Electronic supplementary materials *Mendeleev Commun*., 2020, **30**, 741–743

Anisotropic electrostatic models of nitrogen and phosphorus: the variation and the interpretability of the electrostatic parameters in response to structure variation

Dmitry A. Shulga, Arslan R. Shaimardanov and Vladimir A. Palyulin

Molecular electrostatic potential on a grid around each molecule by:

1. the reference MP2/aug-cc-pVTZ calculation;
2. charge only (m-m) model optimally fit to the reference;
3. anisotropic multipole (m-mdzqz) model optimally fit to the reference;
4. atomic charge potential from the anisotropic multipole (m-mdzqz) model optimally fit to the reference;
5. N/P dipole (dz) potential from the anisotropic multipole (m-mdzqz) model optimally fit to the reference;
6. N/P quadrupole (qz) potential from the anisotropic multipole (m-mdzqz) model optimally fit to the reference.

Red — the positive values, whereas blue — the negative values of the potential. Both sides are clipped (maximum intensity of color) at ±0.01 a.u. (0.27 V).

Table of Contents

Methods details

The structure geometries were obtained using OpenBabel package [1] (–gen3D option). The reference molecular electrostatic potential (MEP) was calculated at MP2/aug-cc-pVTZ level using FireFly [2]. Connolly grid with 5 layers (1.6, 1.8, 2.0, 2.2, 2.4 times VdW radii) and density of 4 points per $Å^2$ were used. All anisotropic electrostatic models were fit to the reference MEP of each molecule using elec_tools [3]. The analysis was conducted using **m-mdzqz** anisotropic model, in which all atoms were represented with point charges (monopoles, **m**), whereas the nitrogen and phosphorus atoms were additionally represented with atomic dipole, **dz**, and quadrupole, **qz**, components, aligned along the *Z* axis, formed by the direction of the bisector of the vectors of all three sigma bonds (Figure S1). The choice of this local coordinate system reduces the number of multipole parameters. The atomic dipole, **dz**, is fully defined by a single number – the magnitude of the vector along the *Z* axis direction. Likewise, the atomic quadrupole 3x3 matrix, expressed in a traceless form ($Q_{XX} + Q_{YY} + Q_{ZZ} = 0$) with the additional symmetry constraint $Q_{XX} = Q_{YY}$ and no off diagonal elements, is fully defined by a single scalar value, Q_{ZZ}, so that $Q_{XX}=Q_{YY}=-0.5 \cdot Q_{ZZ}$ and $Q_{AB}=0$, where $A \neq B$.

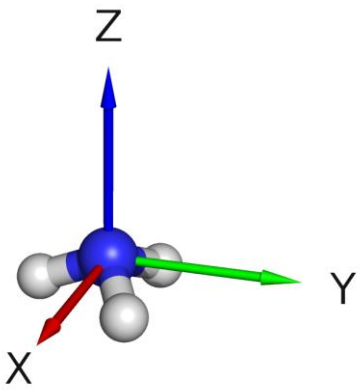


Figure S1. The local multipole axes used for amine N and phosphine P multipole expansion.

The multipole MEP is thus modeled with (1)

$$V_i = \sum_{j}^{Natoms} \frac{q_i}{|\mathbf{R}_{ij}|} + \frac{\mathbf{R}_{ia}\mathbf{d}}{|\mathbf{R}_{ia}|^3} + \frac{\mathbf{R}_{ia}^{\mathrm{T}}\,\mathrm{Q}\,\mathbf{R}_{ia}}{|\mathbf{R}_{ia}|^5}\ , \qquad (1)$$

where V_i is the potential in the *i*-th grid point; $\mathbf{R}_{ij}$ – is the vector from *i*-th grid point to *j*-th atom; q_j – is the charge on *j*-th atom; $\mathbf{R}_{ia}$ – is the vector from *i*-th grid point to the N or P atom, expanded with the dipole vector **d** and quadrupole matrix Q with the components described above.

To estimate the enhancement in quality brought by the introduction of multipoles on N and P atoms the charge only model (**m-m** in Table 1 and 2) was used with the MEP described by the first term of (1).

The simplified Hansch type equation (2) is used to correlate the anisotropic parameters.

$$p_i = a \cdot \mathrm{count}_i(F) + b \cdot \mathrm{count}_i(C) + c, \quad (2)$$

where p_i is the dependent (reference) anisotropic parameter for molecule *i; a*, *b*, *c* the coefficients sought by the regression; $count_i$(F) is the number of F atoms, whereas $count_i$(C) is the number of C atoms (or Me groups) in *i*-th molecule.

The parameters distribution and correlation with substituents were studied using GNU R v3.2.3 [4].

