

Perturbation Expansion Theory Corrected From Basis Set Superposition Error II. Charge-transfer, pair correlation and dispersion terms

Suehiro Iwata

*Center for Quantum Life Sciences and Graduate School of Science,
Hiroshima University, ADSM 701, Kagamiyama, Higashi-Hiroshima 739-8526 Japan*

Takeshi Nagata

*Computing and Communication Center,
Kyusyu University, Fukuoka, Japan**

(Dated:)

Abstract

The second order perturbation theory based on the locally projected molecular orbitals is developed. A few test calculations with cc-pVDZ and aug-cc-pVDZ basis sets are carried out for the dimers, $(\text{H}_2\text{O})_2$ and $(\text{HF})_2$. The charge transfer terms remove the deficiency of the locally projected self-consistent field method for molecular interaction (LP SCF MO MI), and the potential energy curves calculated with aug-cc-pVDZ are very close to the corresponding curves of the counterpoise corrected SCF energy. Only after adding the spin-exchanged dispersion type to the dispersion and intra-molecular pair correlation terms, the calculated potential energy curves become close to those of the counterpoise corrected second order Møller-Plesset (MP2). Pragmatic approaches for reducing the influence of the basis set superposition error are proposed.

* The present address: Department of Chemistry, Iowa State University

I. INTRODUCTION

Molecular clusters bounded with the van der Waals forces and hydrogen bonds are extensively studied both experimentally and theoretically. Every six year Chemical Review publishes a thematic issue on van der Waals molecules[1],[2]. The review articles in these issues report great progress both in experimental and theoretical studies. In particular, the contribution from the theoretical and computational works becomes more important in the last decade. Now, most of the experimental papers on the molecular clusters cite the theoretical counter parts or include some computational studies carried out along with their experiments. In these practical computational studies of the weak bonding, the supermolecule method is mostly used with the self-consistent field molecular orbital (SCF MO), the second order Møller-Plesset method (MP2) or the density functional theory(DFT). Unfortunately it has been known that with the supermolecule method the basis set superposition error (BSSE) has to be removed to properly estimate the binding energy and even to optimize the geometry of the clusters, as was compactly reviewed by van Duijneveldt et al[3]. To remove the BSSE, the counterpoise method by Boys and Bernardi[4] and by Jansen and Ross[5] is widely used. In their review, van Duijneveldt et al concluded that CP correction is a proper procedure despite numerous criticism[3]. Since this review, there are a few works which support this conclusion (see the review by Chalanski and Szczeniak[6]). The calculations for the BSSE correction become a routine job after the last revision of Gaussian 98, although the job requires the extensive cpu, particularly when the number of the constituent molecules is larger than 2. Besides, there is some ambiguity in defining the molecular units for the counterpoise procedure, when the intra-cluster reaction or the charge separation is involved[7].

To avoid the BSSE a priori, several methods are developed. The symmetry-adapted perturbation theory (SAPT) has been applied mostly for simpler systems, and recently the applications become much broader[8]. Another approach is Chemical Hamiltonian Approach (CHA) by Mayer and his collaborators in the second order Møller-Plesset theory[9]. The calculated binding energy of both methods are in good agreement with the CP corrected energy, when properly compared. In the CHA, the agreement persists regardless of the choice of the basis set. Another possible approach is the use of the local Møller-Plesset perturbation[10],[11], which is a more general method applicable to the large covalent bonded

molecules[12].

A decade ago, Gianinnetti et al[13] proposed a method, which they claimed that it is BSSE free. They called their method "Self-Consistent Field molecular orbitals for Molecular Interaction (SCF MI)". The molecular orbitals are expanded strictly in terms of the local basis sets (Strictly Monomer Basis Set). Famulari et al [14] reported extensive test calculations with large basis sets. One of the advantages of SCF MI is that it has a simple form of wave function, and the disadvantage is the non-orthogonality among the occupied orbitals. Using the valence bond approach, Calderoni et al extended the method to take into account the dispersion term, and calculated the binding energies for a few basic systems to compare them with those of the other approaches[15]. The SCF MI has received a strong criticism from Hamza et al [16];they argue that the method excludes the physically important, true charge-transfer effects.

