

Supporting Information: Parallel and low-order scaling implementation of Hartree-Fock exchange using local density fitting

Christoph Köpll and Hans-Joachim Werner*

Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany.

E-mail: werner@theochem.uni-stuttgart.de

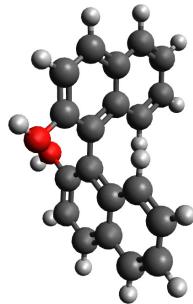
*To whom correspondence should be addressed

1 Optimized geometries

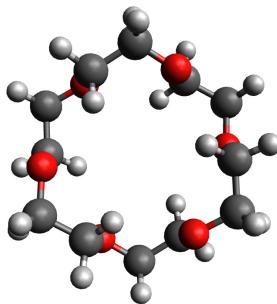
Optimized molecular structures were taken from seven recent publications^{1–7} and used without being further modified. The geometry of water for reaction **I** was optimized using MP2/aug-cc-pVTZ. The molecular structures not shown in the main text are visualized⁸ in the following.

1.1 Structures

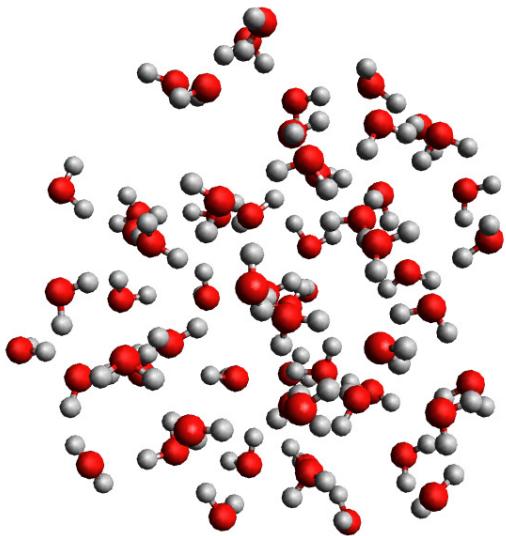
1.1.1 (*S*)-BINOL, C₂₀H₁₄O₂



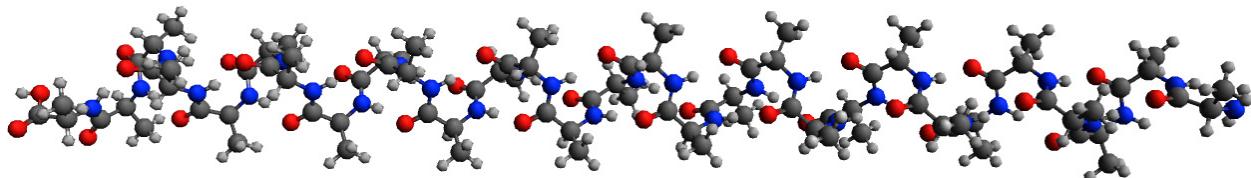
1.1.2 18-crown-6, C₁₂H₂₄O₆



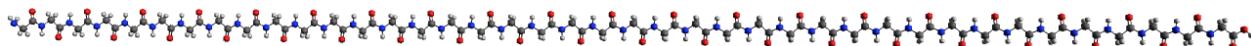
1.1.3 $(\text{H}_2\text{O})_{60}$



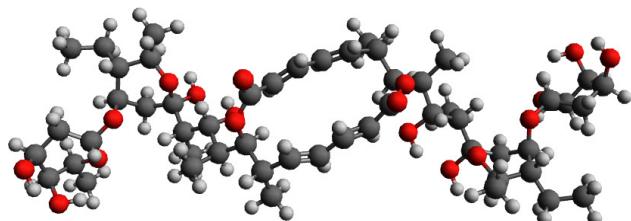
1.1.4 $(\text{Ala})_{29}$, $\text{C}_{87}\text{N}_{29}\text{O}_{30}\text{H}_{147}$



1.1.5 $(\text{Gly})_{40}$, $\text{C}_{80}\text{N}_{40}\text{O}_{41}\text{H}_{122}$



1.1.6 Elaiophylin, $\text{C}_{54}\text{H}_{88}\text{O}_{18}$



2 Errors with orbital relocalization in each iteration

Table 1: Errors (in mE_H) of total LDF-HF energies and PNO-LMP2 correlation energies relative to the energies obtained with non-local DF-HF orbitals. For each heading, the energies in the left column are obtained by localizing the orbitals and adjusting the domains in each HF iteration, while the results in the right columns are obtained by localizing and choosing the domains only once in the beginning. The energies under the heading “last iteration” are obtained using all local approximations, while the “recalculated energies” are computed using the final orbitals without local approximations. The same parameters as in Table 4 of the main text have been used.

