phenols.³

An accurate understanding of the electronic structure (spin and charge) of the Cu_2O_2 core is essential to clarifying the operation of dioxygen transport and would advance the design of synthetic catalysts that employ dioxygen as a terminal oxidant. There is significant interest in the biomimetic application of naturally occurring metal complexes for use in metallodrug design, with Cu(II) complexes recently employed in cancer therapeutics as artificial DNA metallonucleases⁴

and tyrosinase mimics.⁵

However, the formation of oxygenated hemocyanin (oxyHc) via the binding of O₂ to deoxygenated hemocyanin (Hc) remains a challenging problem. In particular, the binding of O₂ falls into the category of a spin-forbidden transition. Molecular O₂ is in a spin triplet configuration, and the Cu ions in deoxyHc are known to be in the Cu(I) d^{10} singlet configuration. The combination of triplet O₂ and singlet deoxyHc, to produce the Cu₂O₂ antiferromagnetic singlet in oxyHc, is believed to occur via a simultaneous charge transfer of one electron from each Cu(I) ion to O₂ forming a hybrid Cu(II)-peroxy-Cu(II) configuration. A superexchange pathway is hypothesised to form across the two Cu atoms.⁶ This mechanism is supported by superconducting quantum interference device (SQUID) measurements that report a large superexchange coupling⁷ between the two Cu centers, and a diamagnetic ground state.⁸

Despite intensive theoretical studies on the nature of the side-on coordinated Cu₂O₂ core, theoretical analysis has so far proved to be challenging for many electronic structure methods including ab initio quantum chemistry, density functional theory (DFT) and mixed quantum/classical (QM/MM) methods due to the complex simultaneous interplay of both the charge and spin, and to the multi-reference character of oxyHc. In particular, DFT and hybrid-DFT do not predict the correct singlet ground state due to the multi-reference nature of the ground state that is not accessible in DFT-based approaches.⁹⁻¹¹ To overcome the limitation of DFT techniques, a spin-projection method (also called spin-mixing) is often applied, whereby the different spin-polarised ground states are calculated individually,¹² and the entangled singlet is reconstructed by linear combination

of the respective Slater determinants (essentially a combination of the spin-broken symmetry state in the up-down, up-up, down-down configurations to extract an effective singlet state). Although this construction yields insights into the energetics, it does not allow for the study of excitations, and limits the scope of comparison with experimental data, such as the optical absorption,¹⁰ that is a stringent test of any theory. Furthermore, the spin-contamination present in spin-polarised hybrid DFT remains an issue,^{6,11,13-15} and typically the broken symmetry state wrongly becomes asymmetric in the oxyHc Cu_2O_2 core.¹⁰ Finally, experiments have also alluded to the necessity of characterising the oxyHc ground state as a mixed valence state,¹⁶ which cannot be accessed in the ground state DFT approach.

Meanwhile, multireference wave function methods have been used to extensively study the oxyHc core,^{13,14,17-22} but these approaches are not feasible for systems containing more than several dozen electrons. Flock and Pierloot¹⁸ argue that the inclusion of the imidazole ligands results in steric effects that are critical for a realistic description of oxygen containing dicopper systems, but most multireference wave function studies of this core consider a simplified model with ammonia ligands (including CC,^{13,14} CASPT2,¹⁸ MRCI,¹⁹ RASPT2,²⁰ and DMRG-CASPT2²¹), while others (such as DMRG²² and DMRG-CT²³) are limited to the experimentally inaccessible bare $\text{Cu}_2\text{O}_2^{2+}$ core. Furthermore, this system has large active space requirements given that it likely suffers from triplet instability,¹¹ and if the number of allowed excitations is too limited, size-extensivity errors arise.²⁰ Some of these methods also lack dynamical correlation contributions,²⁴⁻²⁶ and others strongly over-correct correlation effects.¹⁹

In this work, we apply an alternative strategy: dynamical mean-field theory (DMFT).²⁷ This method belongs to a broad category of embedding approaches and accounts for the limitations discussed above by treating the many-body effects and the superexchange of the di-Cu bridge explicitly (unlike DFT), while limiting this treatment to the correlated subspace of the copper $3d$ electrons, thereby side-stepping the prohibitive scaling of quantum chemistry methods. DMFT is a sophisticated method that includes quantum dynamical effects, and takes into account charge fluctuations, spin fluctuations, and thermal excitations. Although DMFT is routinely used to describe materials, it was also recently extended to molecular systems,^{28,29} and combined with the linear-scaling DFT software ONETEP to extend the applicability of DFT + DMFT to systems of unprecedented size.³⁰⁻³⁴

In order to characterise the superexchange mechanism between the Cu_2 d -orbitals and intermediate O p -orbitals, we must also use non-local DMFT (or “cluster DMFT”). Single-site DMFT invokes a mean-field approximation across the multiple correlated sites, and thus it can only treat the multiplet structure of each Cu atom separately. This is not the case for cluster DMFT, where all correlated sites are directly treated in a multi-site Anderson impurity model (AIM) — hence our approach might be denoted as DFT+AIM. We also carry out a self-consistency cycle over the charge density, as we work at fixed total number of electrons. This produces a feedback to the solution of the model that is similar in its nature to the mean-field approximation used in DMFT.

Our DFT + cluster DMFT approach identifies a regime where many-body effects — as characterised by the Hubbard interaction U — stabilise the singlet for the butterflyed Cu_2O_2 structure,

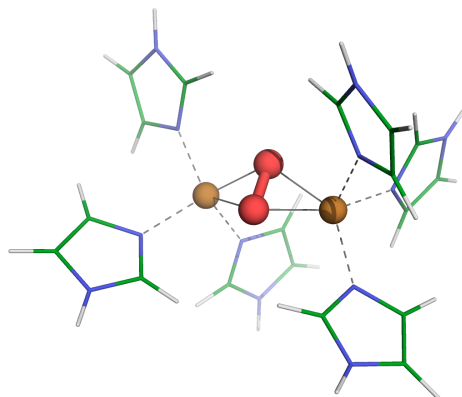


Figure 1: The oxyHc functional complex. Structure showing the Cu_2O_2 correlated subsystem, which is treated using dynamical mean field theory, and the surrounding imidazole rings representing the protein environment. Hydrogen, carbon, nitrogen, oxygen and iron atoms are shown in white, green, blue, red, and orange respectively.

in line with experiment. We show that a dominant spin singlet state is maximised at $U = 8$ eV and discuss the validation of the model by optical experiments. The obtained singlet is in the Heitler-London regime (an entangled quantum superposition of two localised magnetic moments), and is associated with incoherent scattering processes that reduce the lifetime of charge excitations.