Recently we studied the SCF MI and obtained a set of equations for multi-fragment systems using the projection operators[17]. The equations are equivalent to those derived by Gianinnetti et al[13] for the closed shell dimer, although it is not apparent at the first sight. Our equations are more clear in their meanings, and we call the method "locally-projected SCF MO (LP MO) for Molecular Interaction" in place of SCF MI, because "SCF for Molecular Interaction" is too general and does not specify the characteristics of the method and the molecular orbitals obtained. Besides we demonstrated that at the large basis set limit the LP SCF MI binding energies of water clusters converge to the incorrect values as Hamza et al [16] argue. We succeeded in analytically proving that the charge-transfer is indeed zero in LP SCF MI under Mulliken population analysis. To remove this difficulty, we examined the contribution from the charge-transfer terms to the binding energy, using the second order perturbation expansion[18]. For the perturbation expansion we need a proper set of excited orbitals localized on each unit in the cluster, consistent with the locally projected occupied orbitals. Utilizing the projection operator technique again, we defined the locally projected excited orbitals. A few test calculations for the potential energy curves of water dimer, HF dimer and M^+Ar_2 ($M=Na$, and K) demonstrated that the charge-transfer terms indeed mend the ill-behavior of LP SCF MI curves; the curves become close to the corresponding counterpoise corrected SCF curves[18]. Recently using the projection operators, we derived the equations for the high-spin open shell restricted LP MO, applicable to a cluster having one open shell molecule (atom) surrounded by many

closed shell[19]. The resulting equations have different forms from those derived earlier by Gianinetti et al[20].

In the present work we extend the previous work of the closed shell clusters by adding the double excitations both of the dispersion type and intramolecular pair correlation type. The other types are also examined numerically.

II. METHODS

A. Locally projected molecular orbitals

We assume that a cluster consists of closed shell molecules; the interaction is not so strong. More correctly speaking, the electron delocalization among the constituent molecules is not significant. To describe the electronic wave function in the locally projected molecular orbital (LP MO) method, the MOs of a molecule in the cluster are expanded in terms of the basis set located on the molecule; the occupied molecular orbital coefficient matrix is blocked as

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_A & 0 & 0 \\ 0 & \mathbf{T}_B & 0 \\ 0 & 0 & \mathbf{T}_C \end{bmatrix} \quad (1)$$

for instance, for a trimer. The stationary condition for the matrix \mathbf{T}_A can be written [19]

$$\left(1 - \widehat{P}\right) \widehat{F} \left(1 - \widehat{P}_{\notin A}\right) \chi_A \mathbf{T}_A = 0 \quad (2)$$

where \widehat{F} is the Fock operator of the full cluster and χ_A is a set of basis functions located on molecule A . The projection operators \widehat{P} and $\widehat{P}_{\notin A}$ are those defined for all of the occupied MOs and for the occupied orbitals of the molecules other than molecule A , respectively. The MO matrix \mathbf{T}_A can be obtained either by transforming (2) to a matrix general eigenvalue problem[18] or by defining a single Fock operator[19], which includes the stationary condition(2). The full MO coefficient matrix \mathbf{T} is solved self-consistently. One of the characteristics of the LP MO is the non-orthogonality among the occupied orbitals; it is inherent to the local expansion of MOs.

As is mentioned in Introduction, the single determinant description with the LP MO underestimates the binding energy. To go beyond the simplest description of the wave function, the excited (external) orbitals have to be properly determined; they should also

be defined locally. In the previous paper [18], a procedure to determine the local excited orbitals is described; the eigenvalue problem,

$$\left(\mathbf{1} - \widehat{P}\right) \chi_A \mathbf{t}_{A_k}^{ex} = \chi_A \mathbf{t}_{A_k}^{ex} \eta_k \equiv \varphi_{A_k}^{ex} \eta_k, \quad (3)$$

is solved. If $\eta_k = 1$, the excited orbital $\varphi_{A_k}^{ex}$ satisfies the orthogonality to all of the occupied orbitals under the restriction of the local expansion. But, to have $(N_A^{basis} - M_A^{occ})$ excited orbitals (N_A^{basis} and M_A^{occ} being the total numbers of the basis set and the occupied orbitals of molecule A , respectively), all of which are orthogonal to all of the occupied orbitals, a few of them with $\eta_k < 1$ have to be delocalized over the other molecules.[18] Thus, the excited orbitals are grouped to purely local (le) and partially delocalized (pd) excited orbitals. In the present numerical study, a threshold $T_h^\eta = 0.99999$ is chosen; if the eigenvalue η_k is smaller than T_h^η , the orbital $\varphi_{A_k}^{ex}$ is delocalized as

$$\phi_{A_k}^{pd-ex} = N_{A_k} \left(\mathbf{1} - \widehat{P}\right) \varphi_{A_k}^{ex} \quad (4)$$

where N_{A_k} is a normalization constant.

