DOI: 10.1002/anie.201004022

Twin Arginine Pairs

Quantum-Chemical and Combined Quantum-Chemical/Molecular-Mechanical Studies on the Stabilization of a Twin Arginine Pair in Adenovirus Ad11**

*Chris Vanessa Sumowski, Benedikt B. T. Schmitt, Sabine Schweizer, and Christian Ochsenfeld**

Stacked arginine pairs are an important motif in several essential biological processes. Such pairs play a central role, for example, in the anaerobic respiration of many eubacteria, archaea, and chloroplasts,[1–3] where a highly conserved arginine pair in the corresponding signal peptides determines how efficient respiratory enzymes are transported across the membrane barrier (the so-called twin-arginine transport pathway). In the respiratory process of most aerobic systems, a twin arginine motif is also involved as part of the crucial enzyme cytochrome *c* oxidase.[4,5] Furthermore, studies on mutant domains of the crystallin protein class have shown two arginine pairs to be responsible for a significant part of domain interactions.[6]

Recently, Persson et al.[7] were able to show that such a stacked arginine pair is decisive for binding between a species B adenovirus and its receptor. Species B adenoviruses can cause severe infections in several parts of the human body, for example, in the respiratory and urinary system.[8–10] Most of these viruses use the membrane cofactor protein CD46 as a cellular receptor.[11–14] As Persson et al. demonstrated,[7] the presence of the twin arginine pair Arg279/Arg280 in the virus serotype Ad11 is crucial for the efficient interaction with the receptor CD46. Mutation studies have shown that exchanging Arg279 to glutamine (this mutant is named in the following as R279Q) decreases the binding affinity of Ad11 towards CD46 dramatically. At first sight, these observations are surprising, as Arg279 does not contact CD46 directly but is only located below the Arg280 residue interacting with CD46. These observations were supported by our quantum-chemical results that showed a considerable decrease of the interaction energy between virus and receptor upon elimination of Arg279.[7]

From a physical point of view, it may seem surprising how two positively charged arginines could form a stable interacting pair, as electrostatic repulsion is expected to overcompensate possible stabilizing effects within the arginine pair (for example, cation–π and π–π interactions). For this reason, influences by the environment and the protonation states of the arginines need to be studied, because they may lead to a stabilization of the twin arginines. In this context computational methods are expected to be particularly well-suited for studying the possible stabilizing factors of such stacked arginine pairs, as they allow to examine the influence of specific parts of the surrounding structure in detail. Therefore, we present both purely quantum-chemical (QM) and combined quantum-chemical/molecular-mechanical (QM/MM) calculations to investigate the protonation state of the twin arginine pair within the Ad11/CD46 complex, which was shown to be a decisive element upon the binding process.[7] Whilst such stabilizing factors have already been discussed for simpler model systems using quantum-chemical methods and also for related (but different) systems using force field models,[15–27] we employ linear-scaling quantum-chemical methods that allow QM regions with more than 1000 atoms to be described.[28–32] In this way, the influence of the environment and its stabilizing factors can be accounted for in a more reliable manner and systematic convergence studies with respect to the size of the QM region up to more than 1000 atoms are presented. Furthermore, the impact of the local environment upon the central amino acid pair 279/280 is compared for the R279Q mutant and the wild-type Ad11.

As a structural basis, a protonated X-ray structure of the Ad11 knob-D2 complex was employed (provided by David Persson and Thilo Stehle (University of Tübingen); for structural details and protonation, see Ref. [7]). Here, all of the titratable groups were treated as charged residues. Starting from this structure containing zwitterionic salt bridges for Arg279/Glu250 and Arg280/Glu63, all possible protonation states for these pairs were studied (Figure 1a, including the notation of isomers). The notation for the pivotal amino acids used throughout this work can be found in Figure 2.