Results

The spin state of the Cu_2O_2 core

This work presents the first DFT + DMFT simulations on a 58-atom model of the oxyHc functional complex (Fig. 1). While 58 atoms is within the reach of some less accurate quantum chemistry methods, the overhead of extending DFT + DMFT to include much more of the protein environ-

ment would come at an insignificant computational cost given our use of linear-scaling DFT.

In vivo, the Cu_2O_2 core exists in a low-spin (singlet) state, as identified by EPR.³⁵ In the model of Solomon and co-workers, this low-spin state is stabilised by superexchange via the O_2 ligand orbitals, which relies on the Cu_2O_2 core being planar (Fig. 2(a, b)).¹⁰ As the peroxide molecule unbinds, the core butterflies (i.e. the dioxygen moves up out of the plane, leaving the core in a bent configuration). Here, each Cu overlaps with a different oxygen π^* orbital on the peroxide (Fig. 2(c, d)). This removes the superexchange, and the triplet state becomes most favourable. If we measure planarity by $R = \left| \frac{1}{2}(\mathbf{r}_{\text{CuA}} + \mathbf{r}_{\text{CuB}}) - \frac{1}{2}(\mathbf{r}_{\text{O1}} + \mathbf{r}_{\text{O2}}) \right|$ — that is, the distance between the mean position of the two copper atoms and the mean position of the two oxygen atoms — the singlet-to-triplet transition occurs at $R = 0.6 \text{ \AA}$ using the B3LYP DFT functional.¹⁰

However, X-ray structures of the Cu_2O_2 core reveal that the bound singlet state is not planar. In oxyHc $R = 0.47 \text{ \AA}$, and in oxyTy $R = 0.63 \text{ \AA}$ — beyond the predicted singlet-to-triplet transition.^{36,37} QM/MM studies of the entire oxyHc protein (from which our model complex derives) obtain $R = 0.54$ to 0.71 \AA ; evidently, the protein scaffolding around the binding site prevents the core ever reaching the planar structure observed in model complexes.¹¹

With this in mind, we examined the reduced density matrix of our butterflyed model ($R = 0.68 \text{ \AA}$). This provides a detailed picture of the electronic structure of the $3d$ subspace of the Cu atoms (Fig. 3). Within DMFT, the density matrix is obtained by tracing the Anderson impurity model over the bath degrees of freedom, and gives an effective representation of the quantum states of the two Cu atoms. Note that in our approach, the ground-state wave-function is not a

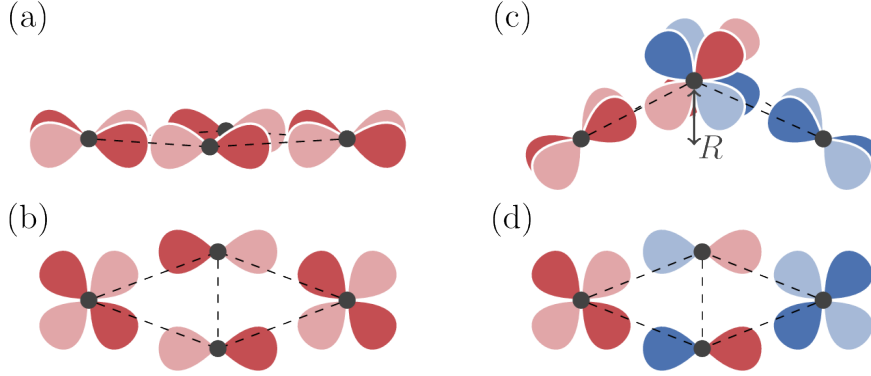


Figure 2: The superexchange model of Solomon and co-workers. The Cu_2O_2 core is depicted from side-on (a, c) and above (b, d). In the planar configuration (a, b), single ligand orbitals bridge the two copper sites, and superexchange is possible. In a bent configuration (c, d), the copper d orbitals overlap with different π^* orbitals. As these two sets of orbitals are orthogonal, hopping between the blue and the red subspaces is not possible and the superexchange mechanism breaks down.

pure state with a single allowed value for the spin states (singlet, doublet, triplet, etc.), yet we can describe the fluctuating spin states of the Cu atom by analysing the spin distribution obtained from the dominant configurations.

DMFT allows us to systematically investigate how many-body and finite-temperature effects alter these results. The Hubbard interaction U is known to be paramount to capture many-body effects. Several competing and relevant many-body phenomena stem from this term, including charge localisation, exchange of electrons, charge-transfer excitations, and stabilisation of magnetic multiplets. Consequently, the electronic structure of the $3d$ subspace is highly sensitive to the magnitude of this parameter. Although typical values for U can be obtained by linear response or constrained RPA,^{38,39} these predictions are typically dependent on the choice of local orbitals and

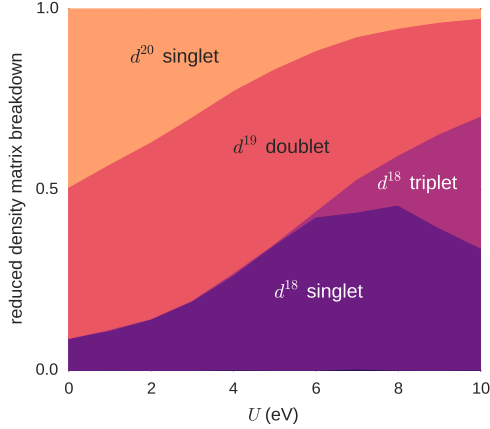


Figure 3: Decomposition of the reduced density matrix of the Cu_2 dimer in the different quantum sectors. The colours correspond to the respective weights of the different contributions for each value of the Coulomb repulsion U (if a colour occupies all the vertical axis, for example, it means that all eigenvectors of the density matrix are in that particular quantum sector). Note that the d occupation is the sum of both Cu sites (for example, d^{20} means both Cu atoms are in the d^{10} configuration).

basis representation. In this work we instead consider a range of values for U (from 0 to 10 eV). By artificially manipulating the magnitude of the local many-body effects, we can systematically investigate their influence on the electronic structure and derived properties.