References

1. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminform.*, 2011, **3**, 33.
2. A. A. Granovsky, *Firefly, version 8*, http://classic.chem.msu.su/gran/firefly/index.html.
3. O. I. Titov, D. A. Shulga and V.A. Palyulin, web http://qsar.chem.msu.ru/elec_tools/.
4. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015, https://www.R-project.org/.

Residual MEP error and anisotropic parameters

Table S1. The residual MEP error and the values of anisotropic multipole parameters for N containing compounds of the training set.

	MEP error, kcal·mol^{-1}el^{-1}		Anisotropic parameters for models, a.u.			
			m	**m-mdzqz**		
Molecule	**m-m**	**m-mdzqz**	**m**	**m**	**dz**	**qz**
NF_3	1.21	0.65	0.09	0.43	-0.36	-0.60
NF_2CH_3	1.55	0.84	-0.09	0.38	-0.42	-1.00
NF_2H	1.69	1.00	-0.31	0.18	-0.38	-0.93
NFH_2	1.89	1.12	-0.63	-0.17	-0.31	-1.14
$NFCH_3H$	2.02	1.29	-0.45	0.18	-0.50	-1.34
$NF(CH_3)_2$	1.57	0.83	-0.09	0.39	-0.47	-1.53
$N(CH_3)_3$	1.47	0.64	-0.07	0.41	-0.45	-2.02
$N(CH_3)_2H$	2.25	1.48	-0.54	0.19	-0.62	-1.85
NCH_3H_2	2.37	1.60	-0.86	-0.19	-0.45	-1.58
NH_3	1.98	0.36	-0.92	-0.95	0.05	-0.86
Mean	1.80	0.98	-0.39	0.09	-0.39	-1.29
Std.dev.	0.36	0.39	0.35	0.43	0.18	0.46

Table S2. The residual MEP error and the values of anisotropic multipole parameters for P containing compounds of the training set.

	MEP error, kcal·mol^{-1}el^{-1}		Anisotropic parameters for models, a.u.			
			m-m	**m-mdzqz**		
Molecule	**m-m**	**m-mdzqz**	**m**	**m**	**dz**	**qz**
PF_3	2.26	0.57	0.41	1.02	-0.96	-1.41
PF_2CH_3	2.66	0.67	0.11	0.90	-1.18	-1.86
PF_2H	3.12	1.55	-0.18	0.56	-0.82	-2.11
PFH_2	3.01	1.27	-0.42	0.41	-0.76	-2.95
$PFCH_3H$	3.14	1.29	-0.35	0.50	-0.90	-2.75
$PF(CH_3)_2$	2.93	0.66	-0.08	0.79	-1.28	-2.40
$P(CH_3)_3$	3.06	0.56	-0.15	0.65	-1.28	-2.94
$P(CH_3)_2H$	3.17	1.00	-0.44	0.42	-0.91	-3.31
PCH_3H_2	2.99	1.19	-0.57	0.21	-0.71	-3.26
NH_3	2.78	0.63	-0.54	-0.10	-0.32	-2.72
Mean	2.91	0.94	-0.22	0.53	-0.91	-2.57
Std.dev.	0.28	0.36	0.31	0.33	0.29	0.62

Parameters of the Hansch-type correlations

Table S3. Parameters of the Hansch-type correlations (2) of the components of the anisotropic multipole models with the variation of the substituents for amine N compounds.

p	R^2	stderr	R^2_{adj}	F	a	b	c
m	0.91	0.14	0.89	36.12	0.39 ± 0.05 (***)	0.39 ± 0.05 (***)	-0.7 ± 0.1 (***)
dz	0.48	0.14	0.34	3.276	-0.07 ± 0.05 (-)	-0.13 ± 0.05 (*)	-0.2 ± 0.1 (-)
qz	0.91	0.16	0.88	35.31	0.15 ± 0.06 (*)	-0.31 ± 0.06 (**)	-1.1 ± 0.1 (***)

