Supporting Information

Computational Insights into Aluminum and Cation Placement in Clinoptilolite: Optimizing DFT Parameters for Natural Zeolites, Part I

Lobna Saeed[1,2], Michael Fischer*[1,2]

[1]Crystallography & Geomaterials Research, Faculty of Geosciences, University of Bremen, Klagenfurter Straße 2-4, 28359 Bremen, Germany

[2]Bremen Center for Computational Materials Science and MAPEX Center for Materials
and Processes, University of Bremen, 28359 Bremen, Germany

Email: lobna@uni-bremen.de, michael.fischer@uni-bremen.de

*Corresponding author (E-mail: michael.fischer@uni-bremen.de)

Index

Optimization of the basis sets S1

Example CP2K input parameters S2

Discussion on bond lengths S3

Least stable configuration for Na-clinoptilolite S4

References

S1:Optimization of the basis sets:

The basis sets optimizations were carried out using the same module 'Quickstep' within CP2K. The Gaussian-type orbitals (GTO) are better suited for analyzing systems characterized by sparse packing, covalently bonded, and have significant variations in the electronic densities as found in zeolites [1]. However, the r^2 in the exponential of the GTOs renders them less effective than the Slater-type orbitals (STO) for example, and thus more GTOs are necessary for achieving a given level of accuracy; needless to say, this is remedied by the simplicity with which the necessary integrals are computed. A linear combination of those orbitals or a 'contraction' of those orbitals can improve their accuracy since a single GTO is not enough to describe nodal characteristics and the spatial extent of atomic orbitals. The Gaussian orbitals utilized here are therefore highly contracted, and they meet the criteria of being small yet highly accurate; additionally, they have their angular momentum states sharing the same exponents. Nevertheless, they are incomplete. The first and foremost important step in the optimization of MOLOPT basis sets is finding suitable training molecules, which can be found here [2, 3]. In addition, based on the energies of the training molecules and the weight of the condition number of the overlap matrix, the contraction coefficients were optimized, after the addition of one extra *d* orbital function, or one could say more contraction coefficients for the *d* orbital, which was followed by the optimization of the exponents. In other words, the main aim of the optimizations of the basis sets is to minimize the objective function ($\Omega(\{\alpha_i, c_i\})$) [4] which is essentially a linear combination of two components: the total energy (E_{tot}) of a set of molecules, denoted as M below, and the logarithm of the condition number (κ) of the overlap matrix for the respective set of molecules M with B as the basis sets to be optimized:

$$\Omega(\{\alpha_i, c_i\}) = \sum_B \sum_M (E_{tot}^{B,M}(\{\alpha_i, c_i\}) + \gamma ln\kappa^{B,M}(\{\alpha_i, c_i\})) \quad (1)$$

The addition of contraction coefficients of an *f* orbital function increased the total energy of the system which was otherwise decreasing when contraction coefficients of a *d* orbital function or two *f* orbital functions were added. Adding to this, the latter showed significant improvement in the SCF cycle behavior in terms of fewer SCF steps to achieve convergence and other factors; however, a more prolonged time for convergence was observed. The significant impact of the two *f* orbitals on the movement of Ca cations is noteworthy. Specifically, in the most stable configuration, Ca was 5-coordinated and migrated to the uppermost section of channel B rather than 3-coordinated when *d* orbitals were utilized. On the same note, the bond lengths were more overestimated with the 2*f* orbitals.

All the contraction coefficients or the angular momentum share the same Gaussian exponents. The optimization of both the training molecules and the MOLOPT basis sets, based on GTH pseudopotentials, were carried out employing the PBE functional, which is recommended as the most suitable option for such purposes, along with uncontracted basis sets for each element of the respective molecule and the GPW approach. Those uncontracted basis sets are available with the distribution of CP2K. Furthermore, optimizations of the wave functions were carried out using the OT method. One should carry on unrestricted KS calculations in the optimizations of some of the isolated molecules along with specifying the multiplicity to avoid the artificial mixing of spins. Optimizing basis sets and finding a solution for a given system is an art that is devoid of a specific formula or any special recipe; however, there can be general guidelines and procedures. First, a suitable set of training molecules with uncontracted basis sets or very large reference basis sets from which wave functions will be calculated and utilized for further optimizing the basis sets should be set up. Second: contraction coefficients should be optimized along with added new orbitals and/or polarization functions followed by optimization of the exponents as well. Third: testing the optimized basis sets with the respective system, if not suitable the whole procedure will be repeated. It is worth noting that in all the above-mentioned calculations, not only the valence electrons but also the semi-core electrons were included. All the optimized basis sets for Ca are shown below. One can also find valuable information here [5, 6].