In QM and QM/MM calculations, the QM part was treated at the HF (Hartree–Fock)[33–35] and RI-MP2 (second-order many-body perturbation theory using the resolution of identity approach)[36–39] levels using SVP[40] and TZVPP[40] basis sets. Although the HF/SVP approach has well-known deficiencies, the energetic order of the various isomers for the

[*] Dr. C. V. Sumowski, B. B. T. Schmitt, Dr. S. Schweizer, Prof. Dr. C. Ochsenfeld
Chair of Theoretical Chemistry, Department of Chemistry
University of Munich (LMU)
Butenandtstrasse 7, 81377 Munich (Germany)
E-mail: christian.ochsenfeld@cup.uni-muenchen.de
Homepage: http://www.cup.uni-muenchen.de/pc/ochsenfeld

[**] We thank Prof. Dr. Thilo Stehle and Dr. David Persson (Universität Tübingen) for useful discussions and providing the X-ray-based structure. C.V.S. thanks the Fonds der Chemischen Industrie, B.B.T.S. the Cusanuswerk—Bischöfliche Studienförderung, and S.S. the Studienstiftung des deutschen Volkes for graduate fellowships. C.O. acknowledges financial support by the Volkswagen Stiftung within the funding initiative "New Conceptual Approaches to Modeling and Simulation of Complex Systems".

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201004022.

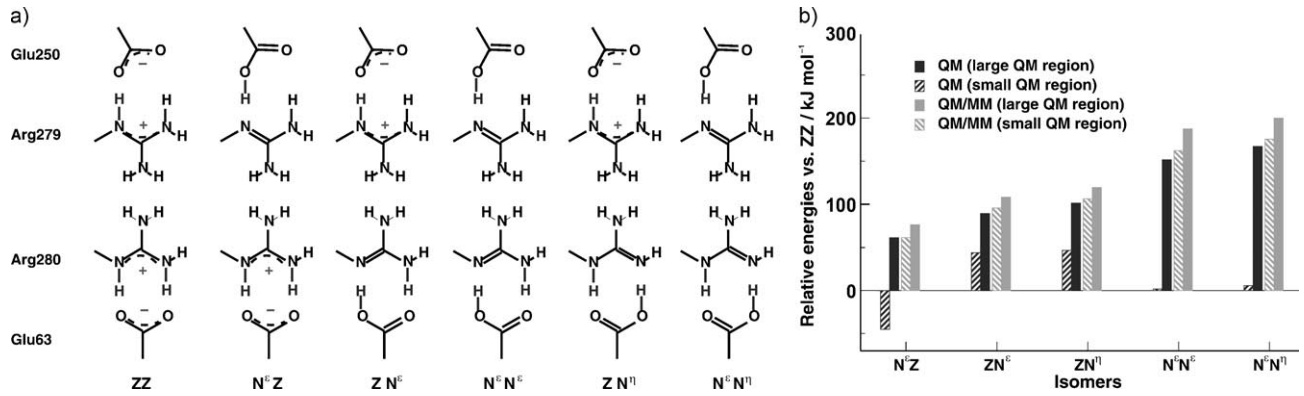

Figure 1. a) Possible isomers for Arg279/Glu250 and Arg280/Glu63 pairs. b) Relative energies between isomer **ZZ** and all other possible isomers. The calculations were performed for two QM region sizes (small region comprises 84 atoms, large region 1035 atoms). The QM part was calculated at the HF/SVP level. Positive values indicate the stabilization of **ZZ** compared to the respective isomer.

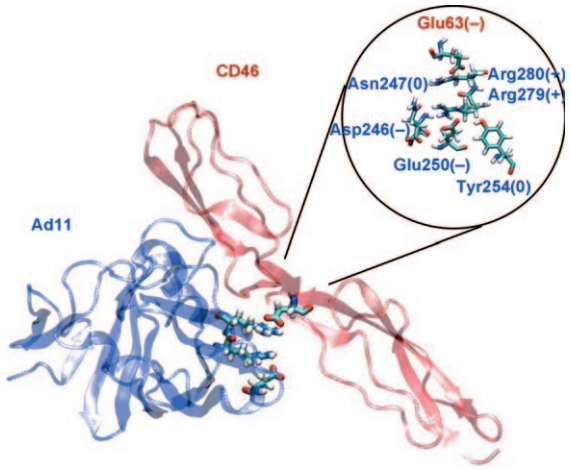

Figure 2. The Ad11/CD46 complex with the arginine pair (crucial for effective binding) and its stabilizing microenvironment.