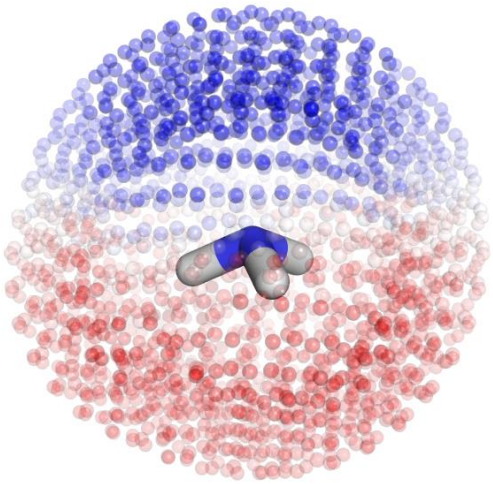
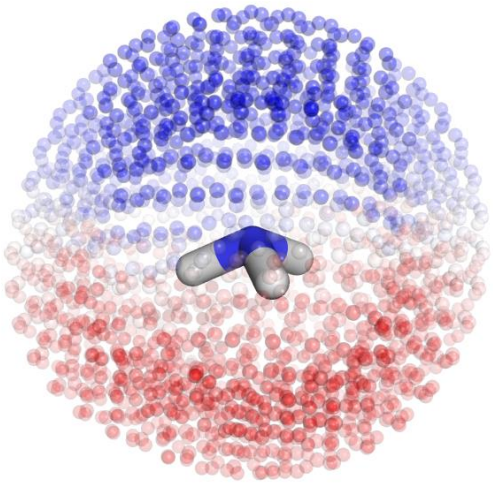
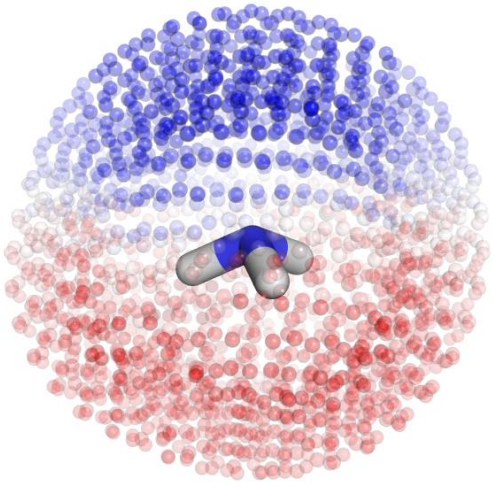
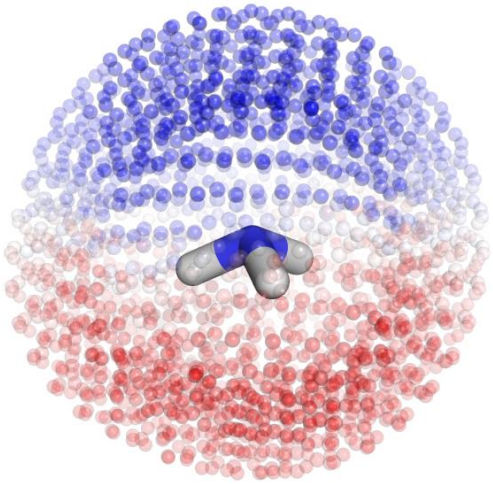
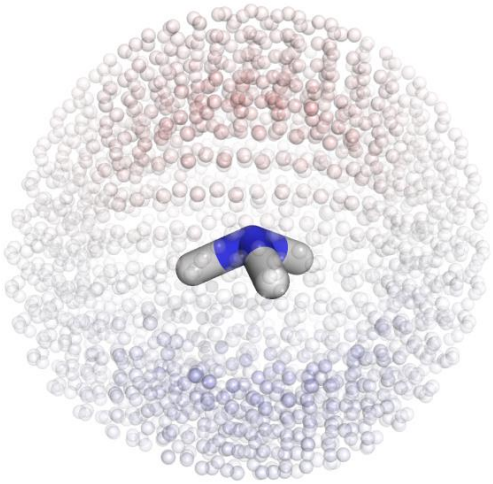
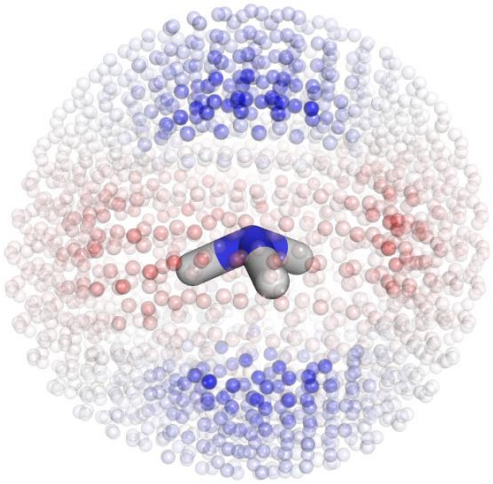
* Parameter significance: *** - high, ** - moderate, * - satisfactory, '-' - low.

Table S4. Parameters of the Hansch-type correlations (2) of the components of the anisotropic multipole models with the variation of the substituents for phosphine P compounds.