The effect of U on the character of the reduced density matrix is shown in Fig. 3. In the weakly correlated regime ($U < 2$ eV) we obtain a large contribution from the d^{20} and d^{19} configurations, indicating that the average charge transfer from the Cu to O_2 involves less than one electron per Cu, thus preventing the formation of a singlet (as the Cu $3d$ orbitals are nearly full).

As U increases, the total electronic occupation of the Cu dimer decreases (Fig. 3 and Supplementary Table 1). In the range $U = 6 - 8$ eV, the d^{18} singlet component is maximised, beyond which d^{18} triplet excitations begin to contribute. The existence of this singlet is corroborated by the observed local magnetic moment and spin correlation (Supplementary Note 1 and Supplementary Figure 1). Interestingly, the von Neumann entropy (Supplementary Note 2 and Supplementary Figure 2) grows as U increases, pointing to the importance of many low-lying quantum states. Noting that $U \approx 8 - 8.5$ eV in the case of both molecules⁴⁰ and solids,⁴¹ it appears that in nature many-body quantum effects stabilise the low-spin singlet in spite of the butterfly structure of the Cu_2O_2 core (Fig. 2(c, d)).

The superexchange mechanism

Having identified this singlet in the Cu_2O_2 core at $U \approx 8$ eV, let us establish how it forms. Direct hopping between localised Cu d -orbitals is very unlikely due to the large distance by which they are separated, and therefore hopping must proceed via an intermediate oxygen p -orbital (i.e. superexchange).⁴²

The superexchange process can be pictured in the canonical hydrogen atom dimer system. This system describes a pair of up and down electrons that form a singlet state. In this picture, two different limits are possible: i) when the H atoms form a bond at short distance, and the up and down electron form a delocalised bound singlet (BS) centred on the bond, with a high degree of double occupancy, ii) in the dissociated case, known as the Heitler-London (HL) limit, where the

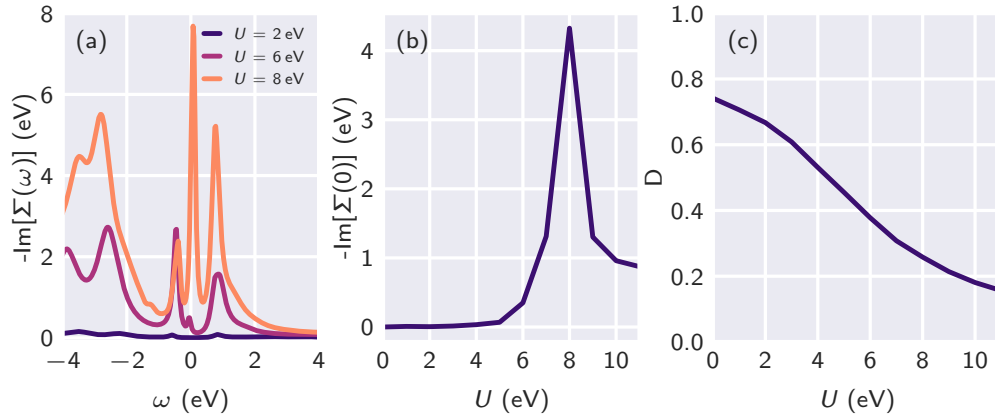


Figure 4: Dynamical mean field self energy. a) Imaginary part of the dynamical mean field local self energy $\Sigma(\omega)$ of the Cu-3d empty orbital for Hubbard $U = 2$ eV, 6 eV, and 8 eV. At $U = 8$ eV, we obtain incoherent excitations at $\omega = 0$ eV. b) The imaginary local self energy at $\omega = 0$ and c) the double occupancy of the partially-filled 3d orbitals D as a function of U . Note that although the double occupancy is evolving smoothly with the Coulomb interaction U , $\Sigma(\omega = 0)$ shows a sharp increase near $U = 8$, associated with the stabilisation of a localised singlet.

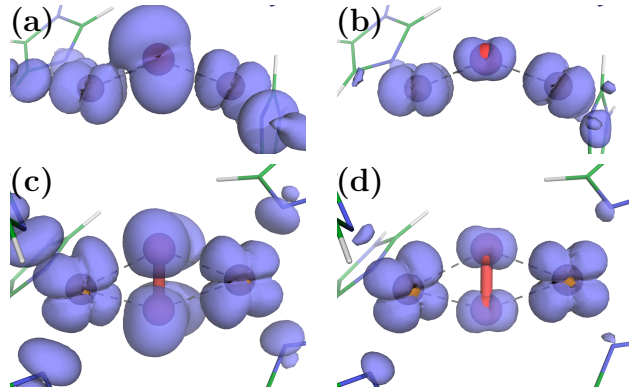


Figure 5: Molecular orbital isosurfaces. The isosurface for the highest occupied molecular orbital (a, c) and lowest unoccupied molecular orbital (b, d) densities for $U = 8$ eV, as viewed side-on (a, b) and above (c, d). Note that because these are extracted from the Green's function via the spectral density, the phase of the orbitals is inaccessible.

H atoms are far apart and the singlet is a true quantum entangled state of the singly occupied H orbitals. In the latter case, the molecular orbital contains only one electron and double occupancy is zero.

In the HL limit, which typically occurs in the limit of dissociation, the charge is localised around the H atoms. However, this effect may also occur in systems where the local Hubbard Coulomb repulsion U acts as a Coulomb blockade: many-body effects prevent long-lived charge transfer excitations, and the Coulomb repulsion energy is reduced at the expense of the kinetic energy. A signature of the blockade is typically a large increase in the self energy at the Fermi level (pole structure), indicating charge localisation and incoherent scattering associated with a short lifetime of charge excitations.

To investigate the nature of the singlet (BS or HL), we report in Fig. 4(a) the computed self energy of the Cu $3d$ subspace, for various values of U . We obtain a qualitative difference between $U = 6$ eV and $U = 8$ eV, where at $U = 8$ eV the self energy develops a pole at $\omega = 0$. The formation of the pole is particular to $U = 8$ eV (Fig. 4(b)), and is associated with the regime where excitations are incoherent, which prevents long-lived charge transfer excitations from the Cu $3d$ orbitals to O_2 . In this limit, the many body effect acts as a Coulomb blockade and the charge is in turn localised, with weak direct coupling (HL limit). For $U = 6$ eV, the singlet is in the BS limit, where charge excitations allow a direct electron transfer across the oxo-bridge. Note that the observation of the BS-HL crossover is not apparent in averaged quantities, such as in the double occupancies (Fig. 4(c)), which evolve smoothly with the Coulomb repulsion.