The overlap and Fock matrices in terms of these orbitals are diagonal dominant with non-zero off-diagonal elements; they are sparse matrices of the following property, for the dimer AB ,

$$\mathcal{M}^{full} = \begin{array}{c|cccccc} & A_{oc} & A_{le} & A_{pd} & B_{oc} & B_{le} & B_{pd} \\ \hline A_{oc} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathcal{M}_{oc,oc}^{AB} & \mathbf{0} & \mathbf{0} \\ A_{le} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathcal{M}_{le,oc}^{AB} & \mathcal{M}_{le,le}^{AB} & \mathcal{M}_{le,pd}^{AB} \\ A_{pd} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathcal{M}_{pd,oc}^{AB} & \mathcal{M}_{pd,le}^{AB} & \mathcal{M}_{pd,pd}^{AB} \\ \hline B_{oc} & \mathcal{M}_{oc,oc}^{BA} & \mathcal{M}_{oc,le}^{BA} & \mathcal{M}_{oc,pd}^{BA} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ B_{le} & \mathbf{0} & \mathcal{M}_{le,le}^{BA} & \mathcal{M}_{le,pd}^{BA} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ B_{pd} & \mathbf{0} & \mathcal{M}_{pd,le}^{BA} & \mathcal{M}_{pd,pd}^{BA} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{array} \quad (5)$$

$$\mathbf{F} = \begin{array}{c|cccccc} & A_{oc} & A_{le} & A_{pd} & B_{oc} & B_{le} & B_{pd} \\ \hline A_{oc} & \Lambda_{occ}^A & 0 & \mathbf{F}_{oc,pd}^A & \mathbf{F}_{oc,oc}^{AB} & \mathbf{F}_{oc,le}^{AB} & \mathbf{F}_{oc,pd}^{AB} \\ A_{le} & 0 & \Lambda_{le}^A & \mathbf{F}_{le,pd}^A & \mathbf{F}_{le,oc}^{AB} & \mathbf{F}_{le,le}^{AB} & \mathbf{F}_{le,pd}^{AB} \\ A_{pd} & \mathbf{F}_{pd,oc}^A & \mathbf{F}_{pd,le}^A & \Lambda_{pd}^A & \mathbf{F}_{pd,oc}^{AB} & \mathbf{F}_{pd,le}^{AB} & \mathbf{F}_{pd,pd}^{AB} \\ \hline B_{oc} & \mathbf{F}_{oc,oc}^{BA} & \mathbf{F}_{oc,le}^{BA} & \mathbf{F}_{oc,pd}^{BA} & \Lambda_{occ}^B & 0 & \mathbf{F}_{oc,pd}^B \\ B_{le} & \mathbf{F}_{le,oc}^{BA} & \mathbf{F}_{le,le}^{BA} & \mathbf{F}_{le,pd}^{BA} & 0 & \Lambda_{le}^B & \mathbf{F}_{le,pd}^B \\ B_{pd} & \mathbf{F}_{pd,oc}^{BA} & \mathbf{F}_{pd,le}^{BA} & \mathbf{F}_{pd,pd}^{BA} & \mathbf{F}_{pd,oc}^B & \mathbf{F}_{pd,le}^B & \Lambda_{pd}^B \end{array} \quad (6)$$

where \mathbf{A} is a diagonal matrix. Although non-zero parts of both matrices are expected to be small for most of the clusters targeted for the studies, they have to be taken into account in the perturbation expansion theory (or the other many-determinant theories).

B. 1st order wave function and 2nd order energy

In this paper we use the Møller-Plesset type of the partition of Hamiltonian \hat{H} , although the LP MOs are not canonical as in (6). The perturbation expansion with non-canonical MOs has been known with the local MP2 method [10]. The difference from the Local MP2 is that in the present case even the occupied orbitals are neither canonical nor orthogonal. Therefore, in evaluating the matrix elements, the contribution from the normalization factor of Slater determinants has to be properly handled. A few examples of the matrix elements are given in Appendix.