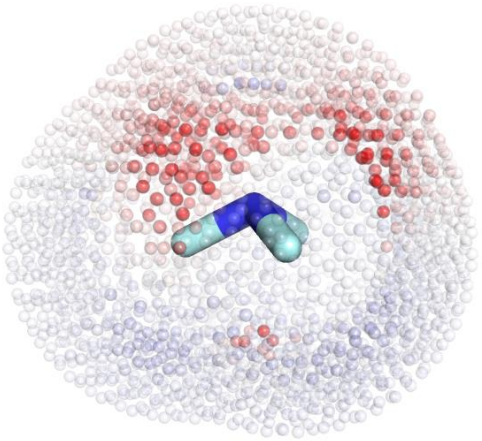
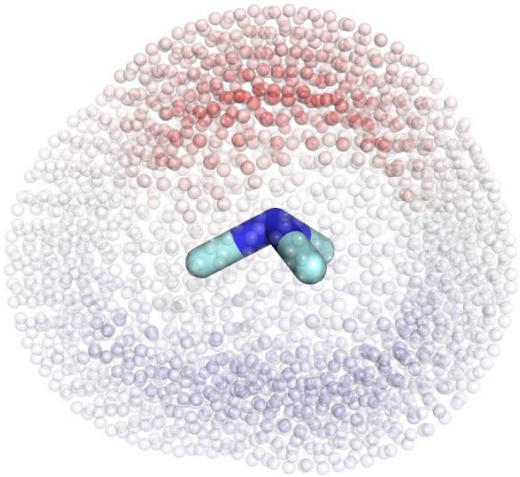
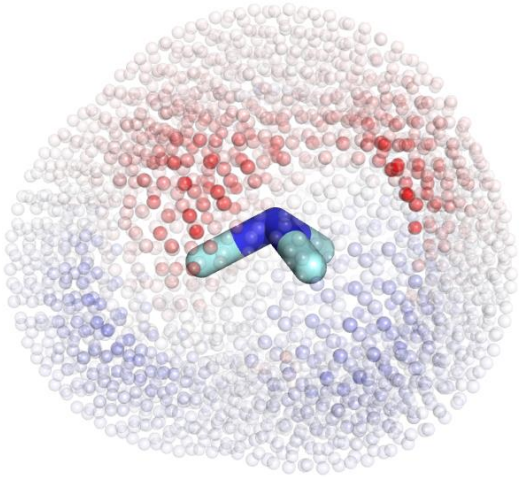
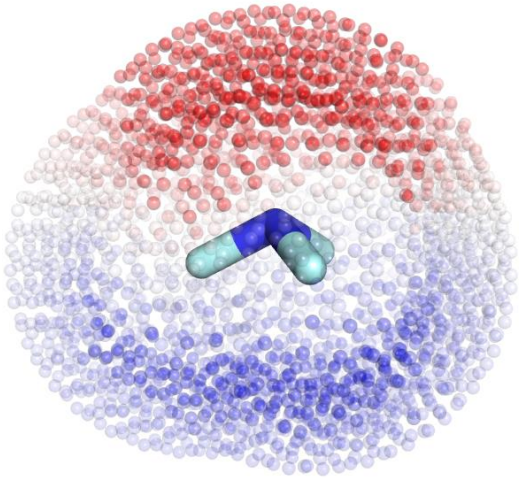
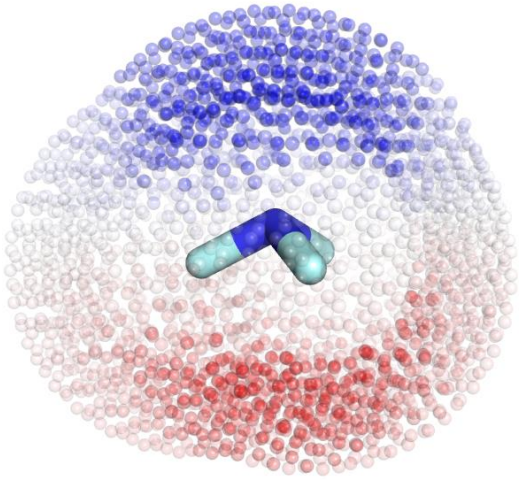
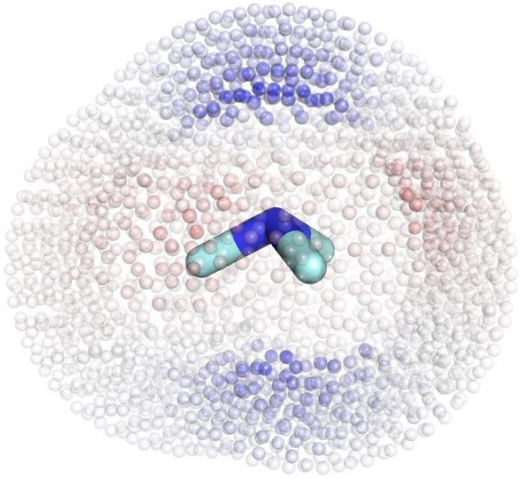
p	R^2	stderr	R^2_{adj}	F	a	b	c
m	0.97	0.06	0.97	125.6	0.35 ± 0.02 (***)	0.24 ± 0.02 (***)	-0.06 ± 0.04 (-)
dz	0.94	0.08	0.92	50.65	-0.22 ± 0.03 (***)	-0.30 ± 0.03 (***)	-0.39 ± 0.06 (***)
qz	0.86	0.26	0.82	21.95	0.56 ± 0.09 (***)	0.04 ± 0.09 (-)	-3.2 ± 0.2 (***)

* Parameter significance: *** - high, ** - moderate, * - satisfactory, '-' - low.