Turning back to the physical system, superexchange relies on a single ligand orbital linking the two copper sites. (Superexchange pathways via two p orbitals are possible, but they give rise to ferromagnetic coupling that is significantly weaker than antiferromagnetic coupling from single-ligand orbital pathways⁴³). Examination of the molecular orbitals near the Fermi level (Fig. 5) reveals that, for the HOMO of the bent Cu_2O_2 structure, electronic density of the oxygen ligand is directed into the copper plane, thus providing a pathway for the antiferromagnetic superexchange that we observe.

Interestingly, we note that these molecular orbitals differ in their energetic ordering compared to those from DFT studies of planar model complexes.⁸ In particular, the HOMO involves hybridisation with oxygen π^* rather than σ^* orbitals (with the σ^* oxygen orbital featuring approximately

3 eV above the Fermi level). This reordering will have substantial ramifications for the potential catalytic pathways, especially considering the importance of the σ^* orbital to bond-breaking. This picture is confirmed by natural bond orbital (NBO) analysis, the results of which are presented in the Supplementary Note 3.

Optical transitions

As a validation of the DFT + DMFT computational model, and to identify the strength of correlations in oxyHc, we extracted the optical absorption spectrum of ligated hemocyanin (Fig. 6).

As experiments are performed in the gaseous phase and not in a single crystal, we have calculated the isotropic and anisotropic components of the dielectric tensor. The former involves only pair correlators along the same spatial directions, whereas the latter also incorporates non-diagonal terms (note that both are causal quantities), which are important in oxyHc as the local coordination axes of the Cu atoms are not aligned. Remarkable agreement is obtained for $U = 8$ eV, with however a blue shift at high frequency ($\omega > 4$ eV), and a concomitant transfer of optical weight to lower energies ($\omega < 3$ eV). The position of the peak at approximately 3.5 eV is in excellent agreement. The experimental features at 3 eV (see arrow) are also visible in the theoretical calculations. The peaks at 1.8 eV and 2.2 eV contribute to the long infrared tail of the optical weight down to 1.5 eV. We note that the main difference between the BS ($U \leq 6$ eV) and the HL singlet ($U = 8$ eV) is the presence of several large peaks associated with charge transfer from ligand to Cu $3d$ orbitals below 2.5 eV, which are significantly smaller/absent in experiment.^{45,46} The suppression of the optical

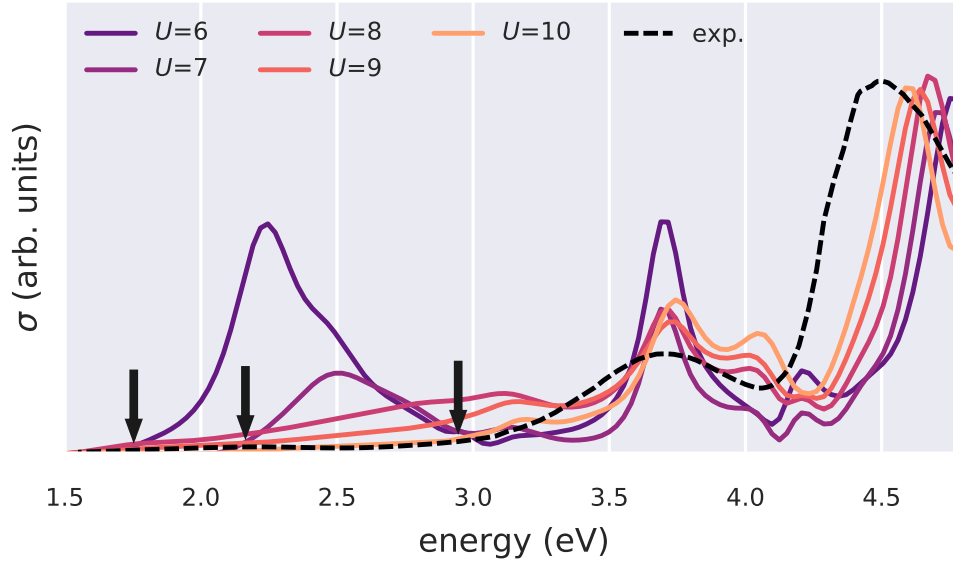


Figure 6: Theoretical optical absorption of the Cu_2O_2 core and imidazole rings obtained by dynamical mean field theory for values of the Coulomb repulsion $U = 6$ to 10 eV. For comparison, we show the experimental optical absorption⁴⁴ in a wide range of wavelengths (infrared to UV). There are several smaller peaks in the experimental spectra that are not visible at this scale (indicated with arrows).

weight at 2 eV is due to a large increase in incoherent scattering at $\omega = 0$ eV at $U = 8$ eV (Fig. 4), associated with the localisation of the holes in the Cu $3d$ shell at $U = 8$ eV. The strong suppression of the optical weight in the near infrared regime and the consistent agreement with our calculations shows that the di-Cu singlet is in the HL regime. In comparison, DFT, without extensions, puts a strong emphasis on the near infrared peak in the optical absorption¹⁰ because the aforementioned scattering processes are absent at this level of theory. (See also the very small HOMO-LUMO gap in Supplementary Fig. 4 for small values of U .)

The absorption spectrum of this protein has been reported experimentally to be qualitatively dependent on its ligation state. In its oxygenated form, a weak peak at 570 nm (approximately 2.2 eV, see arrow) is attributed to ligand-to-metal charge transfer^{45,46} from the O₂ π anti-bond with lobes oriented perpendicular to the metal atoms. This orbital is denoted “ π_v^* ”; the π anti-bond with lobes directed towards the copper atoms is denoted π_σ^* and is responsible for an experimentally observed (and much larger) peak at approximately 3.6 eV. Our calculations accurately reproduce this large peak, associating it with a transition from the weight at -1 eV in the density of states (Supplementary Fig. 4), which is localised on Cu_A/Cu_B/imidazole, to the LUMO, which in our calculations is a hybridised state between the Cu and the π_σ^* O₂, with dominant Cu 3*d* character (Fig. 5b and d). Meanwhile, a small peak at 2 eV exists corresponding to the HOMO-to-LUMO transition (π_v^* to Cu charge transfer). Overall, comparison with the experimental spectrum suggests $U \geq 8$ eV which is compatible with the value of U that gives rise to the singlet. We note that time dependent density functional theory (TDDFT) calculations reproduce the absorption peak at 3.6 eV, but not the peak at 4.5 eV. More importantly, the ground state electronic structure at this level of theory is not the antiferromagnetic singlet state observed experimentally, thus agreement between TDDFT and experiment likely involves some cancellation of errors.