In the previous paper[18], where only the single excitations are taken into account, the diagonal approximation for $\Xi + \mathbf{W}$ defined in (A1) of Appendix is used. The approximation causes a slight numerical differences, but as is shown below, the key conclusion that the charge-transfer terms remove the deficiency of the LP SCF MO MI remains correct.

C. The type of single and double excitations (orbital replacements)

The excited orbitals are also localized on each constituent molecule. Although some of them are slightly delocalized over the basis sets of the other molecules in the cluster, their belongings are easily identified. Thus, the electron configurations formed by the single and double excitations (orbital replacements) from the reference closed shell determinant are classified, for instance, for a tetramer $ABCD$,

Type of excitations	Form	abbrev.	orbital replacement
Local	A^S	LE	$a_A \rightarrow r_A$
Charge-Transfer (CT)	$(A^+B^-)^S$	CT	$a_A \rightarrow r_B$
Intra-molecular pair	A^{**}	Pair(P)	$(a_A, b_A \rightarrow r_A, s_A)^*$
Dispersion	$A^S B^S$	Disp(D)	$(a_A \rightarrow r_A)^S \& (b_B \rightarrow s_B)^S$
Dispersion (triplet)	$A^T B^T$	DispEx(Dex)	$(a_A \rightarrow r_A)^T \& (b_B \rightarrow s_B)^T$
CT&Polarization	$(A^+B^-)^S C^S$	CT&P	$(a_A \rightarrow r_B)^S \& (b_C \rightarrow s_C)^S$
CT&Polarization (triplet)	$(A^+B^-)^T C^T$	CT&PEX	$(a_A \rightarrow r_B)^T \& (b_C \rightarrow s_C)^T$
CT&anion-reorganization	$(A^+B^-)^* B^*$	CT&AR	$(a_A \rightarrow r_B)^* \& (b_B \rightarrow s_B)^*$
CT&cation-reorganization	$(A^+B^-)^* A^*$	CT&CR	$(a_A \rightarrow r_B)^* \& (b_A \rightarrow s_A)^*$
Double CT	$A^{2+} B^{2-}$	DCT	$(a_A, b_A \rightarrow r_B, s_B)^*$
Two-direction CT	$A^{+2} B^- C^-$	TDCT	$(a_A \rightarrow r_B)^* \& (b_A \rightarrow s_C)^*$
Two-direction CT	$A^{-2} B^+ C^+$	TDCT	$(a_B \rightarrow r_A)^* \& (b_C \rightarrow s_A)^*$
A pair of CT	$A^+ B^- C^+ D^-$	PCT	$(a_A \rightarrow r_B)^* \& (b_C \rightarrow s_D)^*$

where the super script $S(T)$ stands for the singlet (triplet) coupled excited configuration, and if the spin coupling is not specified, the script $*$ is added or no script is given. In the table, a_A and b_A stand for the occupied orbitals of molecule A , and r_B and s_B for the excited orbitals of molecule B . The list includes all of the possible types of the single and double excitations. It should, however, be emphasized that only some of the types are retained in the perturbation expansion. If all terms are included in the calculations a part of the basis set superposition errors (BSSE) re-emerges. The most illustrated example is the full configuration interaction (CI). Suppose we use the full CI in evaluating the binding energy of a dimer AB with a basis set. Without the counterpoise correction, no matter what kind molecular orbitals are used, the calculated binding energy contains the BSSE. The LP SCF MO energy is BSSE-free, but once the other electron configurations are added to the ground state wave function by means of MP2, CI, or other many-determinant methods, the possibility to induce the BSSE arises. In particular, the terms involving the charge-transfer (or electron delocalization) may contribute to "improving" the molecular orbitals by borrowing the basis set on the other molecules, and thus the BSSE may arise. Therefore, care should be taken to select the types of the excitations in the calculations, both theoretically and numerically.