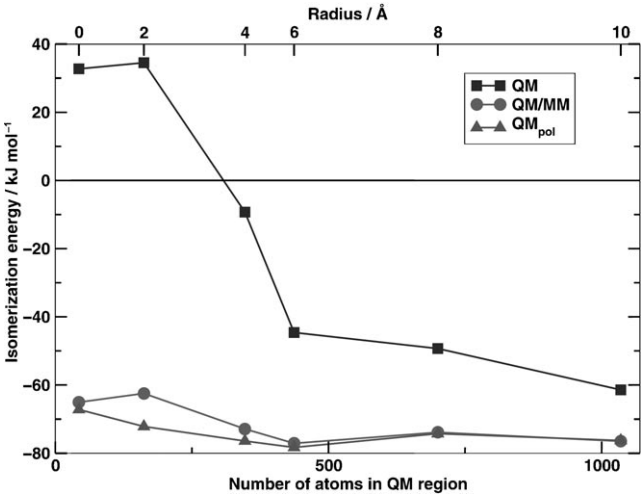

Figure 3. Dependence of the isomerization energy between **ZZ** and **N$^{\varepsilon}$Z** upon increase of the QM region size. The QM part was calculated at the HF/SVP level. The negative sign indicates **ZZ** as the energetically preferred isomer.

present system is correctly described (for a comparison to MP2 data, see the Supporting Information), which shows the usefulness of the approach for large QM fragments. Nevertheless it would be desirable to study also the larger QM fragments using our linear-scaling MP2 methodology[41] in the future. For further computational details, see the Supporting Information.

To study the possible protonation states (Figure 1 a) of the central salt bridges Glu250/Arg279 and Arg280/Glu63 in the Ad11/CD46 complex, the isomerization energies between the states were calculated. As the environment is expected to have an important impact on the protonation state of the arginine pair, the relevant influences have to be determined and a suitable QM region has to be chosen to allow for a reliable calculation of the isomerization energies.

Therefore, as a first example, we studied the dependence of the isomerization energy upon a systematic increase of the calculated QM region size for the two isomers **ZZ** and **N$^{\varepsilon}$Z** (Figure 1 a). The employed QM regions were chosen as spherical cutouts with a defined radius around residues Glu250 and Arg279, which are the proton exchanging units in this case. The notation of the subsystems corresponds to the radius of the spheres included in the QM region. A detailed description of the computational procedure is provided in the Supporting Information.

The QM and QM/MM convergence behavior (Figure 3) shows significant differences: The pure QM results strongly depend on the chosen QM region varying over a range of 95 kJ mol^{-1} (HF/SVP) upon enlarging the QM region. In contrast, the QM/MM data are considerably less influenced by the choice of the QM region with a maximum variation of 15 kJ mol^{-1} between the isomerization energies calculated for the different QM regions. By including all residues in a 6 Å radius around the Glu250/Arg279 pair (437 atoms in QM region), convergence within 3 kJ mol^{-1} compared to the largest QM region is obtained within the QM/MM scheme. A comparison of the QM and QM/MM results for the largest QM region shows a remaining difference between the approaches of 15 kJ mol^{-1}, even though the QM region comprises more than 1000 atoms. This remaining discrepancy most likely originates from long ranging electrostatic effects.

The data clearly show that nearby residues (less than 6 Å) entail a larger influence (90 kJ mol^{-1}) on the isomerization energy than electrostatic effects originating from distant groups (15 kJ mol^{-1}, more than 10 Å). Therefore, residues within a sphere of 6 Å (437 atoms) around the central pair have to be included into the QM region to obtain reliable isomerization energies. Although the faster convergence shows the usefulness of the QM/MM scheme, large QM regions are still needed for a reliable description of the system.