Discussion

We have presented the application of a DFT + cluster DMFT approach, designed to treat strong electronic interaction and multi-reference effects, to oxyHc, a molecule of important biological function. The reduced density matrix of the 3*d* subspace of the two Cu atoms revealed the presence

of fluctuating spin-states, in which a $\text{Cu}_2 d^{18}$ singlet component is maximised at $U = 8 \text{ eV}$ in spite of the butterfly distortion of the Cu_2O_2 core. The Hubbard U is necessary to capture the multi-reference character of the ground-state, placing oxyHc in the limit of a true quantum entangled singlet in the limit of the Heitler-London model, while the highest occupied molecular orbital has been shown to provide a pathway for antiferromagnetic superexchange. This work provides a starting point for studying biological activity of oxyHc and related type 3 Cu-based enzymes by (a) establishing that the singlet can survive the butterfly distortion, thereby resolving a prior inconsistency between structural data, spectroscopy, and first-principles calculations, and (b) by providing a framework for subsequent studies to account for the effect of the protein “scaffolding” in which the active site sits, as well as the effects of strong electronic correlation. Our approach reproduced the experimentally-observed peaks in the absorption spectrum at around 2.2 eV, 3.7 eV, and 4.5 eV.

The DFT + DMFT method provides a useful middle-ground between (a) density-functional theory based simulations of metalloproteins, which depend heavily on the choice of functional and incorrectly describe multi-reference effects, and (b) single- or multi-reference quantum chemistry approaches, whose unfavourable scaling prevents studies on models of realistic size. More generally, our particular implementation of DMFT within linear-scaling DFT software³² can account for strongly correlated electronic behaviour while simultaneously including the effects of protein environments, making it ideally suited for studying biological activity in a wide range of transition metal-containing proteins.^{47,48}

Methods

Geometry

The geometry of the 58 atom system was obtained from the literature.¹¹ This had been optimised using the B3LYP hybrid functional and closely matches the experimentally observed structure.^{11,36}

The coordinates of this structure are provided in Supplementary Note 4.

DFT

The DFT ground-state was obtained using ONETEP,³⁰ which is a linear-scaling DFT code that is formally equivalent to a plane-wave method. Linear-scaling is achieved by the *in situ* variational optimisation of its atom-centered basis set (spatially-truncated nonorthogonal generalised Wannier functions, or NGWFs).⁴⁹ The total energy is directly minimised with respect to the NGWFs and the single-particle density matrix. The use of a minimal, optimised Wannier function representation of the density-matrix allows for the DFT ground state to be solved with relative ease in large systems. This is particularly useful in molecules, since explicit truncation of the basis functions ensures that the addition of vacuum does not increase the computational cost.⁵⁰

The DFT calculations of the oxyHc system were run with an energy cut-off of 897 eV, using the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional.⁵¹ Nine NGWFs were employed on the copper atoms, four on each carbon/nitrogen/oxygen, and one on each hydrogen. Spin symmetry was imposed. NGWFs were truncated using 7 Å cutoff radii. Open boundary conditions were achieved via a padded cell and a Coulomb cut-off.⁵² The Hubbard projectors were constructed

from the Kohn-Sham solutions to an isolated copper pseudopotential.⁵³ The pseudopotentials were generated with the OPIUM pseudopotential generation project.⁵⁴ These pseudopotentials partially account for scalar relativistic effects. (Studies have demonstrated that relativistic effects can play a role in the electronic structure of the Cu_2O_2 core.⁵⁵)

DMFT

We refined our DFT calculations using the DFT+DMFT method^{27,56} in order to obtain a more accurate treatment of strong electronic correlation effects. The oxyHc model was mapped, within DMFT, to an Anderson impurity model (AIM) Hamiltonian,⁵⁷ and we used a recently developed extended Lanczos solver⁵⁸ to obtain the DMFT self energy. The convergence of the mapping is shown in Supplementary Fig. 5 and discussed in Supplementary Note 5.

To identify the best spatial representation of the local d -space in the projected Anderson impurity model, we first identified the orthogonal transformation which reduces the off-diagonal elements of the local Green's function, for respectively the Cu_A and Cu_B sites. Then, we implemented a minimisation procedure which finds the closest corresponding real space $SO(3)$ rotation of the local Cartesian axis corresponding to the $O(5)$ orthonormal transformation in d -space. This provides a set of local axes, different on each Cu atom, which make the local Green's function nearly diagonal in frequency space. These axes are plotted in Supplementary Figure 6 and discussed in Supplementary Note 6. As a result, we observed that the Cu_A and Cu_B holes are of pure orbital character within this local axis representation. The validity of this approach was vindicated by the NBO analysis, which when performed on the DMFT density revealed a hole in one $3d$ or-

bital for each Cu atom, confirming the expected Cu(II) oxidation state $3d^94s^0$. (See Supplementary Note 3 for details.)

Since only a single impurity site ($3d$ orbital subspace) is present, the system becomes crystal momentum independent in the molecular limit, and since the Kohn-Sham Green's function is computed in full before projection onto the impurity subspace, the Anderson impurity mapping is effectively exact, and the necessity of invoking the DMFT self-consistency is not required. However, in DFT + DMFT there is also a charge self-consistency cycle, where (i) the chemical potential can be updated to ensure particle conservation and/or (ii) the DFT + DMFT density kernel can be used to generate a new Kohn-Sham Hamiltonian, which in turn provides a new input to DMFT; the procedure being repeated until convergence is achieved.^{59–61} In this work, we updated the chemical potential but not the Hamiltonian due to computational cost.

To obtain the Kohn-Sham Green's function, we performed the matrix inversion, as well as all matrix multiplications involved in the DMFT algorithm, on graphical processing units (GPUs) using a tailor-made parallel implementation of the Cholesky decomposition written in the CUDA programming language. Electronic correlation effects are described within the localised subspace by the Slater-Kanamori form of the Anderson impurity Hamiltonian,^{62,63} specifically:

$$\begin{aligned} \mathcal{H}_U = & U \sum_{\mathbf{m}} n_{\mathbf{m}\uparrow} n_{\mathbf{m}\downarrow} + \left(U' - \frac{J}{2} \right) \sum_{\mathbf{m} > \mathbf{m}'} n_{\mathbf{m}} n_{\mathbf{m}'} \\ & - J \sum_{\mathbf{m} > \mathbf{m}'} \left(2\mathbf{S}_{\mathbf{m}} \mathbf{S}_{\mathbf{m}'} + \left(d_{\mathbf{m}\uparrow}^\dagger d_{\mathbf{m}\downarrow}^\dagger d_{\mathbf{m}'\uparrow} d_{\mathbf{m}'\downarrow} \right) \right), \end{aligned} \quad (1)$$

where m, m' are orbital indices, $d_{m\sigma}$ ($d_{m\sigma}^\dagger$) annihilates (creates) an electron with spin σ in the orbital m , and n_m is the orbital occupation operator.

