Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

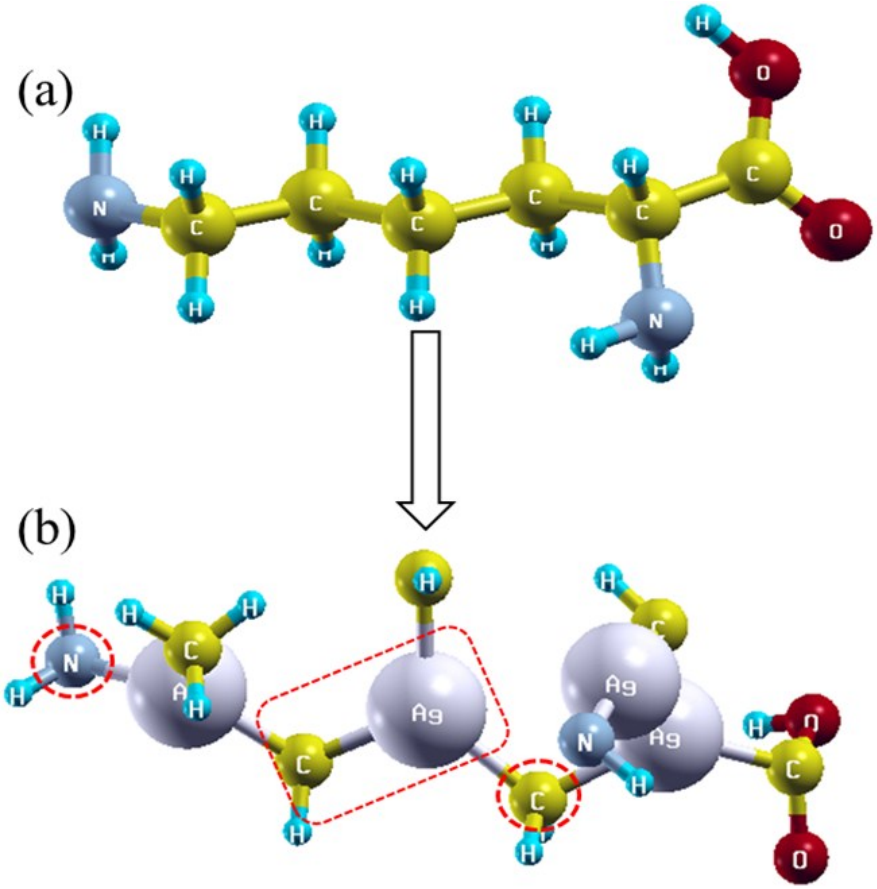
Supplementary Material

Driving the electrocatalytic activity by interface electronic structure control in metalloprotein hybrid catalyst for efficient hydrogen evolution

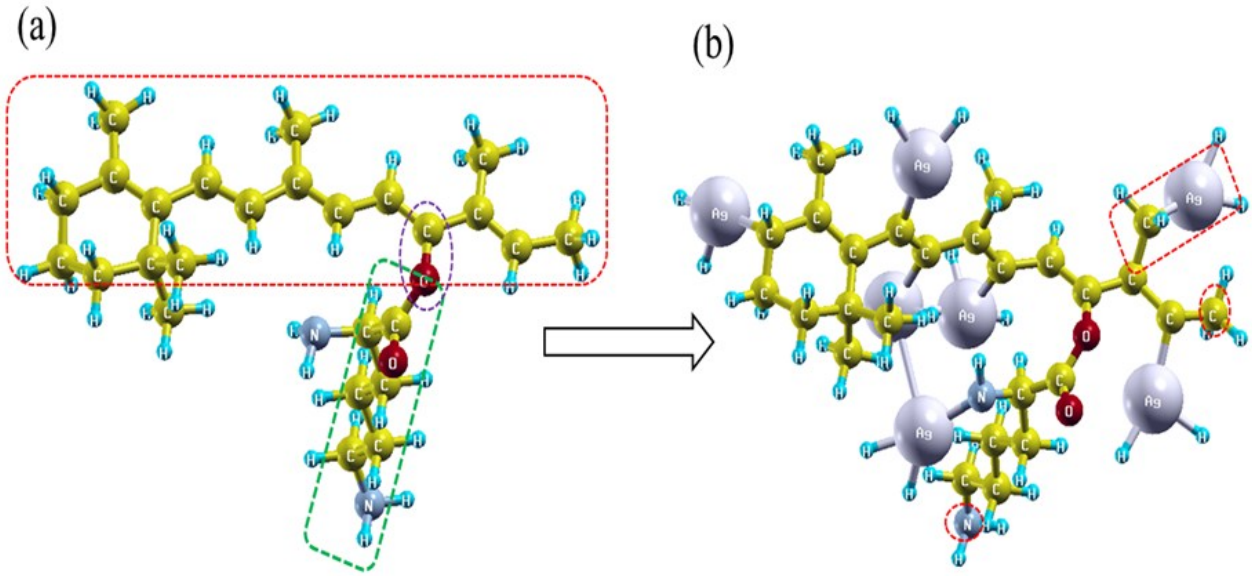
Sushant Kumar Behera, Pritam Deb*, Arghya Ghosh