It should also be noted that pure classical force field models are not suitable for studying the active center in the present case, as the breaking and formation of bonds leads to different topologies and therefore different energy expressions for the respective isomers (which cannot be compared consistently).

In Figure 1, the isomerization energies relative to the zwitterionic isomer **ZZ** are shown for all possible isomers calculated using QM and QM/MM schemes with two different QM region sizes. The size-converged results within the two schemes (using the 10 Å QM region of the systematic convergence study as a reference) consistently reveal the zwitterionic state **ZZ** as the most stable isomer. These results, combined with the findings of the systematic convergence study, demonstrate the importance of the environment and, in particular, of nearby residues in stabilizing the zwitterionic state **ZZ**, in which the two positively charged arginines stay in close contact and no proton transfer occurs. This finding is remarkable, as from a structural point of view, a hydrogen transfer between the positively charged arginines and their negatively charged counterparts would have seemed quite convenient (see orientation of the respective aminoacids in Figure 2).

In this context, it is crucial to gain a more detailed understanding of the specific effects that stabilize the zwitterionic arrangement in **ZZ**. Therefore, the influence of individual amino acids upon the isomerization energy between the two most stable isomers **ZZ** and **N$^{\varepsilon}$Z** was studied. Starting from a cutout comprising only residues Glu250, Arg279, and Arg280, the surrounding amino acids were included in a stepwise fashion in the respective subsystems of **ZZ** and **N$^{\varepsilon}$Z** (denoted as I to V according to Table 1). The calculations were carried out at the HF/SVP and RI-MP2/TZVPP level.

Table 1: Specific influences of surrounding residues upon the QM isomerization energy between **ZZ** and **N$^{\varepsilon}$Z**.[a]

Residue numbers		$\Delta E_{ZZ-N^{\varepsilon}Z}$ [kJ mol^{-1}] RI-MP2/TZVPP	HF/SVP
I	$250^{-}/279^{+}/280^{+}$ (**ZZ**) $250^{0}/279^{0}/280^{+}$ (**N$^{\varepsilon}$Z**)	85.2	83.8
II	$250/279/280/63^{-}$	45.8	45.1
III	$250/279/280/63/254^{0}$	5.7	1.7
IV	$250/279/280/63/254/246^{-}$	−5.7	−8.7
V	$250/279/280/63/254/246/247^{0}$	−20.6	−23.6
overall change		−105.8	−107.4

[a] The superscripts indicate the charges of the newly comprised residues and the central units. Negative signs for $\Delta E_{ZZ-N^{\varepsilon}Z}$ indicate **ZZ** as the energetically preferred isomer.

The decisive residues stabilizing **ZZ** in the molecular surroundings are revealed by analyzing the **ZZ**–**N$^{\varepsilon}$Z** isomerization energy (RI-MP2/TZVPP) with regard to the subsystem size (see Table 1). The results show that an elaborated interplay of the twin arginines with their microenvironment is responsible for the stabilization of the charged arginine pair in **ZZ**: First, for the smallest subsystem (Glu250, Arg279, and Arg280 without surroundings), a preference for the neutral situation (**N$^{\varepsilon}$Z**) is predicted. However, the evaluation of larger fragments shows a stepwise stabilization of **ZZ** (and with it the doubly charged arginine pair) by favorable electrostatic interactions (Glu63 and Asp246) and beneficial hydrogen bonding in **ZZ** (Tyr254 and Asn247; see Figure 4 a,b and Figure 5 a,b). In particular for Tyr254, this effect is remarkable because this residue is not in direct contact to the

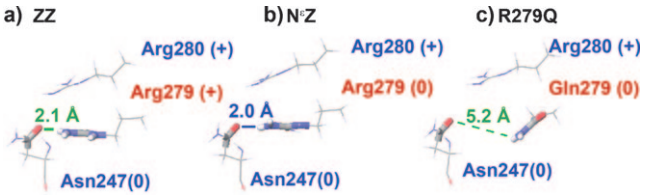

Figure 4. Possible hydrogen bonds between Asn247 and residue 279. The corresponding distances are also given. The proton positions were reoptimized for each structure.