The first term describes the effect of intra-orbital Coulomb repulsion, parameterised by U , and the second term describes the inter-orbital repulsion, proportional to $U' = U - 2J$, which is renormalised by the Hund's exchange coupling parameter J in order to ensure a fully rotationally invariant Hamiltonian (for further information on this topic, we refer the reader to Imada et al.⁶⁴) The third term is the Hund's rule exchange coupling, described by a spin exchange coupling of amplitude J . S_m denotes the spin corresponding to orbital m , so that $S_m = \frac{1}{2}d_{ms}^\dagger \vec{\sigma}_{ss'} d_{ms'}$, where $\vec{\sigma}$ is the vector of Pauli matrices.

Our DMFT calculations were carried out at room temperature, $T = 293$ K and the Hubbard U was varied in the range 0–10 eV, with a fixed Hund's coupling $J = 0.8$ eV. The theoretical optical absorption was obtained in DFT + DMFT within the linear-response regime (Kubo formalism), in the no-vertex-corrections approximation,⁶⁵ which is given by:

$$\sigma(\omega) = \frac{2\pi e^2 \hbar}{\Omega} \int d\omega' \frac{f(\omega' - \omega) - f(\omega')}{\omega} \times (\rho^{\alpha\beta}(\omega' - \omega) \mathbf{v}_{\beta\gamma} \rho^{\gamma\delta}(\omega') \cdot \mathbf{v}_{\delta\alpha}), \quad (2)$$

where the factor of two accounts for spin-degeneracy, Ω is the simulation-cell volume, e is the electron charge, \hbar is the reduced Planck constant, $f(\omega)$ is the Fermi-Dirac distribution, and $\rho^{\alpha\beta}$ is the density-matrix given by the frequency-integral of the interacting DFT + DMFT Green's function. The matrix elements of the velocity operator, $\mathbf{v}_{\alpha\beta}$, noting that we do not invoke the Peierls substitution,⁶⁵ are given by:

$$\mathbf{v}_{\alpha\beta} = -\frac{i\hbar}{m_e} \langle \phi_\alpha | \nabla | \phi_\beta \rangle + \frac{i}{\hbar} \langle \phi_\alpha | [\hat{V}_{nl}, \mathbf{r}] | \phi_\beta \rangle. \quad (3)$$

This expression is general to the NGWF representation used in this work,⁶⁶ where the contribution

to the non-interacting Hamiltonian due to the non-local part of the norm-conserving pseudopotentials,^{67,68} represented by \hat{V}_{nl} , is included.

Finally, the double-counting correction E_{DC} must be introduced, since the contribution of interactions between the correlated orbitals to the total energy is already partially included in the exchange-correlation potential derived from DFT. The most commonly used form of the double-counting term is⁶⁹

$$E_{dc} = \frac{U^{av}}{2}n_d(n_d - 1) - \frac{J}{2} \sum_{\sigma} n_{d\sigma}(n_{d\sigma} - 1). \quad (4)$$

where

$$U^{av} = \frac{U + 2(N - 1)U'}{2N - 1} \quad (5)$$

with N being the number of orbitals spanning the correlated subspace (and recall that $U' = U - 2J$). This double-counting is derived by attempting to subtract the DFT contributions in an average way; U^{av} is the average of the intra- and inter-orbital Coulomb parameters.

Data availability

The data underlying this publication are available on the Apollo - University of Cambridge Repository: <https://doi.org/10.17863/CAM.46223>

Code availability

The ONETEP code version 5 is available from www.onetep.org.

1. Ghiretti, F. The decomposition of hydrogen peroxide by hemocyanin and by its dissociation products. *Arch. Biochem. Biophys.* **63**, 165–176 (1956).
2. Halfen, J. A. *et al.* Reversible Cleavage and Formation of the Dioxygen O-O Bond Within a Dicopper Complex. *Science* **271**, 1397–1400 (1996).
3. Duckworth, H. W. & Coleman, J. E. Physicochemical and kinetic properties of mushroom tyrosinase. *J. Biol. Chem.* **245**, 1613–25 (1970).
4. McGivern, T. J. P., Afsharpour, S. & Marmion, C. J. Copper complexes as artificial DNA metallonucleases: from Sigman's reagent to next generation anti-cancer agent? *Inorganica Chim. Acta* **472**, 12–39 (2018).
5. Nunes, C. J. *et al.* Reactivity of dinuclear copper(II) complexes towards melanoma cells: correlation with its stability, tyrosinase mimicking and nuclease activity. *J. Inorg. Biochem.* **149**, 49–58 (2015).
6. Gherman, B. F. & Cramer, C. J. Quantum chemical studies of molecules incorporating a $\text{Cu}_2\text{O}_2^{2+}$ core. *Coord. Chem. Rev.* **253**, 723–753 (2009).
7. Dooley, D. M., Scott, R. A., Ellinghaust, J., Solomont, E. I. & Gray, H. B. Magnetic susceptibility studies of laccase and oxyhemocyanin (copper proteins/variable temperature measurements/antiferromagnetism). *Proc. Natl. Acad. Sci. U. S. A.* **75**, 3019–3022 (1978).
8. Solomon, E. I. *et al.* Copper dioxygen (bio)inorganic chemistry. *Faraday Discuss.* **148**, 11–108 (2011).