In the present study, from the above list, the first five types (LE, CT, Pair, Disp and

DispEx) are examined. To demonstrate the effects from the other types, the CT&AR and CT&CR) types are also examined. The pair type terms (Pair) are essential in correlating the energy with the sum of the MP2 energy of the isolated molecules. The dispersion term (Disp), resulting from the simultaneous singlet excitations $(a_A \rightarrow r_A)^S$ & $(b_B \rightarrow s_B)^S$, is essential to obtain the van der Waals interaction. It is worth examining the contribution from the triplet-triplet excitation $(a_A \rightarrow r_A)^T$ & $(b_B \rightarrow s_B)^T$. The contribution from these terms is almost equivalent to the dispersion-exchange term of SAPT.[8] To separate the dispersion and dispersion-exchange terms, the CSF form of the many-electron basis is suitable. In the LP MO theory, the classic electrostatic and polarization interaction are within the SCF theory. Except for the CT terms no electron delocalization between the molecules is apparently included. Note that through W_{KL} (A2) in Appendix, various types of the excitations are weakly coupled.

In the matrix elements in Appendix, the sparse nature of MO overlap matrix \mathcal{M} in (5) and Fock matrix in (6) is not used in the present code.

III. TEST CALCULATIONS AND DISCUSSION

The present code is based on the spin-adapted configuration state function (CSF). To test the numerical results, the determinant based code was also written. The CSF version is useful in differentiating the ordinary dispersion term (singlet-singlet simultaneous excitation, Disp) from the spin-exchanged dispersion (triplet-triplet excitation, DispEx). The numerical results shown in this paper are all calculated with the CSF-based version.

The potential energy curves of a linear form of water dimer $(\text{H}_2\text{O})_2$ and HF dimer $(\text{HF})_2$ were calculated; the geometric parameters other than O-O or F-F distance are optimized with MP2/aug-cc-pDVZ without the CP correction. The reported MP2 binding energy estimated by the complete basis set limit extrapolation is 20.85 kJ/mol for water dimer and 18.42 kJ/mol for HF dimer[22]. The aug-cc-pVDZ basis set used in the present is poor for evaluating the binding energy. All calculations in this report were carried out without the frozen core approximation. The results are shown in Figures 1-4.

Figure 1 compares the potential energy curves of the SCF level of approximations with the aug-cc-pVDZ basis set. The curve of LP SCF both for $(\text{H}_2\text{O})_2$ and $(\text{HF})_2$ is very different from the others; at the bottom of the curve, the binding energy is by about 2.5 kJ/mol smaller

than the others, and the shape is also slightly different. The deepest curve is the one for the uncorrected SCF. The counterpoise correction (CP) slightly reduces the binding energy with this basis set for both molecules. By adding the charge-transfer (CT) and the locally excitation (LE) terms in the second order perturbation (MP2), the curves (S in the figure) become very close to the curve of CP SCF; they are almost indistinguishable to each other (Note that the energy unit is kJ/mol, not kcal/mol). The LE does not contribute much to the binding energy. It is because the direct matrix element $\langle \Phi(a_A \rightarrow r_A) | \hat{V} | \Phi_0 \rangle = \sqrt{2} \langle r | \hat{F} | \underline{a} \rangle$ in (A10) of Appendix is nearly zero (note that $\langle r_A | \hat{F} | a_A \rangle = 0$, but $\langle r_A | \hat{F} | \underline{a}_A \rangle$ is not exactly zero. See Appendix). The required computation time for calculating the CT MP2 is very small; less than one cycle of the SCF calculation.

For the small basis set cc-pVDZ without diffuse functions, the CT contribution overshoots the CP corrected value as shown in Fig.2, and the curve (LP SCF + S) both for $(\text{H}_2\text{O})_2$ and $(\text{HF})_2$ are close to the uncorrected SCF curve, indicating that the CT terms almost "recover" the BSSE. As described in Subsection IIA, there are two types of the local excited orbitals, purely local (*le*) and partially delocalized (*pd*). The curve (LP SCF + S in the figure) is calculated using both of the excited orbitals. By removing the partially delocalized orbitals from the single excitations, the curve, LP SCF + S[Deloc excld], is obtained. The curve lies between the CP SCF and LP SCF curves in both dimers. Thus, it is the partially delocalized orbitals which induce the BSSE. The number of purely local orbitals, M_A^{le} , is varied with the bond lengths O-O and F-F. For $(\text{H}_2\text{O})_2$ the total number of excited orbitals for each monomer is 20. For $R_{O-O} \geq 6\text{\AA}$, all of them are purely local, and M_A^{le} eventually becomes 16 for $3.4 \geq R_{O-O} \geq 2.5\text{\AA}$. For $(\text{HF})_2$, for $R_{F-F} \geq 5\text{\AA}$, $M_A^{le} = 15$, which implies that all of the excited orbitals can be defined locally. It is reduced to 11 for $R_{F-F} \leq 2.9\text{\AA}$. The comparison of the two curves (LP SCF +S and LP SCF + S[Deloc excld]) demonstrates that the small number of partially delocalized orbitals is responsible to inducing the BSSE. In the present test calculations, the threshold value T_h^η for defining the purely local excited orbitals is 0.99999, which might be so severe.