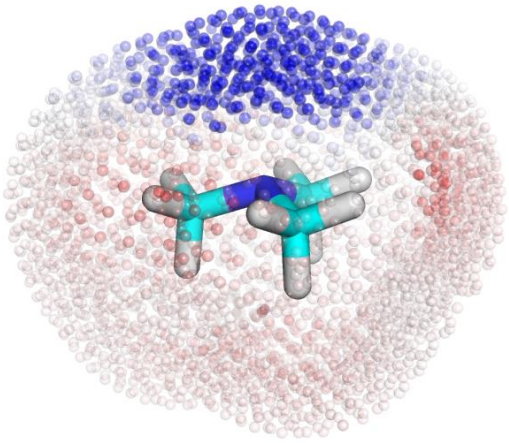
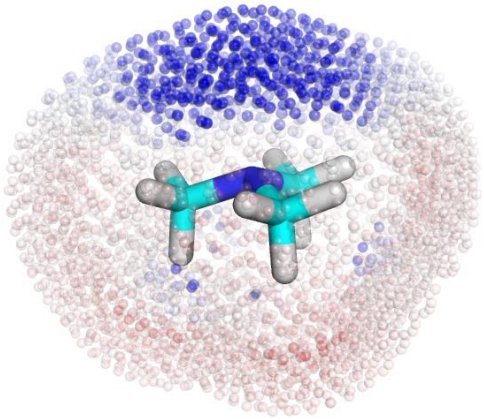
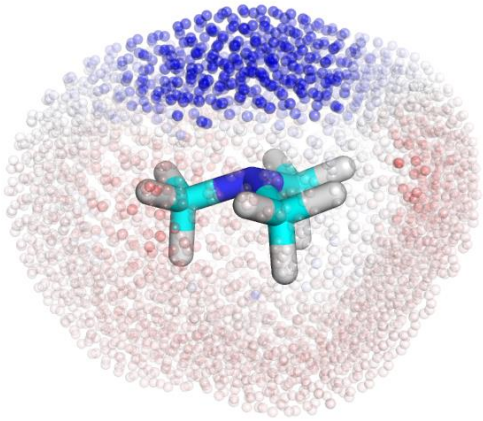
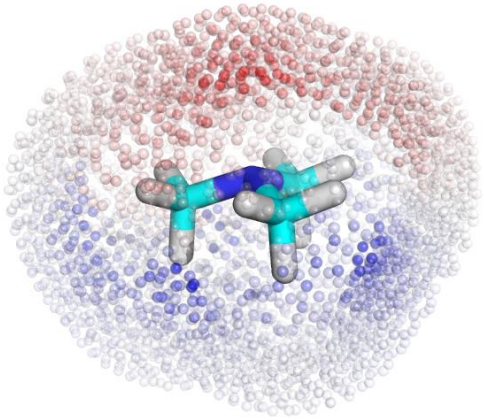
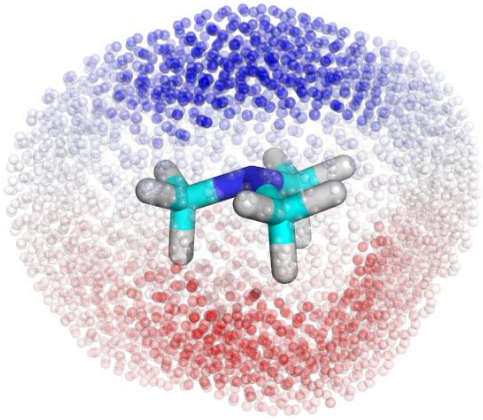
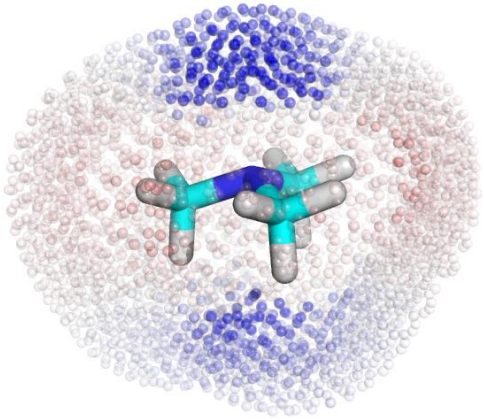
NH_3

