Supporting Information

The CO–N_2 intermolecular potential energy surface revisited

Dariusz Kędziera, Omar Rodríguez-López,
Ad van der Avoird, and Berta Fernández

1 Interaction sites and asymptotic site parameters

In addition to the atomic sites, we introduced five off-atomic sites on CO and six on N_2, arranged in three symmetry-related pairs. All off-atomic sites lie on the molecular axes. The CO and N_2 bond lengths are fixed at 1.128323 and 1.09768 Å, respectively.

Table S1 gives the positions, charges, and polarizabilities of all sites. Their indices match the sums in Eq. S1: $a = 1, \ldots, 7$ labels the sites on CO and $b = 1, \ldots, 8$ those on N_2. Coordinates are measured from the center of mass of each monomer. For CO, positive z points from C to O; for N_2, the N atoms are at $z = \pm 0.5488400$ Å. All sites have $x = y = 0$. We use short site labels throughout; `XC1` and `XC2` distinguish the two off-atomic sites beyond the C atom.

The two members of each symmetry-related N_2 pair have the same charge, polarizability, and interaction parameters. Charges q are given in units of the elementary charge, and isotropic polarizabilities α in atomic units (a_0^3). Charges assigned to off-atomic sites are fitted model parameters.

Table S1: Positions and electrostatic/polarization parameters of the 15 sites in the final model.

Monomer	Index a/b	Site	z (Å)	q (e)	α (a_0^3)
CO	1	`C`	-0.6444855	-0.116782230625	10.8298922098
CO	2	`O`	+0.4838375	+0.0378425574501	6.06379945585
CO	3	`XC1`	-1.2144855	+15.672823247	0
CO	4	`MaCO`	-0.2303240	+15.4202361584	0
CO	5	`XO`	+0.6038375	+8.67516776181	0
CO	6	`MCO`	-0.4903240	-53.4317221825	0
CO	7	`XC2`	-0.8344855	+13.7424346885	0
N_2	1	`N1`	+0.5488400	-0.121375610909	0.103893077258
N_2	2	`N2`	-0.5488400	-0.121375610909	0.103893077258
N_2	3	`XaN1`	+0.9888400	-0.00647327578332	0
N_2	4	`XaN2`	-0.9888400	-0.00647327578332	0
N_2	5	`MaN1N2`	+0.3600000	+0.0553866602215	0
N_2	6	`MaN2N1`	-0.3600000	+0.0553866602215	0
N_2	7	`XN1`	+0.7288400	+0.072462226471	0
N_2	8	`XN2`	-0.7288400	+0.072462226471	0

Only the atomic sites have nonzero dispersion coefficients C_6^s, C_8^s, and C_{10}^s (Table S2). The index s distinguishes the three atomic-site types: $s = 1$ for C, $s = 2$

for O, and $s = 3$ for either N atom. The coefficients vanish at all off-atomic sites. Pairwise coefficients follow the combination rule given below.

Table S2 Atomic-site dispersion coefficients C_n^s (kcal mol^{-1} Å^n).

Atomic site	s	C_6^s	C_8^s	C_{10}^s
`C`	1	362.212893892	3866.57343078	48471.3021647
`O`	2	150.820927291	1066.56237907	8441.51401275
`N1`, `N2`	3	206.887925295	3550.11677514	63260.5500339

2 Analytical potential and fitted parameters

We use the site–site form of the final AutoPES potential. The total interaction energy is written as a sum of a polarization term and contributions from all pairs of sites on different monomers:

$$V_{\text{int}} = V_{\text{ind}}(\text{CO}, \text{N}_2) + \sum_{a=1}^{7} \sum_{b=1}^{8} u_{ab}(r_{ab}). \tag{S1}$$

For sites a on CO and b on N_2, separated by $r_{ab} = |\boldsymbol{r}_a - \boldsymbol{r}_b|$, the pair interaction is

$$\begin{aligned} u_{ab}(r) = & \left[1 + a_1^{ab} r + a_2^{ab} r^2 + a_3^{ab} r^3\right] A^{ab} e^{-\beta^{ab} r} \\ & + k_{\text{e}}\, f_1(\delta_1^{ab}, r) \frac{q_a q_b}{r} - \sum_{n=6,8,10} f_n(\delta_n^{ab}, r) \frac{C_n^{ab}}{r^n} + \frac{A_{12}^{ab}}{r^{12}}, \end{aligned} \tag{S2}$$

Distances are in Å and energies in kcal mol^{-1}. The Coulomb conversion factor is $k_{\text{e}} = 332.06371$ when charges are in units of e and distances in Å. The dispersion coefficients are obtained from the atomic-site values in Table S2:

$$C_n^{ab} = \sqrt{C_n^a C_n^b} \qquad (n = 6, 8, 10). \tag{S3}$$

We use the Tang–Toennies damping functions,

$$f_n(\delta, r) = 1 - e^{-\delta r} \sum_{k=0}^{n} \frac{(\delta r)^k}{k!}. \tag{S4}$$