Advanced Functional Material Laboratory (AFML), Department of Physics, Tezpur University (Central University), Tezpur, Assam-784028, India.

* Corresponding author
Email address: pdeb@tezu.ernet.in (Pritam Deb



Supplementary Figure 1 Geometry optimized atomic configurations. (a) The configuration of LYS unit with atoms named in the figure itself. **(b)** Its counterpart configuration incorporated with silver ions. The red dashed circles show the active hydrogen absorption sites both at carbon and nitrogen atoms. The rectangular red dashed circle shows the interface between silver and protein active site.



Supplementary Figure 2 Geometry optimized atomic configurations. (a) The configuration of LYS unit with atoms named in the figure itself. The rectangular red and green dashed lines show the configuration of retinal chromophore (rc) and LYS unit respectively. The violet dashed circle shows the attachment of rc with LYS unit. **(b)** Its counterpart configuration incorporated with silver ions. The red dashed circles show the active hydrogen absorption sites both at carbon and nitrogen atoms. The rectangular red dashed circle shows the interface between silver and protein active site.

Absolute Co-ordinates of atoms

Atoms	x	y	z
N	0.07002	0.03774	-0.00423
C	1.53710	0.00253	-0.01319
C	2.06619	-0.02067	1.40813
O	1.29546	-0.07594	2.35443
C	2.06907	-1.21920	-0.80163
C	1.69820	-1.18321	-2.29278
H	1.91002	0.93845	-0.48367
H	1.68521	-2.15604	-0.33974
H	3.17829	-1.23363	-0.73821
H	0.59534	-1.19943	-2.41854
C	2.28861	-2.39022	-3.03250
H	-0.31773	-0.87099	0.34171
H	2.09975	-0.24734	-2.73999
H	1.87066	-3.32527	-2.59856
H	3.39332	-2.39951	-2.90439
C	1.96205	-2.33395	-4.53119
N	2.51497	-3.48879	-5.23416
H	2.38749	-1.40603	-4.97230
H	0.86029	-2.31650	-4.67811
H	2.10846	-4.36805	-4.83957
H	3.55381	-3.51075	-5.11311
H	-0.30516	0.23680	-0.95798
O	3.39461	0.04530	1.64579
H	4.03630	0.12866	0.92674

Optimized value of Energy: 16.4418533 kJ/mol

Supplementary Table 1 The list of atoms and their absolute coordinates of lysine (LYS) unit of seventh helix (gh) obtained from geometry optimization. This coordinates show the stable structure of active sites with minimum energy configuration with value 16.44185 kJ/mol.

Absolute Co-ordinates of atoms			
Atoms	x	y	z
N	0.33089	0.22737	-0.21404
C	1.76498	-0.05226	-0.07112
C	2.14752	-0.12754	1.40102
O	1.81636	0.86104	2.04134
C	2.12247	-1.38004	-0.78107
C	1.86550	-1.34326	-2.29064
H	2.34225	0.77936	-0.53573
H	1.54933	-2.21795	-0.32700
H	3.20726	-1.57887	-0.64054
C	2.30740	-2.64873	-2.95677
H	1.72328	-3.49424	-2.53087
H	3.38745	-2.81850	-2.75839
C	2.08716	-2.58677	-4.47140
N	2.49079	-3.83365	-5.11399
H	1.93700	-4.62917	-4.72088
H	3.50717	-4.00127	-4.93531
O	3.44119	-0.48748	1.57354
H	3.97281	-0.14636	2.30772
Ag	0.97877	-1.74768	2.09750
Ag	2.97529	0.28935	-3.11383
H	0.78341	-1.20033	-2.48649
Ag	-0.05066	0.89067	-2.13112
H	0.06319	1.05257	0.36871
Ag	3.24784	-0.99260	-5.30018
H	1.01451	-2.39717	-4.69373