Molecule (basis), point group	$dE^{\text{LDF-HF}}$			$dE^{\text{PNO-LMP2}}$			LDF-HF	
	last iteration	recalculated energy	dE_{corr}	dE_{corr}	$dE^{\text{PNO-LMP2}}$	dE_{corr}	iterations	iterations
(S)-BINOL (aug-cc-pVTZ), C ₂	-3.0510	-3.3546	-0.0005	-0.0006	-0.0884	-0.1018	12	12
18-crown-6 (aug-cc-pVTZ), C _i	-0.3364	-0.3683	0.0000	0.0000	-0.0035	-0.0024	8	8
Glabrescol (cc-pVTZ), C ₂	-1.3659	-1.1767	-0.0002	-0.0002	-0.0210	-0.0203	9	9
(R,R)-Jacobsen’s ligand (cc-pVTZ), C ₂	-2.2051	-2.7699	-0.0003	-0.0004	-0.0489	-0.0593	11	11
α -cyclodextrin (aug-cc-pVDZ), C ₂	-2.9351	-3.4036	0.0000	0.0000	-0.0605	-0.0523	10	10
(H ₂ O) ₆₀ (cc-pVTZ), C ₁	1.6139	0.5181	-0.0018	-0.0034	-0.0738	-0.0801	11	10
(Ala) ₂₉ (cc-pVTZ), C ₁	-8.0001	-7.3291	-0.0007	-0.0009	-0.0621	-0.1267	16	11
(Gly) ₄₀ (cc-pVTZ), C ₁	-6.1437	-5.8269	-0.0008	-0.0007	-0.0117	-0.0288	12	11
Nonactin (cc-pVDZ), C ₂	-1.2174	-1.1881	-0.0008	-0.0012	-0.0132	-0.0076	11	11
Nonactin (aug-cc-pVDZ), C ₂	-1.5824	-1.7362	0.0000	0.0000	-0.0310	-0.0284	11	11
Nonactin (cc-pVTZ), C ₂	-1.4650	-1.3865	-0.0003	-0.0003	-0.0247	-0.0212	11	11
Nonactin (aug-cc-pVTZ), C ₂	-1.4570	-1.5989	-0.0001	-0.0001	-0.0237	-0.0205	11	11
Elaiophylin (cc-pVDZ), C ₁	-3.1120	-4.3737	-0.0014	-0.0017	-0.0611	-0.0556	11	10
Elaiophylin (aug-cc-pVDZ), C ₁	-3.7483	-5.3083	-0.0005	-0.0008	-0.0692	-0.0649	11	10
Elaiophylin (cc-pVTZ), C ₁	-3.0409	-4.5013	-0.0006	-0.0008	-0.0586	-0.0556	11	10
Elaiophylin (aug-cc-pVTZ), C ₁	-3.4791	-5.0085	-0.0012	-0.0015	-0.0508	-0.0456	11	10

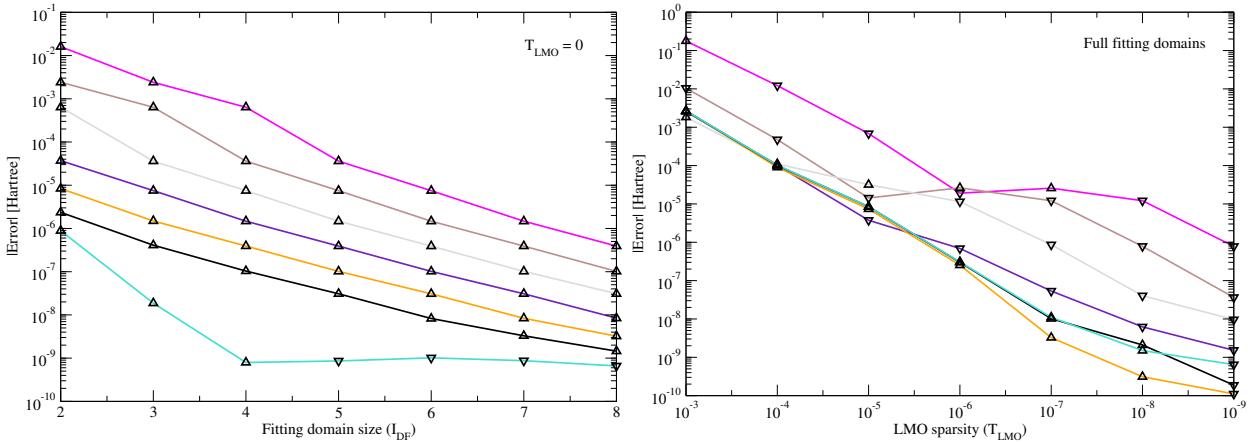
3 Absolute energies and reaction energies

Table 2: Total Hartree-Fock (reaction) energies calculated with our new Hartree-Fock (LDF-HF) implementation, as described in the main text. The PNO-LMP2 correlation energies were calculated as described in Refs. 1 and 3.