9. Takano, Y. *et al.* Theoretical studies on the magnetic interaction and reversible dioxygen binding of the active site in hemocyanin. *Chem. Phys. Lett.* **335**, 395–403 (2001).
10. Metz, M. & Solomon, E. I. Dioxygen binding to deoxyhemocyanin: electronic structure and mechanism of the spin-forbidden two-electron reduction of O_2 . *J. Am. Chem. Soc.* **123**, 4938–4950 (2001).
11. Saito, T. & Thiel, W. Quantum mechanics/molecular mechanics study of oxygen binding in hemocyanin. *J. Phys. Chem. B* **118**, 5034–5043 (2014).
12. Cohen, A. J., Tozer, D. J. & Handy, N. C. Evaluation of $\langle \hat{S}^2 \rangle$ in density functional theory. *J. Chem. Phys.* **126**, 214104 (2007).
13. Cramer, C. J., Kinal, A., Włoch, M., Piecuch, P. & Gagliardi, L. Theoretical characterization of end-on and side-on peroxide coordination in ligated Cu_2O_2 models. *J. Phys. Chem. A* **110**, 11557–11568 (2006).
14. Cramer, C. J., Włoch, M., Piecuch, P., Puzzarini, C. & Gagliardi, L. Theoretical models on the Cu_2O_2 torture track: mechanistic implications for oxytyrosinase and small-molecule analogues. *J. Phys. Chem. A* **110**, 1991–2004 (2006).
15. Siegbahn, P. E. M. The performance of hybrid DFT for mechanisms involving transition metal complexes in enzymes. *J. Biol. Inorg. Chem.* **11**, 695–701 (2006).
16. Loeb L, B., Crivelli P, I. & Andrade P, C. Oxy-hemocyanin: a peroxo copper (II) complex? A mixed-valence alternative view. *Comments Inorg. Chem.* **20**, 1–26 (1998).

17. Roemelt, M., Guo, S. & K-L Chan, G. A projected approximation to strongly contracted N-electron valence perturbation theory for DMRG wavefunctions. *J. Chem. Phys.* **144**, 204113 (2016).
18. Flock, M. & Pierloot, K. Theoretical study of the interconversion of O₂-binding dicopper complexes. *J. Phys. Chem. A* **103**, 95–102 (1999).
19. Rode, M. F. & Werner, H.-J. Ab initio study of the O₂ binding in dicopper complexes. *Theor Chem Acc* **114**, 309–317 (2005).
20. Malmqvist, Å. *et al.* Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **128**, 5648 (2008).
21. Kurashige, Y., Chalupský, J., Lan, T. N. & Yanai, T. Complete active space second-order perturbation theory with cumulant approximation for extended active-space wavefunction from density matrix renormalization group. *J. Chem. Phys.* **141**, 174111 (2014).
22. Kurashige, Y. & Yanai, T. High-performance ab initio density matrix renormalization group method: Applicability to large-scale multireference problems for metal compounds. *J. Chem. Phys.* **130**, 234114 (2009).
23. Yanai, T., Kurashige, Y., Neuscamman, E. & Chan, G. K.-L. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J. Chem. Phys.* **132**, 024105 (2010).
24. Malmqvist, P.-Å. & Roos, B. O. The CASSCF state interaction method. *Chem. Phys. Lett.* **155**, 189–194 (1989).

25. Roos, B. O. The complete active space self-consistent field method and its applications in electronic structure calculations. In *Adv. Chem. Phys. Ab Initio Methods Quantum Chem. Part 2, Vol. 69*, 399–445 (John Wiley & Sons, Ltd, 1987).
26. Roos, B. O., Taylor, P. R. & Siegbahn, P. E. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **48**, 157–173 (1980).
27. Georges, A., Kotliar, G., Krauth, W. & Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.* **68**, 13–125 (1996).
28. Lin, N., Marianetti, C. A., Millis, A. J. & Reichman, D. R. Dynamical mean-field theory for quantum chemistry. *Phys. Rev. Lett.* **106**, 96402 (2011).
29. Jacob, D., Haule, K. & Kotliar, G. Dynamical mean-field theory for molecular electronics: electronic structure and transport properties. *Phys. Rev. B* **82**, 195115 (2010).
30. Skylaris, C.-K., Haynes, P. D., Mostofi, A. A. & Payne, M. C. Introducing ONETEP: linear-scaling density functional simulations on parallel computers. *J. Chem. Phys.* **122**, 084119 (2005).
31. Weber, C. *et al.* Vanadium dioxide: a Peierls-Mott insulator stable against disorder. *Phys. Rev. Lett.* **108**, 256402 (2012).
32. Linscott, E. B., Cole, D. J., Hine, N. D. M., Payne, M. C. & Weber, C. ONETEP + TOSCAM: uniting dynamical mean field theory and linear-scaling density functional theory (2019). [arXiv:1911.07752](https://arxiv.org/abs/1911.07752).

33. Weber, C. *et al.* Importance of many-body effects in the kernel of hemoglobin for ligand binding. *Phys. Rev. Lett.* **110**, 106402 (2013).
34. Weber, C., Cole, D. J., O'Regan, D. D. & Payne, M. C. Renormalization of myoglobin–ligand binding energetics by quantum many-body effects. *Proc. Natl. Acad. Sci.* **111**, 5790–5795 (2014).
35. Nakamura, T. & Mason, H. An electron spin resonance study of copper valence in oxyhemocyanin. *Biochem. Biophys. Res. Commun.* **3**, 297–299 (1960).
36. Magnus, K. A. *et al.* Crystallographic analysis of oxygenated and deoxygenated states of arthropod hemocyanin shows unusual differences. *Proteins Struct. Funct. Genet.* **19**, 302–309 (1994).
37. Matoba, Y., Kumagai, T., Yamamoto, A., Yoshitsu, H. & Sugiyama, M. Crystallographic evidence that the dinuclear copper center of tyrosinase is flexible during catalysis. *J. Biol. Chem.* **281**, 8981 (2006).
38. Scherlis, D. A., Cococcioni, M., Sit, P. & Marzari, N. Simulation of heme using DFT + U: a step toward accurate spin-state energetics. *J. Phys. Chem. B* **111**, 7384–7391 (2007).
39. Linscott, E. B., Cole, D. J., Payne, M. C. & O'Regan, D. D. Role of spin in the calculation of Hubbard U and Hund's J parameters from first principles. *Phys. Rev. B* **98**, 235157 (2018).
40. Didziulis, S. V., Cohen, S. L., Gewirth, A. A. & Solomon, E. I. Variable photon energy photoelectron spectroscopic studies of copper chlorides: an experimental probe of metal-ligand