The final potential includes a self-consistent induced point-dipole polarization model, enabled by the `POLARIZABLE = T` option in AutoPES. Only the atomic sites have nonzero polarizabilities; the off-atomic sites contribute to the electric field

through their permanent charges. For each polarizable site, the induced dipole is determined by the field of the other monomer and the induced dipoles on that monomer. In atomic units,

$$\boldsymbol{\mu}_a^{\mathrm{ind}} = \alpha_a \left(\boldsymbol{E}_a + \sum_{b \in \mathrm{N_2}} \boldsymbol{T}_{ab} \boldsymbol{\mu}_b^{\mathrm{ind}} \right), \tag{S5}$$

$$\boldsymbol{E}_a = \sum_{b \in \mathrm{N_2}} f_1(\delta_1^{ab}, r_{ab}) \frac{q_b \boldsymbol{r}_{ab}}{r_{ab}^3}, \tag{S6}$$

$$\boldsymbol{T}_{ab} = f_3(\delta_{\mathrm{p}}^{ab}, r_{ab}) \left[\frac{3 \boldsymbol{r}_{ab} \otimes \boldsymbol{r}_{ab}}{r_{ab}^5} - \frac{\boldsymbol{I}}{r_{ab}^3} \right]. \tag{S7}$$

For sites on $\mathrm{N_2}$, the corresponding sums run over the CO sites. The field and dipole expressions use atomic units, with $\boldsymbol{r}_{ab} = \boldsymbol{r}_a - \boldsymbol{r}_b$ in bohr. In the damping functions, however, distances are expressed in Å and the damping parameters in Å^{-1}. The polarization energy in $\mathrm{kcal\,mol^{-1}}$ is

$$V_{\mathrm{ind}} = -\frac{E_{\mathrm{h}}^{\mathrm{kcal}}}{2} \left(\sum_{a \in \mathrm{CO}} \boldsymbol{E}_a \cdot \boldsymbol{\mu}_a^{\mathrm{ind}} + \sum_{b \in \mathrm{N_2}} \boldsymbol{E}_b \cdot \boldsymbol{\mu}_b^{\mathrm{ind}} \right), \tag{S8}$$

where $E_{\mathrm{h}}^{\mathrm{kcal}} = 627.50947\,\mathrm{kcal\,mol^{-1}}$ is the Hartree-to-kcal/mol conversion factor. The symmetry-related sites on $\mathrm{N_2}$ share the same pair parameters, leaving 28 distinct CO–$\mathrm{N_2}$ parameter sets. In Table S3, each interaction pair is identified by the site indices a and b given in Table S1. Only four representative $\mathrm{N_2}$ sites ($b = 1, 3, 5, 7$) are listed; the same parameters apply to their symmetry-related partners ($b = 2, 4, 6, 8$, respectively).

The amplitudes A^{ab} are given in $\mathrm{kcal\,mol^{-1}}$, the exponents β^{ab} and δ_1^{ab} in Å^{-1}, and the polynomial coefficients a_i^{ab} in Å^{-i}. The amplitudes are reported directly, rather than as their fitted logarithms.

Table S3: Short-range and electrostatic damping parameters for the 28 distinct CO–N_2 site pairs. Site indices refer to Table S1.