Molecule (basis set), point group	Total energies [E _H]	
	LDF-HF	PNO-LMP2 (corr)
(<i>S</i>)-BINOL (aug-cc-pVTZ), C ₂	-915.5597155	-3.6654150
18-crown-6 (aug-cc-pVTZ), C _i	-917.8118186	-3.5306489
Glabrescol (cc-pVTZ), C ₂	-1690.9257086	-6.7953254
(<i>R</i> , <i>R</i>)-Jacobsen's ligand (cc-pVTZ), C ₂	-1653.8380874	-6.9497957
α -cyclodextrin (aug-cc-pVDZ), C ₂	-3644.3548060	-11.1409318
Nonactin (cc-pVDZ), C ₂	-2450.5843477	-7.7478473
Nonactin (aug-cc-pVDZ), C ₂	-2450.7117311	-8.1239801
Nonactin (cc-pVTZ), C ₂	-2451.2435252	-9.5524566
Nonactin (aug-cc-pVTZ), C ₂	-2451.2627601	-9.7267292
Elaeophylin (cc-pVDZ), C ₁	-3443.7655062	-10.8578836
Elaeophylin (aug-cc-pVDZ), C ₁	-3443.9690804	-11.4176254
Elaeophylin (cc-pVTZ), C ₁	-3444.7026349	-13.4139745
Elaeophylin (aug-cc-pVTZ), C ₁	-3444.7339549	-13.6648477
(Gly) ₄₀ (cc-pVTZ), C ₁	-8351.7897628	-30.2787817
(Ala) ₂₉ (cc-pVTZ), C ₁	-7208.3234100	-27.4455815
(H ₂ O) ₆₀ (cc-pVTZ), C ₁	-4563.5254685	-16.0130087
Reaction	Total reaction energies [E _H]	
I	-0.0035467	-0.0088503
II	-0.0105071	0.0016853
III	-0.0233981	-0.0140105
IV	0.0320290	0.0521821