In the enlarged plots of Fig. 1(b), the curve, LP SCF + S[Deloc excld], is added for comparison. In these bond lengths, $M_A^{le} = 24 \sim 25$ (the total number of the excited orbitals is 29). As is seen in the figure, by excluding the partially delocalized orbitals (*pd*) for aug-cc-pVDZ, the curve does shift upward, but only by 0.5kJ/mol, which is compared with the large change seen in Fig. 2(b). The contribution from the *pd* orbitals is much smaller in the

basis set with the diffuse functions than without them. If the basis set is sufficiently large to describe a molecule, the use of the basis set on the other molecule is less necessary. In the other word, this difference indicates that for smaller basis sets, the electron delocalization over molecules is over-emphasized in the super molecule calculations and induces the BSSE.

Figure 3 compares the potential energy curves of the MP2 level of approximations with the aug-cc-pVDZ. As is known, the counterpoise correction for MP2 is large even for aug-cc-pVDZ, as is seen in Fig. 3. The curve obtained by adding the intra-molecular pair type (Pair) and dispersion type (Disp) is far less stable than the curve of the CP corrected MP2. By further adding the dispersion-exchange (triplet-triplet excitation, Dex) terms, the curve becomes closer to that of CP MP2. In particular, at the longer O-O and F-F distance, the two curves (CP MP2 and LP MO MP2 with DispEx) practically coincide to each other in both dimers. The contribution from the dispersion-exchange term has been discussed in the SAPT[8]. Figure 3 also demonstrates the importance of exchange type dispersion term, which is rather discouraging in practical purposes. It implies that the applicability of the dispersion energy evaluated by the frequency-dependent polarizability is limited only at the distances far longer than near the bottom of the potential energy curve, unless the dispersion-exchange terms are included. This is rather obvious since near the bottom of the potential energy curves for the hydrogen bonding system the orbital overlap are significant and therefore the spin-coupling within each molecule is meaningless.

Figure 3 shows the curve calculated including the terms CT&AR and CT&CR. It is apparent that the inclusion of these terms causes the large error. In the local MP2 methods, the similar terms are also excluded [10], [11]. A kind of the basis set superposition error cannot be avoided once the orbital relaxation of the occupied orbitals through the charge-transfer terms is included. More careful analyses are needed, in particular, on the contribution from the off-diagonal matrix elements (A2) and on the basis set dependence.

Figure 4 is a comparison among the methods used with the cc-pVDZ basis set. The substantially large CP correction is found both for $(\text{H}_2\text{O})_2$ and $(\text{HF})_2$; nearly a half of the binding energy at the bottom of the curves. The curves of the LP MO MP2 of both levels of approximation without and with the dispersion-exchange (DispEx) terms are closer to the CP corrected MP2 than to the uncorrected MP2. But because of the over-correction in the single excitation terms, the curves lie below the CP MP2 curve. By removing the contribution from the partially delocalized excited orbitals in the single excitations, the

curves move to the upper than the CP corrected curve as in Fig.2. The corresponding curves with and without the partially delocalize excited orbitals bracket the CP MP2 curve. If one or two delocalized excited orbitals are included by relaxing the threshold value T_h^η , the curve (S[deloc exclud]+P+D+Dex) becomes closer to the CP MP2 for both dimers.