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	Anisotropic (charges on *)
	
Anisotropic dipole (dz)	Anisotropic quadrupole (qz)
	

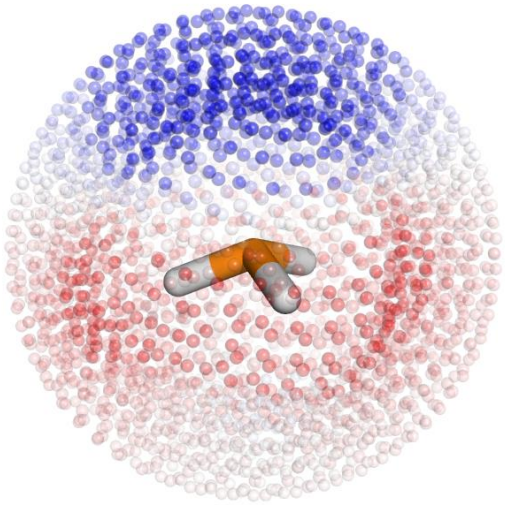
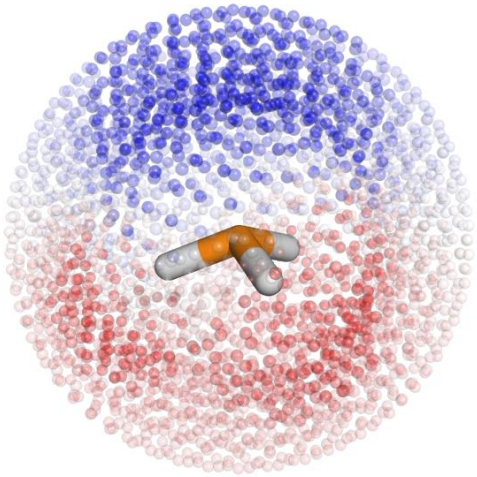
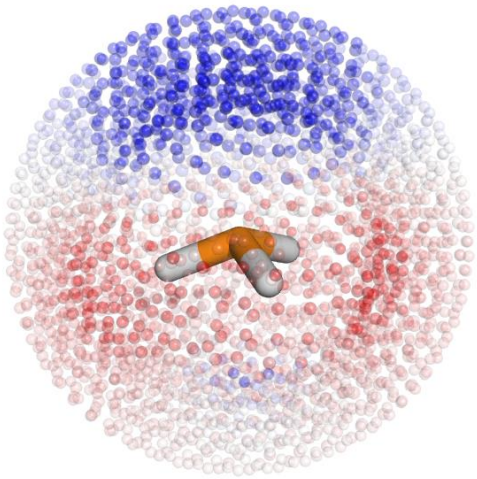
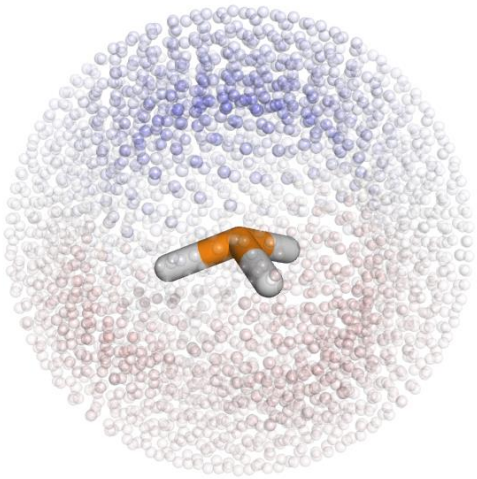
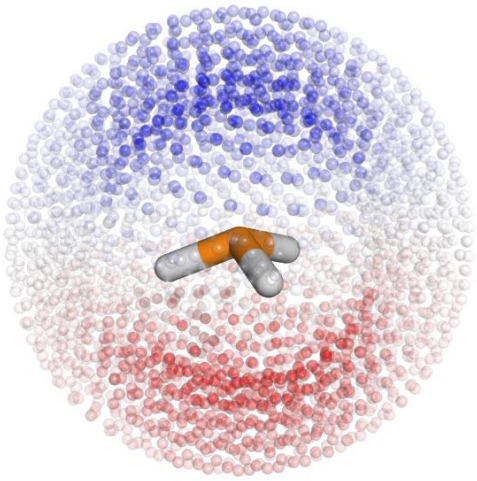
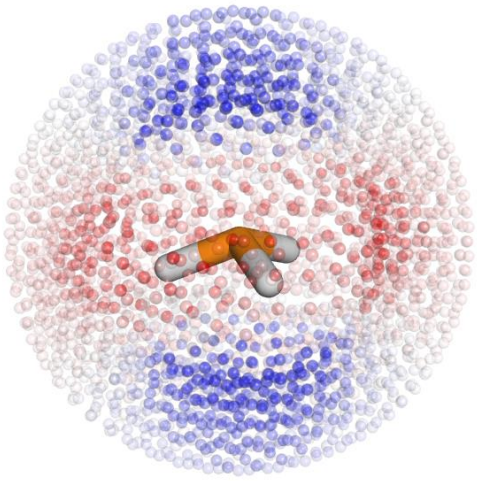
NF_3