Ca with extra d-orbital
Ca DZVP-MOLOPT-PBE-GTH-Q10
2 0 2 6 3 2 2

Exponents	*s*-orbitals		
4.60969561119830	-4.16779140132101E-01	2.02721176838159E-01	-2.46797676278451E-01
4.07522365691220	6.44401618836558E-01	-2.67914636629039E-01	2.92105638082871E-01
1.22341351595228	-2.75632783290642E-01	4.77036941035633E-02	2.26816684078908E-02
0.48657501357383	-5.36401324892427E-01	-1.49652678636945E-01	5.13253972784606E-01
0.17461145783866	1.05069982225753E-01	1.45011866797011E-01	-6.92577026176508E-01
0.04707158534723	1.90531331428070E-01	9.17292681631653E-01	3.31900839612247E-01

p-orbitals	
2.43431267203838E-01	8.06459501148548E-02
-1.82671844869678E-01	-9.32197046402658E-02
-6.42389254349290E-01	-1.31505967482886E-01
-6.37002479694419E-01	-1.92140795917026E-01
-2.20485178419722E-01	2.57958876352657E-01
-2.00804352782480E-01	9.29543775498761E-01

d-orbitals	
-3.97963023464865E-01	8.44733292117897E-01
5.83758398884686E-01	1.00766203661576E-01
3.30684092006633E-01	2.64471926186627E-01
4.14642962512910E-01	8.50604408788367E-02
4.33524739883338E-01	1.91303429813635E-01
1.77840120691707E-01	4.03105639734135E-01

Ca with extra f-orbital
Ca DZVP-MOLOPT-PBE-GTH-Q10
2 0 3 6 3 2 1 1

Exponents	*s*-orbitals		
4.50798328825480	-4.35073323833690E-01	2.15245730662016E-01	-2.92700059715182E-01
4.09136741100836	6.28737920242540E-01	-2.65293575064148E-01	3.30593789296264E-01
1.20722427602211	-2.58726010297236E-01	3.29592413809241E-02	4.65368204647540E-02
0.47552507378954	-5.30691924130877E-01	-1.54764368746541E-01	4.82923949115056E-01
0.16618372233072	1.87490274229611E-01	1.25230825958132E-01	-6.75772776627323E-01
0.04670560826516	1.77972359833315E-01	9.17914763616634E-01	3.36131461786926E-01

p-orbitals	
3.11569326587742E-01	8.72156595597027E-02
-2.56801196871591E-01	-1.01281034982567E-01
-6.26031251904546E-01	-1.14881049264135E-01
-6.09938903850515E-01	-1.64822238316717E-01
-2.01429432246825E-01	2.92467533191001E-01
-1.80175717143462E-01	9.25329291109838E-01

d-orbitals
-4.81098478403694E-01
6.01171652053338E-01
2.88488639039091E-01
3.66936978814153E-01
4.01345395465083E-01
1.67899765062976E-01

f-orbitals
1.33240665751123E-01
-1.18539912727538E-01
2.05212010100707E-01
4.09225488517396E-01
5.51132258840644E-01
6.74441233766205E-01

Ca with extra f-orbitals
Ca DZVP-MOLOPT-PBE-GTH-Q10
2 0 3 6 3 2 1 2

Exponents	*s*-orbitals		
4.50802022976258	-4.35050696626062E-01	2.15305495252464E-01	-2.92691259197007E-01
4.09137508799607	6.28696466272250E-01	-2.65334621848619E-01	3.30662943499679E-01
1.20719488070510	-2.58786482461149E-01	3.29450513590746E-02	4.65382963386126E-02
0.47540438695886	-5.30701308748047E-01	-1.54783511424308E-01	4.82926931335503E-01
0.16600342022583	1.87551986900360E-01	1.25236299895012E-01	-6.75706332912976E-01
0.04669984446780	1.77993185993697E-01	9.17885417431423E-01	3.36200181754666E-01

p-orbitals	
3.11534751907668E-01	8.72684515519671E-02
-2.56781160996079E-01	-1.01288487967710E-01
-6.26019345111863E-01	-1.14791975423779E-01
-6.10029764409852E-01	-1.64825727689873E-01
-2.01259844630328E-01	2.92545878632918E-01
-1.80187332325232E-01	9.25309164730168E-01

d-orbitals
-4.81133063808088E-01
6.01147604150075E-01
2.88489539036501E-01
3.66859412446906E-01
4.01420147688763E-01
1.67876011777789E-01

f-orbitals	
1.42867047265852E-01	4.81014819456806E-01
-1.18259867173701E-01	7.66366490484168E-01
2.04713641375036E-01	3.98778407223170E-01
4.08588317600017E-01	6.60831890784634E-02
5.50479265466481E-01	6.06870028322653E-03
6.73590454704370E-01	-1.33712791671401E-01