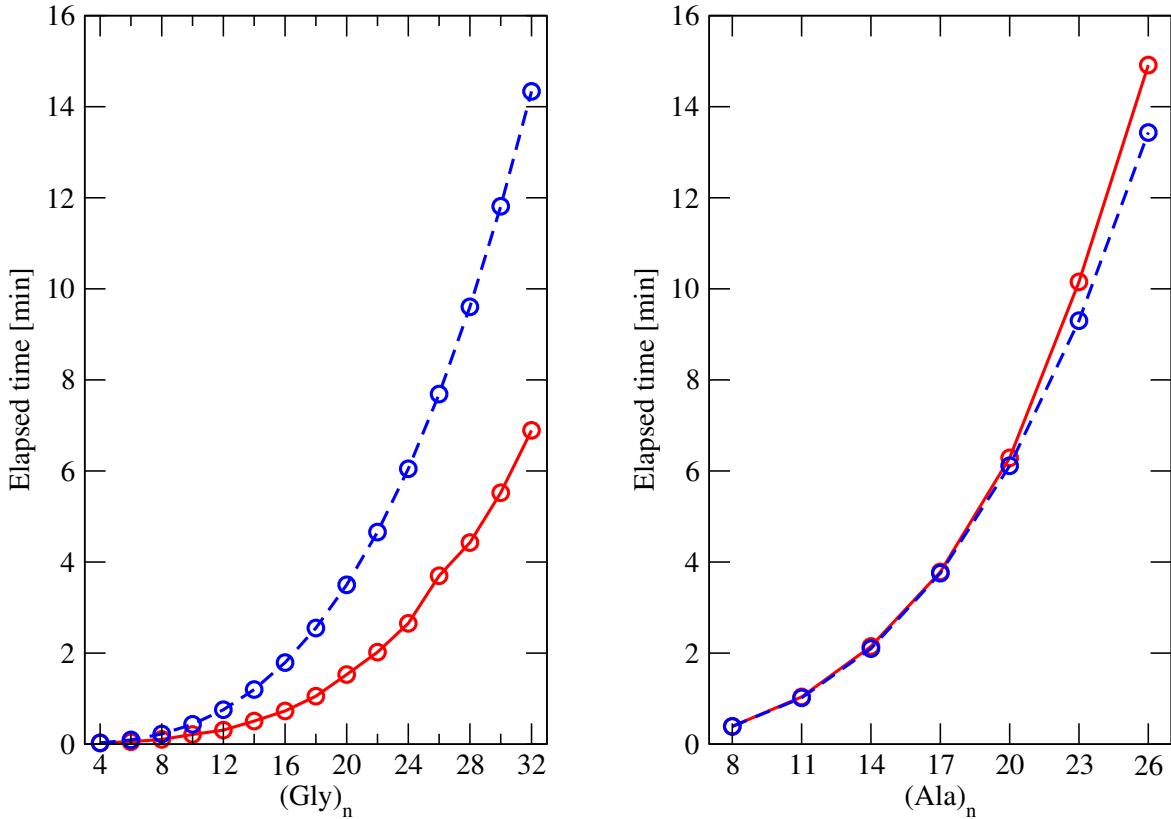
4 Domain thresholds and accuracy

Figure 1: Visualization of the error (E_H) in the total Hartree-Fock energy of the $(\text{Gly})_{16}$ molecule with respect to the canonical DF-HF program in MOLPRO^{9,10} using the VTZ basis set. Several fitting domain sizes (defined through connectivity and distance criteria, see main text) were used (left). The upper-most line (purple) shows the errors if the orbitals are localized and the Fock-matrix is recalculated using the same fitting domains (I_{DF}) as used during the iterations after convergence is reached. The lines below represent the errors if the orbitals are localized and the Fock-matrix is recalculated after convergence is reached, but using larger fitting domain sizes, where I_{DF} is increased by one for each subsequent line. The lowest line (blue) shows the errors obtained when purely the energy is recalculated without using any local approximations. Using the same color coding, the DF-HF errors using LMO sparsity thresholds are shown in the second graph (right). Here, full fitting domains are used and the orbitals are localized, but the sparsity threshold is reduced by a factor of 10 in each step once convergence is reached. The blue line represents the errors obtained when the energy is recalculated without any local approximations once convergence is reached. Triangles indicate if the calculated energy is above (triangle pointing up) or below (triangle pointing down) the corresponding reference energy.



5 Cost and scaling of recalculating the energy

Figure 2: Visualization of the elapsed times to recalculate the energy after the LDF-HF procedure is converged (solid red lines) for glycine chains and alanine helices (left, right) respectively. For all calculations the VTZ basis set and 20 CPU cores were used on a single compute node. To demonstrate the scaling behaviour, the dashed blue lines show a N^3 scaling behaviour. This theoretical N^3 scaling behaviour is based on the elapsed times to recalculate the energy for $(\text{Gly})_4$ (1.68 s) and $(\text{Ala})_8$ (23.48 s). For the largest glycine chains and alanine helices, the recalculation of the energy consumes about 30% of the total elapsed time while for smaller chain lengths the elapsed times for this step are (relatively) smaller compared to the overall elapsed times.



References

- (1) Köppl, C.; Werner, H.-J. On the use of Abelian point group symmetry in density-fitted local MP2 using various types of virtual orbitals. *J. Chem. Phys.* **2015**, *142*, 164108.
- (2) Friedrich, J. Incremental Scheme for Intermolecular Interactions: Benchmarking the Accuracy and the Efficiency. *J. Chem. Theory Comput.* **2012**, *8*, 1597–1607.
- (3) Werner, H.-J.; Knizia, G.; Krause, C.; Schwilke, M.; Dornbach, M. Scalable electron correlation methods I. Linear scaling local MP2-F12 using pair natural orbitals. *J. Chem. Theory Comput.* **2015**, *11*, 484–507.
- (4) Zuend, S. J.; Jacobsen, E. N. Mechanism of Amido-Thiourea Catalyzed Enantioselective Imine Hydrocyanation: Transition State Stabilization via Multiple Non-Covalent Interactions. *J. Am. Chem. Soc.* **2009**, *131*, 15358–15374.
- (5) Adler, T. B.; Werner, H.-J.; Manby, F. R. Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules. *J. Chem. Phys.* **2009**, *130*, 054106.
- (6) Fedorov, A.; Batiste, L.; Couzijn, E. P. A.; Chen, P. Experimental and Theoretical Study of a Gold(I) Aminonitrene Complex in the Gas Phase. *ChemPhysChem* **2010**, *11*, 1002–1005.
- (7) Schwilke, M.; Usyat, D.; Werner, H.-J. Communication: Improved pair approximations in local coupled-cluster methods. *J. Chem. Phys.* **2015**, *142*.
- (8) Hanwell, M.; Curtis, D.; Lonie, D.; Vandermeersch, T.; Zurek, E.; Hutchison, G. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminf.* **2012**, *4*, 17.
- (9) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Konna, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.;

- Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. MOLPRO, development version 2015, a package of ab initio programs. 2015.
- (10) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, 2, 242–253.