Optimized value of Energy: 57.2011111 kJ/mol

Supplementary Table 2 The list of atoms and their absolute coordinates of lysine (LYS) unit of seventh helix (gh) incorporated with silver ions obtained from geometry optimization. This coordinates show the stable structure of active sites with minimum energy configuration with value 57.2011 kJ/mol. This shows the excess energy of 40.7593 kJ/mol. indicating that the presence of silver makes the active site to stable at higher energy value.

Absolute Co-ordinates of atoms

Atoms	x	y	z
C	-3.57329	1.45665	1.06811
C	-5.10838	1.25519	1.27565
C	-5.89929	1.77317	0.10477
H	-5.47012	1.81515	2.16963
H	-5.33720	0.18138	1.45821
C	-5.70736	3.27841	0.05349
H	-6.98174	1.55220	0.22710
H	-5.56243	1.29698	-0.83833
C	-4.25579	3.72509	0.08692
C	-4.07303	5.16402	-0.37930
H	-4.42616	5.25338	-1.42801
H	-3.07819	5.62472	-0.33765
H	-4.70942	5.81636	0.25532
C	-3.22736	2.88270	0.50151
C	-1.77986	3.34996	0.34399
C	-0.66245	2.64033	0.61898
H	-1.60500	4.31692	-0.06828
C	0.72609	3.14574	0.35070
H	-0.72490	1.64820	1.00589
C	0.91781	4.50581	-0.30057
H	0.40650	4.51500	-1.28591
H	1.95753	4.81490	-0.48467
H	0.47227	5.28777	0.34955
C	1.75789	2.32157	0.67531
C	3.20386	2.58368	0.41749
C	4.18656	1.67309	0.67803
C	5.64241	2.01331	0.44286
C	6.00730	3.38607	-0.10336
H	5.52242	3.52733	-1.09206
H	7.07707	3.58583	-0.25732
H	5.64630	4.16704	0.59838
C	6.58183	1.06632	0.72855
C	8.07771	1.19074	0.56801
H	6.25389	0.10627	1.11412
H	8.45886	2.14657	0.17899
H	8.42509	0.39489	-0.12303
H	8.55369	1.02688	1.55698
H	-6.23731	3.78119	0.89198
H	-6.18464	3.61210	-0.89395
C	-3.06723	0.36593	0.08454
H	-3.33748	-0.64705	0.45258

H	-1.97633	0.36111	-0.06440
H	-3.50930	0.49546	-0.92344
C	-2.98987	1.22658	2.49233
H	-3.38207	1.99456	3.19366
H	-1.89966	1.25655	2.59647
H	-3.29245	0.22663	2.87428
N	1.03319	-0.42672	-0.63088
C	2.49653	-0.46148	-0.72117
C	3.07321	-0.52157	0.68457
O	2.81194	-1.48822	1.38405
C	2.98675	-1.67831	-1.53963
C	2.46209	-1.68140	-2.98399
H	2.84373	0.45913	-1.23407
H	2.66622	-2.61882	-1.03855
H	4.09904	-1.65770	-1.56501
H	1.35304	-1.73499	-2.98055
C	2.99326	-2.89139	-3.76017
H	0.65715	-1.33979	-0.28494
H	2.77836	-0.74508	-3.49403
H	2.68092	-3.82472	-3.24175
H	4.10417	-2.85335	-3.79060
C	2.44332	-2.90808	-5.19324
N	2.93644	-4.06648	-5.93331
H	2.74920	-1.98124	-5.72553
H	1.33207	-2.94493	-5.16862
H	2.64238	-4.94516	-5.44772
H	3.98090	-4.03604	-5.97826
H	0.60437	-0.19121	-1.55342
O	3.89901	0.43969	1.19245
H	1.50467	1.38390	1.14902
H	3.48713	3.53587	0.01178

Optimized value of Energy: 120.5158133 kJ/mol

Supplementary Table 3 The list of atoms and their absolute coordinates of lysine (LYS) unit of seventh helix (gh) attached with retinal chromophore (rc). This coordinates show the stable structure of active sites with minimum energy configuration with value 120.5158 kJ/mol.