S2: Sample example CP2K input parameters for Na-clinoptilolite:

```
@SET nrep cf

&GLOBAL
  PROJECT suisi-Na
  PRINT_LEVEL MEDIUM
  RUN_TYPE CELL_OPT
&END GLOBAL

&MOTION
  &CELL_OPT
   TYPE DIRECT_CELL_OPT
   KEEP_SYMMETRY
   MAX_DR    1.0E-04
   MAX_FORCE 1.0E-05
   RMS_DR    1.0E-05
   RMS_FORCE 1.0E-04
   MAX_ITER 500
   OPTIMIZER BFGS
   PRESSURE_TOLERANCE 5
  &END CELL_OPT
  &PRINT
   &TRAJECTORY
     FORMAT PDB
   &END
 &END
&END MOTION
```

```
&FORCE_EVAL
 METHOD Quickstep
 &DFT
  BASIS_SET_FILE_NAME BASIS_MOLOPT
  BASIS_SET_FILE_NAME        BASIS_MOLOPT_UCL
  POTENTIAL_FILE_NAME GTH_POTENTIALS
  &MGRID
   NGRIDS 4
   CUTOFF 350              !Ry
   REL_CUTOFF 70
  &END MGRID
  &QS
   EPS_DEFAULT 1.0E-10
  &END QS
  &SCF
   EPS_SCF 1.0E-10                  !Ha
   SCF_GUESS ATOMIC
   MAX_SCF 300
   &DIAGONALIZATION  ON
    ALGORITHM STANDARD
   &END DIAGONALIZATION
    &MIXING
    METHOD BROYDEN_MIXING
    ALPHA 0.4
    NBROYDEN 8
   &END MIXING
  &END SCF
  &XC
```

```
    &XC_FUNCTIONAL
     &BECKE97
      PARAMETRIZATION B97GRIMME
     &END BECKE97
    &END XC_FUNCTIONAL
     &vdW_POTENTIAL
     POTENTIAL_TYPE PAIR_POTENTIAL
     &PAIR_POTENTIAL
      TYPE DFTD3
      PARAMETER_FILE_NAME dftd3.dat
      REFERENCE_FUNCTIONAL B97-D
      R_CUTOFF 25.0
     &END PAIR_POTENTIAL
  &END vdW_POTENTIAL
  &END XC
&END DFT
&SUBSYS
  &CELL
    CELL_FILE_FORMAT CIF
    CELL_FILE_NAME ${nrep}
  &END CELL
  &TOPOLOGY
   CONNECTIVITY OFF
   COORD_FILE_FORMAT CIF
   COORD_FILE_NAME ${nrep}
  &END TOPOLOGY
  &KIND Si
   BASIS_SET TZVP-MOLOPT-GTH-q4
```

```
    POTENTIAL GTH-PBE-q4
   &END KIND
   &KIND O
    BASIS_SET DZVP-MOLOPT-GTH-q6
    POTENTIAL GTH-PBE-q6
   &END KIND
   &KIND Al
    BASIS_SET TZV2P-MOLOPT-SR-GTH-q3
    POTENTIAL GTH-PBE-q3
   &END KIND
  &KIND Na
    BASIS_SET SZV-MOLOPT-SR-GTH-q9
    POTENTIAL GTH-PBE-q9
   &END KIND
 &END SUBSYS
 STRESS_TENSOR ANALYTICAL
 &PRINT
   &STRESS_TENSOR
   &END STRESS_TENSOR
   &FORCES ON
   &END FORCES
 &END PRINT
&END FORCE_EVAL
```

N.B: LS recommends testing the parameters first on the corresponding system.

S3: Discussion on bond lengths:

For cations in Na-clinoptilolite, unlike Si-O distances, there was a shortening of the bond lengths and stronger deviations for both samples. The Na1-O and Na2-O distances showed less deviation with revPBE-D3 for the Agoura sample. In contrast, the B97-D3 showed less deviation for the other sample for both cationic positions. For both samples, the distances were significantly shortened, as depicted in Table S3_a. This could be attributed to the absence of water molecules. The extra-framework cations tend to move closer to the sides of the channels to approach the framework oxygens and form stronger bonds to compensate for the absence of water molecules, as stated earlier. In addition, each experimental cationic site consists of several cations, and those sites are not 100% occupied. Therefore, a direct comparison cannot be made. The lack of the presence of different cations besides water molecules might lead to underbonding or overbonding which both lead to either shortening or lengthening of the bonds, respectively. However, it will always be worth mentioning that experimentally, in dehydration, the cations tend to have shorter distances with the framework, which was confirmed here in this study. All the information and values regarding Si-O and exchangeable cations for the most stable configuration can be found in the SI_Bond-lengths_most-stable-configurations folder.