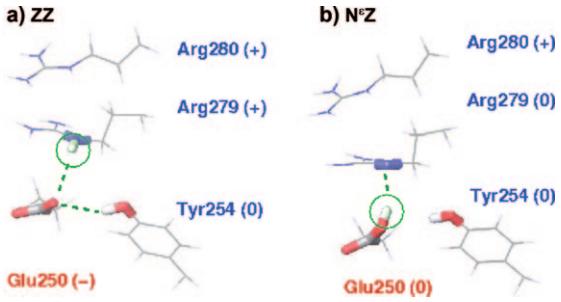

Figure 5. Possible hydrogen bonds between Tyr254 and Glu250. The structure of **R279Q** is similar to **ZZ**.

arginine pair but stabilizes instead the unprotonated form of Glu250. In this way, an overall stabilization of **ZZ** by 106 kJ mol^{-1} (RI-MP2/TZVPP, corresponding value at HF/SVP: 107 kJ mol^{-1}) is observed. This result shows that the few residues chosen in our model already cover a substantial part of the total stabilizing energy determined within the QM/MM convergence studies (Figure 3).

As mentioned above, mutation studies in the work of Persson et al.[7] have shown the exchange of Arg279 by a glutamine residue (Gln279) to reduce the binding affinity of Ad11 toward CD46 drastically. To study to what extent the R279Q mutant profits from the stabilizing network discussed above for the wild-type Ad11 (**ZZ**), the influence of surrounding residues on the stability of the amino acid pair 279/280 was also studied for the mutant.

To obtain a comparable structure for the **R279Q** mutant, residue Arg279 was substituted in isomer **ZZ** by a glutamine residue (for further details, see the Supporting Information). Starting from a small subsystem containing only residues 279 and 280, the other QM subsystems I to V were considered as described above (Table 1). The calculations were performed at the HF/SVP and RI-MP2/TZVPP level. As absolute energies of **ZZ** and **R279Q** cannot be compared directly owing to differing numbers of atoms, the change in the interaction energies ($\Delta\Delta E^{\text{interact}}_{P/P'}$) between the central pairs Arg279/Arg280 (P) and Gln279/Arg280 (P′) and their environment is analyzed upon inclusion of individual amino acids for both **ZZ** and **R279Q**. A detailed description of this procedure can be found in the Supporting Information.

The values in Table 2 show the overall larger stabilization of the wild-type **ZZ** in comparison to the mutant structure **R279Q** by inclusion of the negatively charged residues Glu250, Glu63, and Asp246. However, addition of the neutral residues Tyr254 and Asn247 also stabilizes **ZZ** better than **R279Q**. Again, the different ability of the central pairs to form stabilizing hydrogen bonds entails the better stabilization of Arg279/Arg280 compared to Gln279/Arg280 (Figure 4a,c and Figure 5a) so that the mutant is not able to benefit from the stabilizing hydrogen bond network by the surrounding residues.

Table 2: Change in the interaction energies between the central pair 279/280 and its environment upon increase of the calculated subsystem.[a]

	Added residue	$\Delta\Delta E^{\text{interact}}_{P/P'}$ RI-MP2/TZVPP	$\Delta\Delta E^{\text{interact}}_{P/P'}$ HF/SVP
CPS–I	250^{-}	−269.0	−258.4
I–II	63^{-}	−223.7	−226.7
II–III	254^{0}	−2.6	−4.9
III–IV	246^{-}	−183.1	−176.5
IV–V	247^{0}	−55.2	−50.9
sum $\Delta\Delta E$		−733.6	−717.4

[a] CPS = central pair system. The results for **R279Q** are referred to the zwitterionic state **ZZ** and indicate which structure is energetically preferred by the environmental interactions. For a detailed explanation, see text and Supporting Information. The results are given in kJ mol^{-1}. Negative values for $\Delta\Delta E^{\text{interact}}_{P/P'}$ indicate that enlarging the environment, for example from CPS to I, leads to interaction energies that are more in favor for Arg279/Arg280 (P) rather than for Gln279/Arg280 (P′). Along with our most reliable RI-MP2/TZVPP results, HF/SVP data are also listed for comparison.