Absolute Co-ordinates of atoms

Atoms	x	y	z
C	-3.77852	1.39111	1.25265
C	-5.32145	1.53171	1.11709
C	-5.70046	2.10463	-0.23219
H	-5.71734	2.22772	1.89188
H	-5.81068	0.54521	1.27538
C	-5.14686	3.51934	-0.33485
H	-6.80334	2.13974	-0.35984
H	-5.29035	1.47459	-1.04587
C	-3.69086	3.64875	0.06518
C	-3.04514	4.96554	-0.31421
H	-3.05619	5.08250	-1.41774
H	-2.01026	5.10485	0.04546
H	-3.63680	5.79354	0.12976
C	-3.03504	2.66204	0.77188
C	-1.57053	2.75868	1.06942
C	-0.68944	2.97726	0.07508
C	0.78791	3.00841	0.27041
C	1.35357	2.98875	1.67289
H	0.88961	3.81352	2.25299
H	2.43593	3.19707	1.72649
H	1.14251	2.01624	2.15989
C	1.56701	3.09701	-0.83719
C	3.05440	3.06069	-0.90815
C	3.89690	2.27737	-0.17992
C	5.37358	2.45204	-0.35877
C	5.83276	3.84817	0.04065
C	6.18111	1.48641	0.14019
C	7.65953	1.43522	-0.11035
H	8.00897	2.22668	-0.79533
H	7.90011	0.46421	-0.58910
H	8.19983	1.50230	0.85650
C	-3.26235	0.19777	0.41167
H	-3.77540	-0.74379	0.70322
H	-2.17085	0.05137	0.56306
H	-3.41892	0.35845	-0.67431
C	-3.50008	1.07857	2.73780
H	-3.78427	1.93689	3.38350
H	-2.43637	0.82828	2.90497
H	-4.07733	0.18571	3.06351
N	0.73664	-0.27755	-0.97340
C	2.19466	-0.14510	-0.95019

C	2.65562	0.25062	0.44476
O	2.34634	-0.46806	1.38314
C	2.84929	-1.48347	-1.33310
C	2.44989	-1.96210	-2.73270
H	2.50510	0.60767	-1.69883
H	1.35737	-2.15810	-2.75865
C	3.18369	-3.25389	-3.10014
H	2.70227	-1.17454	-3.47556
H	2.94117	-4.04187	-2.35343
H	4.28037	-3.07136	-3.08570
C	2.76552	-3.73841	-4.49593
N	3.46038	-4.97087	-4.85718
H	2.99918	-2.95746	-5.25210
H	1.66927	-3.92472	-4.51600
H	3.24004	-5.71701	-4.15803
H	4.49269	-4.80297	-4.86250
O	3.45359	1.33222	0.69758
H	3.49067	3.69626	-1.67232
Ag	-0.91247	2.40835	3.06021
Ag	0.64628	3.36483	-2.71951
Ag	4.74098	0.76883	-3.29171
Ag	6.04387	3.04731	-2.27794
Ag	7.87062	4.37727	-0.36663
H	5.67597	3.99212	1.13215
H	5.24473	4.62643	-0.48827
Ag	5.40251	-0.13167	1.26555
Ag	0.07709	-0.21015	-2.92808
H	0.28721	0.53790	-0.50845
Ag	-5.40115	4.12496	-2.37014
H	-5.73077	4.21778	0.30362
Ag	4.97102	-1.21596	-1.28689
H	2.57751	-2.26760	-0.59115
Ag	-1.49124	3.15282	-1.87739

Optimized value of Energy: 281.7356627 kJ/mol

Supplementary Table 4 The list of atoms and their absolute coordinates of lysine (LYS) unit of seventh helix (gh) attached with retinal chromophore and incorporated with silver ions obtained from geometry optimization. This coordinates show the stable structure of active sites with minimum energy configuration with value 281.7356 kJ/mol. This shows the excess energy of 161.2198 kJ/mol. indicating that the presence of silver makes the active site to stable at higher energy value.