a	b	A^{ab}	β^{ab}	a_1^{ab}	a_2^{ab}	a_3^{ab}	δ_1^{ab}
1	1	1908.02024241	3.9694800744	293.15799785	-327.75468743	82.747734043	2.0142701641
1	3	1092.97366828	4.0312724737	-65.77036111	101.26794892	-35.330282965	1.8799976605
1	5	66760.5274745	4.0744289034	-8.5196603679	7.4616721486	-1.3718487035	1.9695695777
1	7	1242.97518565	2.5029299685	24.12733491	-19.87839444	4.6896388677	1.9705128156
2	1	462901.246629	4.0007598748	-1.4159078877	0.65730834891	-0.054258595279	1.999667341
2	3	501.918742169	3.1286508553	-7.0441684403	3.6711176067	-0.25818988816	1.8631836421
2	5	85789.0866616	4.0626918165	-1.7538222001	0.9984573765	-0.3363926573	1.9274293074
2	7	24872.6050302	3.9475486534	-1.6780142757	1.3978356334	-0.87716931649	2.1706545069
3	1	1637.40761522	3.9276580135	2.5442622626	-5.7368865532	1.5610624923	1.896129035
3	3	125.672183913	2.0352752033	-3.3611177832	1.9971068448	-0.35629235687	1.9999205924
3	5	166.781863207	1.8343440876	1.8716730846	-0.76585849109	0.073363719693	1.8262761009
3	7	235.44426099	2.2539996792	7.6178045654	-5.7717610117	1.5329426697	1.9165427107
4	1	20780.4981717	4.0913401472	-29.9729792	21.620769034	-9.5705166214	1.9125726411
4	3	1115.2621855	1.9232023743	-0.60743263416	0.12051743586	-0.004700531989	1.9769207108
4	5	48616.9434943	2.1655312135	-0.56228561722	0.10997447968	-0.0076685789533	1.8723022725
4	7	71639.9117127	2.5523442145	-0.58282453258	0.18663359724	-0.03198432129	2.0839979734
5	1	305.575279478	4.0539677061	-470.39410537	425.12458493	-200.82209903	1.9248948083
5	3	193.386842343	4.0873441228	-70.897230389	94.78400716	-34.073174467	1.8872308757
5	5	963.387041812	3.8811801477	92.839725339	-63.118813417	23.187824752	1.925337956
5	7	6690.01428516	4.0477129536	13.897874473	-12.85652368	5.9729538153	1.972485952
6	1	235.983843784	4.0166223574	150.90035483	827.97802036	223.09635654	2.0351438227
6	3	54135.8680484	4.0073260983	-1.1711621867	-0.0066729026567	0.11415642449	1.9119398704
6	5	63831.7301452	3.941997929	1.8710620935	-4.5624359165	0.45825821488	1.9393274628
6	7	170.403513191	3.9125049449	-677.52769762	-428.18755389	6.4142743754	1.9687770284
7	1	23925.6729761	3.84020815	-21.003426353	16.269454856	-5.3890194104	2.0131100644
7	3	500.073700465	4.0995716629	40.85111462	-82.871658112	34.463081458	2.0645074488
7	5	91805.3298141	4.009749651	1.8315154431	-1.5574519686	0.58195949861	1.926923155
7	7	60367.9626259	2.4879407111	-1.0056031103	0.48469700117	-0.096182173127	1.8603801458

Only the atomic C–N and O–N pairs have nonzero A_{12}^{ab} coefficients and the associated damping parameters δ_6^{ab}, δ_8^{ab}, δ_{10}^{ab}, and δ_{p}^{ab}. Their values are collected in Table S4. For all other site pairs, $A_{12}^{ab} = C_n^{ab} = 0$, so these damping parameters do not enter the potential. The coefficients A_{12}^{ab} are given in $\mathrm{kcal\,mol^{-1}\,\AA^{12}}$, while the damping parameters δ_6^{ab}, δ_8^{ab}, δ_{10}^{ab}, and δ_{p}^{ab} are expressed in $\mathrm{\AA^{-1}}$.

Table S4 Additional atomic-site pair parameters. Values for N1 also apply to the symmetry-equivalent site N2.

CO site	N_2 site	A_{12}^{ab}	δ_6^{ab}	δ_8^{ab}	δ_{10}^{ab}	δ_{p}^{ab}
C	N1	1274.81936701	1.9229663467	1.9180051701	1.8260303857	1.9882845081
O	N1	925.691582456	2.0747067155	2.0730057457	1.8681942032	2.0165988944

3 Final fitting statistics

The final potential was fitted to 3879 *ab initio* interaction energies using a third-order polynomial and 84 fitted linear coefficients. Table S5 gives the fitting RMSE for four energy ranges. These are errors on the data used to fit the potential, not independent validation errors.

The lowest *ab initio* energy in this dataset is $E_{\mathrm{min}} = -0.3382000012\,\mathrm{kcal\,mol^{-1}}$. As described in the main text, off-site positions were selected by minimizing the sum of the RMSE values for $E_{\mathrm{int}} < E_{\mathrm{min}}/2$ and $E_{\mathrm{int}} < 0$. For the final model this sum is $F = 0.002223\,\mathrm{kcal\,mol^{-1}}$.

Table S5 Final AutoPES fitting statistics. The cutoffs refer to the *ab initio* interaction energies.

Energy range	Number of points	RMSE ($\mathrm{kcal\,mol^{-1}}$)	RMSE ($\mathrm{cm^{-1}}$)
$E_{\mathrm{int}} < E_{\mathrm{min}}/2$	1430	0.001025	0.36
$E_{\mathrm{int}} < 0$	2811	0.001198	0.42
$E_{\mathrm{int}} < 10\,\mathrm{kcal\,mol^{-1}}$	3838	0.011793	4.13
$E_{\mathrm{int}} < 100\,\mathrm{kcal\,mol^{-1}}$	3879	0.026554	9.29