In conclusion, we were able to elucidate by QM and QM/MM calculations the contributions of specific amino acids in the microenvironment of a twin arginine pair to the stabilization of this important structural motif in an Ad11/CD46 complex. In particular, electrostatic effects and a complex network of hydrogen bonds lead to a significant stabilization of the charged arrangement of the two arginines and their respective counterparts. This conclusion is noteworthy, as it might otherwise be assumed that the reduction of the electrostatic repulsion between the two positively charged arginines by a proton transfer would have seemed as a suitable option.

Furthermore, by comparing the influences of different amino acids on interaction energies within the zwitterionic wild-type and the mutant R279Q complex, it was shown that the mutant complex is energetically disfavored owing to electrostatic effects and cannot benefit from the network of hydrogen bonds provided by the surrounding residues.

As the twin arginine motif is crucial in some essential biomolecular processes our results provide insights into the stabilizing factors of such arrangements and the respective virus–receptor interactions. The analysis of stabilizing factors has been possible by systematic QM cluster calculations for studying local effects. Furthermore, the usefulness of the QM/MM approach for describing complex biomolecular systems is illustrated: The convergence with the QM-region size is clearly faster within QM/MM than within pure QM approaches, although even within QM/MM schemes large QM regions are required for a reliable description.

Received: July 1, 2010
Revised: August 11, 2010
Published online: November 18, 2010

Keywords: ab initio calculations · hydrogen bonds · protonation · twin arginine pairs · virus–cell interactions

[1] F. Sargent, *Microbiology* **2007**, *153*, 633.
[2] N. R. Stanley, T. Palmer, B. C. Berks, *J. Biol. Chem.* **2000**, *275*, 11591.
[3] K. Hatzixanthis, T. A. Clarke, A. Oubrie, D. J. Richardson, R. J. Turner, F. Sargent, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8460.
[4] J. Xu, G. A. Voth, *Biochim. Biophys. Acta Bioenerg.* **2008**, *1777*, 196.
[5] J. Qian, D. A. Mills, L. Geren, K. Wang, C. W. Hoganson, B. Schmidt, C. Hiser, G. T. Babcock, B. Durham, F. Millett, S. Ferguson-Miller, *Biochemistry* **2004**, *43*, 5748.
[6] B. V. Norledge, E.- M. Mayr, R. Glockshuber, O. A. Bateman, C. Slingsby, R. Jaenicke, H. P. C. Driessen, *Nat. Struct. Biol.* **1996**, *3*, 267.
[7] B. D. Persson, S. Müller, D. M. Reiter, B. B. T. Schmitt, M. Marttila, C. V. Sumowski, S. Schweizer, U. Scheu, C. Ochsenfeld, N. Arnberg, T. Stehle, *J. Virol.* **2009**, *83*, 673.
[8] H. Jalal, D. F. Bibby, J. W. Tang, J. Bennett, C. Kyriakou, K. Peggs, D. Cubitt, N. S. Brink, K. N. Ward, R. S. Tedder, *J. Clin. Microbiol.* **2005**, *43*, 2575.
[9] S. A. Lukashok, M. S. Horwitz, *Curr. Clin. Top. Infect. Dis.* **1998**, *18*, 286.
[10] R. Y. Moon, *Pediatrics Rev.* **1999**, *20*, 56.
[11] A. Gaggar, D. M. Shayakhmetov, A. Lieber, *Nat. Med.* **2003**, *9*, 1408.
[12] M. Marttila, D. Persson, D. Gustafsson, M. K. Liszewski, J. P. Atkinson, G. Wadell, N. Arnberg, *J. Virol.* **2005**, *79*, 14429.
[13] A. Segerman, J. P. Atkinson, M. Marttila, V. Dennerquist, G. Wadell, N. Arnberg, *J. Virol.* **2003**, *77*, 9183.
[14] Y. Zhang, J. M. Bergelson, *J. Virol.* **2005**, *79*, 12125.
[15] D. Pednekar, A. Tendulkar, S. Durani, *Proteins Struct. Funct. Bioinf.* **2009**, *74*, 155.
[16] J.-C. Soetens, C. Millot, C. Chipot, G. Jansen, J. G. Ángyán, B. Maigret, *J. Phys. Chem. B* **1997**, *101*, 10910.
[17] S. Boudon, G. Wipff, B. Maigret, *J. Phys. Chem.* **1990**, *94*, 6056.
[18] A. L. Magalhães, B. Maigret, J. Hoflack, J. A. N. F. Gomes, H. A. Scheraga, *J. Protein Chem.* **1994**, *13*, 195.
[19] P. E. Mason, C. E. Dempsey, G. W. Neilson, J. W. Brady, *J. Phys. Chem. B* **2005**, *109*, 24185.