Energy Components (Eh)	chu	chuag	fh	fhag	gh	ghag	ghrc	ghrcag
Kinetic	**353.4**	**577.7**	**572.1**	**675.1**	**317.9**	**572.8**	**418.3**	**591.3**
Pseudopotential (local)	**-3518.3**	**-4766.0**	**-4213.8**	**-5734.6**	**-3101.5**	**-4266.6**	**-4302.3**	**-4651.9**
Pseudopotential (non-local)	**61.9**	**-179.4**	**76.4**	**-281.1**	**58.5**	**-213.1**	**74.2**	**-144.9**
Hartree	**1572.5**	**2220.7**	**2091.7**	**2706.1**	**1379.3**	**1988.7**	**1928.1**	**2141.9**
Exchange-correlation	**-140.9**	**-182.2**	**-157.0**	**-198.3**	**-127.4**	**-172.6**	**-173.6**	**-201.3**
Ewald	**1469.9**	**1562.9**	**19227.9**	**1917.0**	**1031.0**	**1336.1**	**1519.4**	**1486.3**

Supplementary Table 5 The energy components of protein samples obtained after optimization to get stable structure. The geometry optimization is performed following the various energy terms to achieve stable configurations.

Supplementary Methods

Computational details

Linear Scaling density functional theory (LS-DFT)

The theoretical framework is based on linear scaling density functional theory (LS-DFT) and its implementation has been done by using Order N Electronic Total Energy Package (ONETEP) codes [1]. Within ONETEP, the single-particle density matrix, ρ(r, r′), is expressed in a separable form via atom-centred functions (Non-orthogonal Generalized Wannier functions, NWGFs), Φα(r), as:

$$\rho(r, r') = \sum_{\alpha\beta} \phi_\alpha(r) K^{\alpha\beta} \phi^*_\beta(r') \quad (1)$$

Where, $K^{\alpha\beta}$ is the matrix element of the density kernel, which is non-zero only if $|r_\alpha - r_\beta| < r_c$, where r_α and r_β indicate the coordinates of the centres of the NWGFs, $\phi_\alpha$and ϕ_β. r_c is the real-space cutoff length. The truncation of the density Kernel ($K^{\alpha\beta}$), leading to sparse density matrix ρ(r, r'), is justified by the known exponential decay of ρ(r, r') with respect to $|r_\alpha - r_\beta|$ for systems with an electronic band gap, which makes any insulating or semiconducting systems amenable to LS-DFT simulation. The NGWFs are centered on the nuclear coordinates and strictly localized within a sphere of radius (r_α) .Being non-orthogonal, the NGWFs are characterized by a non-diagonal overlap matrix, $S_{\alpha\beta}$

$$S_{\alpha\beta} = \int dr \phi_\alpha(r) \phi_\beta(r) \quad (2)$$

The NGWFs are in turn expanded as a linear combination of coefficients ($C_{m\alpha}$) of localized and orthogonal periodic cardinal sine (*psinc*) functions, $D_m(r)$, given by:

$$\phi_\alpha(r) = \sum_m C_{m\alpha} D_m(r - r_m) \quad (3)$$

with *m* indexing the real-space Cartesian grid inside the spherical localization region of ϕ_α. The *psinc* functions are formed from discrete sum of plane waves, which makes the set of $D_m(r)$ independent from the nuclear coordinates and systematically improvable on increase of

the kinetic energy cutoff. In this approach using ONETEP codes, the total energy of the system under study is minimized self-consistently with respect to $K^{\alpha\beta}$ and $C_{m\alpha}$in two nested loops. The inner loop contains the codes to optimize total energy (E) with respect to the density kernel ($K^{\alpha\beta}$) and outer loop does the same task with respect to the Wannier functions (ϕ_α and ϕ_β). As a result, the NGWFs are optimized in situ by iteratively improving the set of coefficients $(C_{m\alpha})$ that minimize the total energy under the constraints of idempotency of the density matrix (equation (1)) and conservation of the number of electrons in the simulated system. The NGWFs are represented in terms of the underlying orthogonal *psinc* basis, $D_i(r)$, in ONETEP code.

$$\phi_\alpha(r) = \sum_{i \epsilon LR(\alpha)} C_{i\alpha} D_i(r) \tag{4}$$

Where $LR(\alpha)$ is the spherical and atom-centered localization region of NGWF (ϕ_α) and $C_{i\alpha}$ is its expansion coefficients in the *psinc* basis. This approach has been used to lead to near complete convergence of the density functional theory optimization with respect to the basis set even for minimal number of NGWFs employed in the simulations.