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	**Anisotropic (charges on *)**
	
Anisotropic dipole (dz)	**Anisotropic quadrupole (qz)**
	

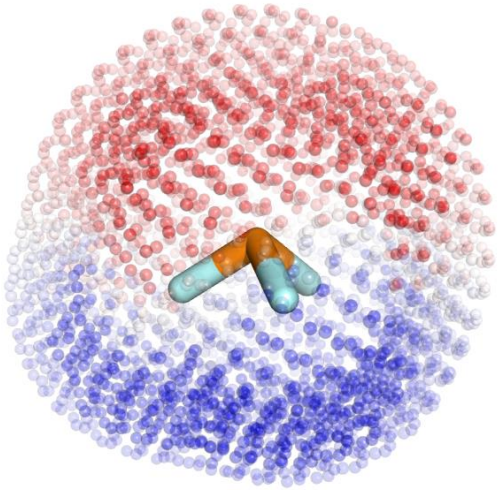
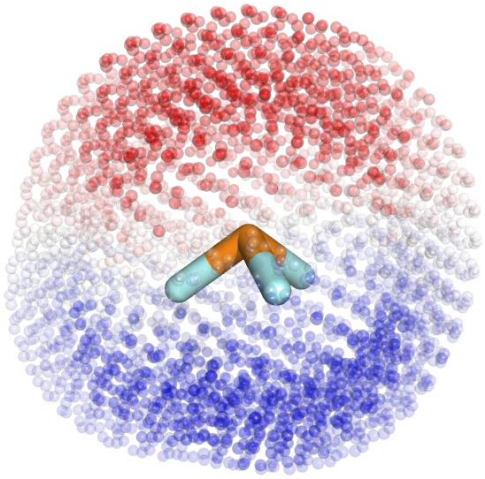
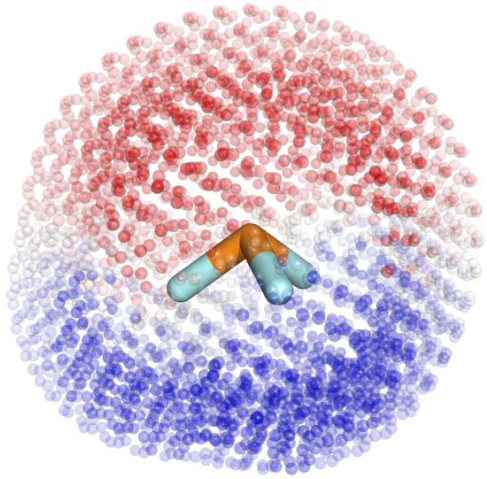
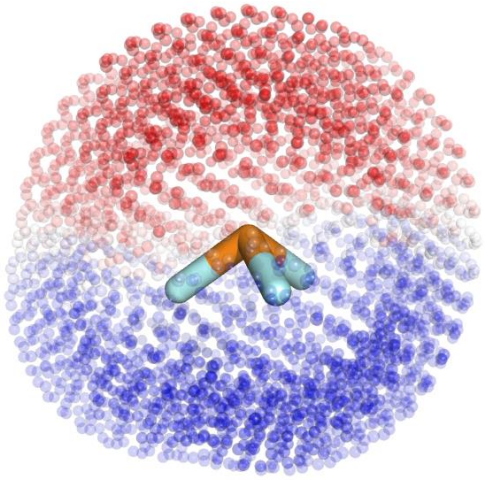
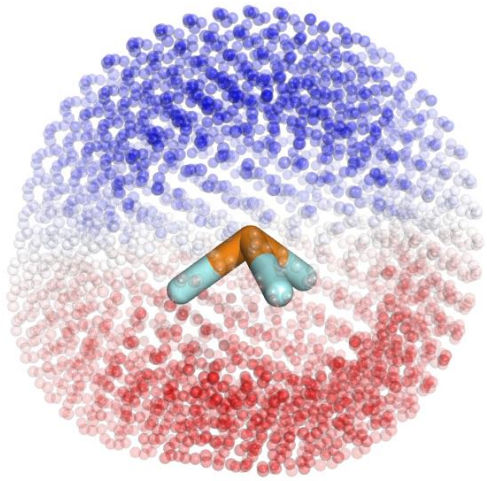
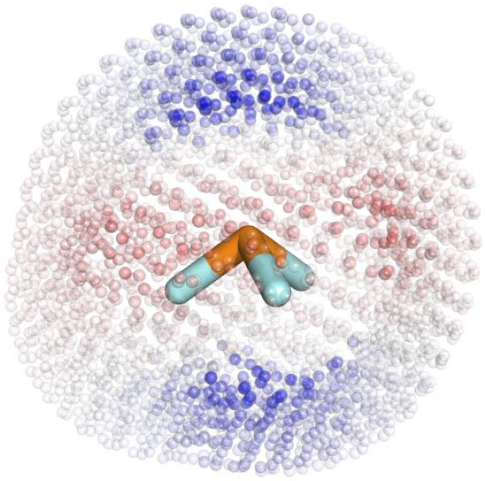
$N(CH_3)_3$