[20] A. Masunov, T. Lazaridis, *J. Am. Chem. Soc.* **2003**, *125*, 1722.
[21] J. Singh, J. M. Thornton, *J. Mol. Biol.* **1990**, *211*, 595.
[22] C. L. Nandi, J. Singh, J. M. Thornton, *Protein Eng.* **1993**, *6*, 247.
[23] V. Král, P. Mader, R. Collard, M. Fábry, M. Hořejší, P. Řezáčová, M. Kožíšek, J. Závada, J. Sedláček, L. Rulíšek, J. Brynda, *Proteins Struct. Funct. Bioinf.* **2008**, *71*, 1275.
[24] M. J. Ramos, A. Melo, E. S. Henriques, J. A. N. F. Gomes, N. Reuter, B. Maigret, W. B. Floriano, M. A. C. Nascimento, *Int. J. Quantum Chem.* **1999**, *74*, 299.
[25] A. Melo, M. Ramos, W. B. Floriano, J. Gomes, J. Leão, A. Magalhães, B. Maigret, M. C. Nascimento, N. Reuter, *J. Mol. Struct.* **1999**, *463*, 8190.
[26] Y.-J. Zheng, R. L. Ornstein, *J. Am. Chem. Soc.* **1996**, *118*, 11237.
[27] X. Barril, C. Alemán, M. Orozco, F. Luque, *Proteins Struct. Funct. Genet.* **1998**, *32*, 67.
[28] C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon, *Chem. Phys. Lett.* **1994**, *230*, 8.
[29] C. Ochsenfeld, C. A. White, M. Head-Gordon, *J. Chem. Phys.* **1998**, *109*, 1663.
[30] C. Ochsenfeld, *Chem. Phys. Lett.* **2000**, *327*, 216.
[31] "Linear-Scaling Methods in Quantum Chemistry": C. Ochsenfeld, J. Kussmann, D. S. Lambrecht, *Reviews in Computational Chemistry*, *Vol. 23*, VCH Publishers, New York, **2007**, pp. 1 – 82.
[32] C. V. Sumowski, C. Ochsenfeld, *J. Phys. Chem. A* **2009**, *113*, 11734.
[33] D. R. Hartree, *Proc. Cambridge Philos. Soc.* **1928**, *24*, 89.
[34] V. Fock, *Z. Phys.* **1930**, *61*, 126.
[35] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69.
[36] J. L. Whitten, *J. Chem. Phys.* **1973**, *58*, 4496.
[37] M. Feyereisen, G. Fitzgerald, A. Komornicki, *Chem. Phys. Lett.* **1993**, *208*, 359.
[38] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143.
[39] R. A. Kendall, H. A. Früchtl, *Theor. Chem. Acc.* **1997**, *97*, 158.
[40] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571.
[41] B. Doser, D. S. Lambrecht, J. Kussmann, C. Ochsenfeld, *J. Chem. Phys.* **2009**, *130*, 064107.