- bonding and changes in electronic structure on ionization. *J. Am. Chem. Soc.* **110**, 250–268 (1988).
41. Anisimov, V. I., Zaanen, J. & Andersen, O. K. Band theory and Mott insulators: Hubbard U instead of Stoner I . *Phys. Rev. B* **44**, 943–954 (1991).
 42. Pavarini, E., Koch, E., Anders, F. & Jarrel, M. (eds.) *Correlated electrons: from models to materials*, vol. 2 (Forschungszentrum Jülich GmbH, Jülich, 2012).
 43. Anderson, P. W. Theory of magnetic exchange interactions: Exchange in insulators and semiconductors. *Solid State Phys.* **14**, 99–214 (1963).
 44. Andersen, N. H., Zoppellaro, G., Bubacco, L., Casella, L. & Andersson, K. K. Raman, UV–vis, and CD spectroscopic studies of dodecameric oxyhemocyanin from *Carcinus aestuarii*. *Chem. Lett.* **40**, 1360–1362 (2011).
 45. Himmelwright, R. S., Eickman, N. C., LuBien, C. D. & Solomon, E. I. Chemical and spectroscopic comparison of the binuclear copper active site of mollusc and arthropod hemocyanins. *J. Am. Chem. Soc.* **102**, 5378–5388 (1980).
 46. Heirwegh, K., Borginon, H. & Lontie, R. Separation and absorption spectra of α - and β -haemocyanin of *Helix pomatia*. *Biochim. Biophys. Acta* **48**, 517–526 (1961).
 47. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M. & Thauer, R. K. Active sites of transition-metal enzymes with a focus on nickel. *Curr. Opin. Struct. Biol.* **8**, 749–758 (1998).

48. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103 (2015).
49. Skylaris, C.-K., Mostofi, A. A., Haynes, P. D., Diéguez, O. & Payne, M. C. Nonorthogonal generalized Wannier function pseudopotential plane-wave method. *Phys. Rev. B* **66**, 035119 (2002).
50. Cole, D. J. & Hine, N. D. M. Applications of large-scale density functional theory in biology. *J. Phys. Condens. Matter* **28**, 393001 (2016).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
52. Hine, N. D. M., Dzielic, J., Haynes, P. D. & Skylaris, C.-K. Electrostatic interactions in finite systems treated with periodic boundary conditions: application to linear-scaling density functional theory. *J. Chem. Phys.* **135**, 204103 (2011).
53. Ruiz-Serrano, Á., Hine, N. D. M. & Skylaris, C.-K. Pulay forces from localized orbitals optimized in situ using a psinc basis set. *J. Chem. Phys.* **136**, 234101 (2012).
54. Rappe, A. M., Rabe, K. M., Kaxiras, E. & Joannopoulos, J. D. Optimized pseudopotentials. *Phys. Rev. B* **41**, 1227–1230 (1990).
55. Liakos, D. G. & Neese, F. Interplay of Correlation and Relativistic Effects in Correlated Calculations on Transition-Metal Complexes: The $(\text{Cu}_2\text{O}_2)^{2+}$ Core Revisited. *J. Chem. Theory Comput.* **7**, 1511–1523 (2011).

56. Maier, T. A., Pruschke, T. & Jarrell, M. Angle-resolved photoemission spectra of the Hubbard model. *Phys. Rev. B* **66**, 075102 (2002).
57. da Silva, L. G. G. V. D., Tiago, M. L., Ulloa, S. E., Reboredo, F. A. & Dagotto, E. Many-body electronic structure and Kondo properties of cobalt-porphyrin molecules. *Phys. Rev. B* **80**, 155443 (2009).
58. Aichhorn, M., Daghofer, M., Evertz, H. G. & von der Linden, W. Low-temperature Lanczos method for strongly correlated systems. *Phys. Rev. B* **67**, 161103 (2003).
59. Pourovskii, L. V., Amadon, B., Biermann, S. & Georges, A. Self-consistency over the charge density in dynamical mean-field theory: a linear muffin-tin implementation and some physical implications. *Phys. Rev. B* **76**, 235101 (2007).
60. Park, H., Millis, A. J. & Marianetti, C. A. Computing total energies in complex materials using charge self-consistent DFT + DMFT. *Phys. Rev. B* **90**, 235103 (2014).
61. Bhandary, S., Assmann, E., Aichhorn, M. & Held, K. Charge self-consistency in density functional theory combined with dynamical mean field theory: k -space reoccupation and orbital order. *Phys. Rev. B* **94**, 155131 (2016).
62. Slater, J. C. The ferromagnetism of nickel. *Phys. Rev.* **49**, 537–545 (1936).
63. Kanamori, J. Superexchange interaction and symmetry properties of electron orbitals. *J. Phys. Chem. Solids* **10**, 87–98 (1959).

64. Imada, M., Fujimori, A. & Tokura, Y. Metal-insulator transitions. *Rev. Mod. Phys.* **70**, 1039–1263 (1998).
65. Millis, A. J. Optical conductivity and correlated electron physics. In *Strong interactions in low dimensions.*, vol. 25, 195–235 (Springer, 2004).
66. Halpern, V. & Bergmann, A. Calculation of electronic green functions using nonorthogonal basis functions: application to crystals. *J. Phys. C Solid State Phys.* **5**, 1953 (1972).
67. Ratcliff, L. E., Hine, N. D. M. & Haynes, P. D. Calculating optical absorption spectra for large systems using linear-scaling density functional theory. *Phys. Rev. B* **84**, 165131 (2011).
68. Read, A. J. & Needs, R. J. Calculation of optical matrix elements with nonlocal pseudopotentials. *Phys. Rev. B* **44**, 13071–13073 (1991).
69. Pruschke, T. & Zöfl, M. *Electronic structure and ordered phases in transition metal oxides: application of the dynamical mean-field theory*, 251–265 (Springer, 2000).

Acknowledgements This work was supported by BBSRC (grant BB/M009513/1), EPSRC (grants EP/N02396X/1, EP/L015552/1) and the Rutherford Foundation Trust. The Flatiron Institute is a division of the Simons Foundation. C.W. gratefully acknowledges the support of NVIDIA Corporation with the donation of the Tesla K40 GPUs used for this research. For computational resources, we were supported by the ARCHER UK National Supercomputing Service and the UK Materials and Molecular Modelling Hub for computational resources (EPSRC Grant No. EP/P020194/1). We are grateful to Prof. Christopher Dennison (Newcastle University) for comments.

Author contributions C. W. and M. A. B. conceived and planned the research. M. A. B., E. B. L. and C. W. performed the calculations. M. A. B. , E. B. L. , A. G. , D. J. C. and C. W. analysed the data and contributed to the final manuscript.

Competing Interests The authors declare no competing interests.