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	Anisotropic (charges on *)
	
Anisotropic dipole (dz)	Anisotropic quadrupole (qz)
	

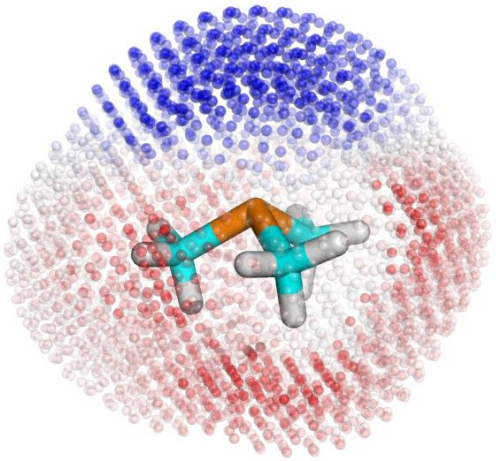
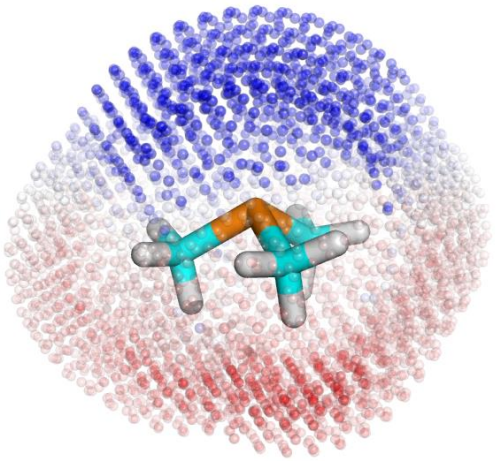
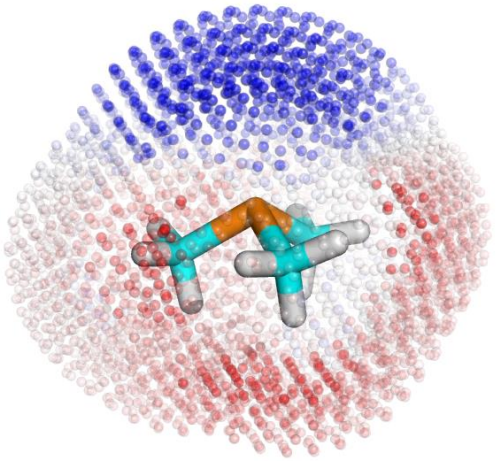
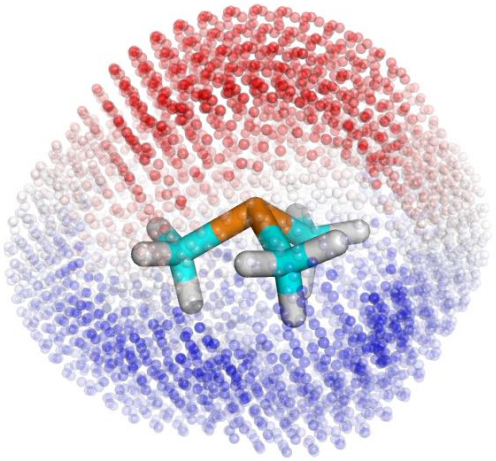
PH_3

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	Anisotropic (charges on *)
	
Anisotropic dipole (dz)	Anisotropic quadrupole (qz)
	

PF_3

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	Anisotropic (charges on *)
	
Anisotropic dipole (dz)	Anisotropic quadrupole (qz)
	

$P(CH_3)_3$

Reference	Charge only (m-m)
	
Anisotropic (m-mdzqz)	Anisotropic (charges on *)
	
Anisotropic dipole (dz)	Anisotropic quadrupole (qz)