The design of catalysts for water electrolysis is based on the first-principles calculation. The hydrogen evolution reaction of the active sites is studied to improve its catalytic ability. In our current work, the linear scaling based ONETEP codes are incorporated with the plane wave scheme, which is based on the density functional theory [2] and the Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [3], is used for our calculations. Supercells with lattices larger than 8 Å are used to investigate the catalytic effect and hydrogen-density dependent HER ability. A 2 × 2 × 1 grid for k point sampling, based on the Monkhorst and Pack scheme [4], for geometry optimization of supercells, and an energy cutoff of 800 eV are consistently used in our calculations. Densities of states (DOSs) are calculated based on a k-point sampling of 3 × 3 × 1. Good convergence is obtained with these parameters, and the total energy was converged to 2.0×10−6 eV/atom.

Ionic forces and Ground state geometry optimization

The forces on ions at R_γ can be calculated by taking the derivative of the total energy with respect to ionic positions following the equation as

$$F_\gamma = -\left(\frac{dE}{dR_\gamma}\right) = -\left(\frac{\partial E}{\partial R_\gamma}\right) - \sum_{\alpha\beta}\left(\frac{\partial E}{\partial K^{\alpha\beta}}\right)\left(\frac{dK^{\alpha\beta}}{dR_\gamma}\right) - \sum_{\alpha}\int\left(\frac{\delta E}{\delta\phi_\alpha(r)}\right)\left(\frac{d\phi_\alpha(r)}{dR_\gamma}\right)d^3r \quad (5)$$

To achieve total electronic energy minimization with respect to the degree of freedom [5] of the density, the following conditions are satisfied during simulation given as

$$\frac{\partial E}{\partial K^{\alpha\beta}} = 0 \quad \text{and} \quad \frac{\partial E}{\partial C_{i\alpha}} = 0 \quad (6)$$

These calculations lie on the Born-Oppenheimer surface for specific ionic calculations. Our calculations converges to Hellmann-Feynman force by vanishing second and third term of equation (5). The components of the total energy (Supplementary Table 5), which will be considered for minimization, consists of ion-ion and ion-electron terms with both local and non-local pseudopotential contributions. This non-local contributions to the energy and ionic forces are calculated in linear scaling $\{O(N)\}$ computational effort. This $\{O(N)\}$ has larger pre-factor of force calculation and consumes less memory compared to its square scaling counterpart $\{O(N^2)\}$. As a result, larger systems like biological systems can be handled smoothly for DFT calculations using this linear scaling methods [6].

Hydrogen Coverage calculation

The hydrogen adsorption energies were calculated using the 2×2×1 surface supercell. The total energy from DFT calculations were obtained as $E^{tot} = [E(NR + nH) - E(NR) - \frac{n}{2}E(H_2)]$, where $E(NR + nH)$, $E(NR)$ and $E(H_2)$ are the total energies for the protein active sites with n hydrogen atoms adsorbed, pristine sites and the molecular hydrogen in the gaseous phase, respectively. We define H coverage as the ratio of the number of H and C atoms in the interfacial plane of the metalloprotein complex. Thermal corrections were done to calculate adsorption free energy. ΔE_{ZPE} and $T\Delta S_H$ are the differences in the zero point energy and entropy contribution between the H adsorbed state and H_2 in the gas phase respectively. The Gibbs free energy of hydrogen absorption was calculated as $\Delta G_H = \Delta E_H + 0.22$ eV.

Supplementary References

1. N. D. M. Hine, et al., *Phys. Rev. B,* 2011, **83**, 195102.
2. P. Hohenberg and W. Kohn, *Phys. Rev.,* 1964, **136**, B864.
3. P. E. Blöch, *Phys. Rev. B,* 1994, **50**, 17953.
4. H. J. Monkhorst and J. Pack, *Phys. Rev. B* 1976, **13**, 5188.
5. C. –K. Skylaris, et al., *Phys. Rev. B* 2002, **66**, 035119.
6. L. Greencard and V. A. Rokhlin, *J. Comput. Phys.,* 1987, **73**, 315.