Table S3_a: Exchangeable cations distances with the framework oxygens of the most stable structures in (a) Agoura and (b) Suisi samples. It should be noted that any cation-O distance within a radius of 4 Å is reported here.

(a) Optimized Agoura Na-clinoptilolite (Exchangeable cations)			
Cations (Å)	**Exp.**	**revPBE-D3**	**B97-D3**
Na (1) – O1	3.392	3.39, 3.15, 2.89, 2.96	2.97, 3.07, 3.02, 3.09
– O2	2.861	2.88, 2.56, 2.67, 2.53, 2.48, 2.76, 2.6, 2.50	2.62, 2.71, 2.54, 2.5, 2.66, 2.52, 2.77, 2.48
– O3	3.121	2.63, 2.64, 2.762, 2.91, 3.03, 2.65, 2.86	2.66, 2.87, 2.75, 2.97, 2.73
Na (2) – O1	2.54	2.45, 2.34	2.35, 2.34
– O10	2.78	2.64, 2.50, 2.58, 2.45	2.48, 2.82, 2.60, 2.66

(b) Optimized Siusi Na-clinoptilolite (Exchangeable cations)			
Cations (Å)	**Exp.**	**revPBE-D3**	**B97-D3**
Na (1) – O1	3.392	2.88, 2.52, 2.82	2.67, 3.25, 2.81, 2.80
– O2	2.861	2.84, 2.49, 2.54, 2.61, 2.83, 2.48, 2.51, 2.59	2.3, 2.69, 2.6, 2.8, 2.89, 2.61, 2.72, 2.37
– O3	3.121	2.55, 2.72, 2.87, 2.76, 2.65, 2.89, 2.80	2.75, 2.99, 2.96, 2.7, 3.10, 2.75
Na (2) – O1	2.54	2.47, 2.40	3.6, 2.36
– O10	2.78	2.36, 2.64, 2.58	2.79, 2.47, 2.64

Table S3_b below illustrates the distance between Na1 and O1 in the 10 MR of channel A of the most stable structure in Na-Ca clinoptilolite to be ~3.33 Å; this is a better agreement with its experimental counterpart which is ~3.39 Å, compared to when Na was the only cation in the cell as shown previously in table S3_a. It is worth noting that this deviation in the first place is attributed to the fact that our model is dehydrated. As mentioned earlier, cations tend to move closer to the framework during the absence of water molecules.

TABLE S3_b: Siusi sample cations-oxygens interatomic distances comparison. DFT calculations for the most stable configuration, as well as experimental refinements, are shown.

Optimized Siusi Na, Ca-clinoptilolite (Exchangeable cations)		
Cations (Å)	**Exp.**	**B97-D3**
Na (1) – O1	3.392	3.017, 3.337
– O2	2.861	2.786, 2.576, 2.915, 2.580
– O3	3.121	2.842, 3.114, 2.876, 3.038
Ca (2) – O1	2.54	2.258, 2.260
– O10	2.78	2.420, 2.309, 2.438, 2.270

S4: Least stable configuration for Na-clinoptilolite:

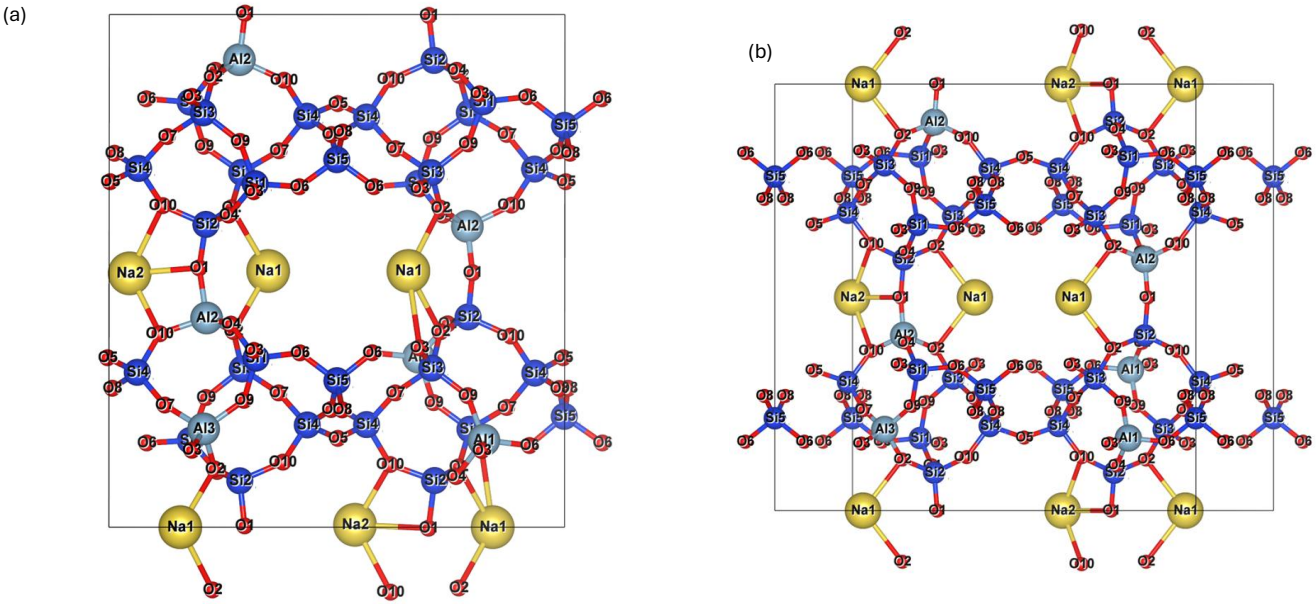


Figure S4: the least stable configuration among the 84 configurations, oriented along the c-direction: (a) portrays the movement of the extra-framework cations post-optimizations, and (b) shows the sites of extra-framework cations prior to optimizations.

In the configuration shown in Figure S4, the left Na1 atom in the 10 MR of channel A is surrounded by Al2 from the left bottom at a distance of 3.56 Å, and another Na2 is at a closer distance of 3.27 Å. Simultaneously, Na1 at the right side of the channel is enclosed by Al2 from above and Al1 from the bottom. The energetically more favorable situation is that the Na2, having originally closer distance to Al2, will move to coordinate with some framework oxygens surrounding the Al2 atom in the 8MR of channel B. This inherently will cause an increase in the distance between the Na1, at the left side of the 10 MR, and Al2 to be ~3.60 Å rather than ~3.56 Å before optimizations. Likewise, the Na1 atom at the right side of the channel will move closer and coordinate with the framework oxygens surrounding Al2. Moreover, this Na1 atom will have poor coordination with the oxygens adjacent to the Al1 atoms resulting in a distance of ~5.52 Å, compared to the previous ~5.39 Å before optimizations, with the Al1. This will lead to frustration for both Na1 atoms in the 10MR. Thereby, they will not be able to move further away from each other as much as in the other configuration, resulting in very short Na-Na separations of 4.90 Å, which was otherwise ~4.67 Å before optimizations. This will cause instability in the energetics of the structure, thus deeming this structure the most unstable of the 84 configurations. In addition, the framework oxygens surrounding Al2 in the left upper part of the structure are completely unsatisfied since Na1 would prefer to coordinate with the oxygens surrounding Al3 at the bottom left. Indeed, while the structure maintains an overall charge balance, there exists local charge dissatisfaction around the Al2 atoms

and the extra-framework cations. The latter also has a dissatisfaction which arises from their exposed positive portions that lack bonding with water molecules. Moreover, at the lower right section of the structure, both Na2 and Na1 atoms shifted away to prevent proximity to the Al1 atom to have distances of ~5.58 Å and ~6.14 Å rather than ~5.39 Å and ~6.10 Å, respectively. The above discussion is relevant to both samples.

References

1. Kudin, K.N.; Scuseria, G.E. Linear-Scaling Density-Functional Theory with Gaussian Orbitals and Periodic Boundary Conditions: Efficient Evaluation of Energy and Forces via the Fast Multipole Method. *Phys. Rev. B* **2000**, *61*, 16440–16453, doi:10.1103/PhysRevB.61.16440.
2. CCCBDB Compare Calculated Bond Lengths Available online: https://cccbdb.nist.gov/calcbondcomp1x.asp (accessed on 16 April 2024).
3. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305, doi:10.1039/B508541A.
4. VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *The Journal of Chemical Physics* **2007**, *127*, 114105, doi:10.1063/1.2770708.
5. Exercises:2015_cecam_tutorial:Basis_set_optimisation_using_optimize_basis [CP2K Open Source Molecular Dynamics] Available online: https://www.cp2k.org/exercises:2015_cecam_tutorial:basis_set_optimisation_using_optimize_basis (accessed on 16 April 2024).
6. Ling, S. Optimisation of Basis Sets and Pseudopotentials. **2007**.