IV. CONCLUSION

More careful examination for restricting the types of the excitations and for removing some of the partially delocalized excited orbitals is required both numerically and theoretically. The present test calculations suggest some promising pragmatic procedures to reduce the BSSE. For instance, with the basis set having the diffuse functions (aug sets or ++ sets), the LP MO + S calculations can be used without the CP correction. This approach is useful for clusters consisting of a large number of molecules. The test calculations for $\text{Mg}(\text{H}_2\text{O})_6$, which are recently found in the molecular beam experiments by Fuke's group[23], proved that the LP MO + S calculation required less cpu time than the supermolecule calculation, and its binding energy is practically equal to the CP corrected energy; the CP correction required a large cpu time by more than one order of magnitude than the LP MO +S calculation[7]. For smaller basis sets, with an appropriate threshold value T_h^η , the LP MO +S[deloc exclud] or LP MO +S[deloc exclud]+P+D+Dex calculations can replace the CP correction procedure. These procedures are useful in the ab initio Monte Carlo simulations of the molecular cluster.

It is worth examining a part of the higher order perturbation expansion, because in the LP MO MP2 the molecular orbitals are not canonical.

Acknowledgements

The work is partially supported by the Grants-in-Aid for Science Research (No. 1454079 and 17550012) of Japan Society for Promotion of Science. TN thanks Prof. Aoyagi of Kyusyu University and the late Prof. K. Saito of Hiroshima University for continuous encouragement. SI is appointed as a special professor under the Special Coordination Funds for Promoting Science and Technology from MEXT, Japan. He thanks Prof. M. Aida of Hiroshima University for giving him an opportunity to return to the research. A part of

computations was carried out at the Research Center for Computational Science(RCCS), Okazaki Research Facilities, National Institutes of Natural Sciences(NINS).

Appendix

The Møller-Plesset type of the partition of Hamiltonian \widehat{H} is used

$$\begin{aligned}\widehat{H} &= \widehat{H}_0 + \widehat{V} \\ \widehat{H}_0 &\equiv \sum_{b,c}^{occ} \widehat{a}_b^\dagger \langle b | \widehat{F} | c \rangle \widehat{a}_c + \sum_{r,s}^{ex} \widehat{a}_r^\dagger \langle r | \widehat{F} | s \rangle \widehat{a}_s\end{aligned}$$

although the LP MOs are not canonical as in (6). Here, \widehat{a}_b^\dagger and \widehat{a}_c are the creation and annihilation operators for the occupied orbitals, and \widehat{a}_r^\dagger and \widehat{a}_s are the corresponding operators of the excited orbitals. This choice is ensure $\langle \Phi_0 | \widehat{H}_0 | \Phi_K \rangle = 0$ [21]. For the first order coefficient vector $\mathbf{A}^{(1)}$, the equation to be solved is

$$(\mathbf{\Xi} + \mathbf{W}) \mathbf{A}^{(1)} = -\mathbf{v} ,$$

where $\mathbf{\Xi}$ is a diagonal matrix,

$$[\mathbf{\Xi}]_{K,L} = \delta_{K,L} \left\{ \langle \Phi_K | \widehat{H}_0 | \Phi_K \rangle - E_0^0 \langle \Phi_K | \Phi_K \rangle \right\} , \quad (\text{A1})$$

which is not a simple difference of the "orbital energy" because of the non-orthogonality and of the non-canonical nature of the orbitals. The many-electron basis functions Φ_K require a proper normalization factor. The off-diagonal matrix elements are defined as

$$[\mathbf{W}]_{K,L} = (1 - \delta_{K,L}) \left\{ \langle \Phi_K | \widehat{H}_0 | \Phi_L \rangle - E_0^0 \langle \Phi_K | \Phi_L \rangle \right\} , \quad (\text{A2})$$

in which the term $E_0^0 \langle \Phi_K | \Phi_L \rangle$ is proved to be canceled out with a part of the first term $\langle \Phi_K | \widehat{H}_0 | \Phi_L \rangle$. The inhomogeneity term \mathbf{v} is defined as

$$[\mathbf{v}]_L = \langle \Phi_L | \widehat{V} | \Phi_0 \rangle . \quad (\text{A3})$$

In evaluating the matrix elements, the required MO integrals involving the occupied MOs $|c\rangle$ are calculated using the biorthogonal transformed MO $|\underline{a}\rangle$,

$$|\underline{a}\rangle = \frac{1}{\sqrt{\mathcal{R}_{aa}}} \sum_c^{occ} |c\rangle \mathcal{R}_{ca}, \quad (\text{A4})$$

where \mathcal{R} is the inverse of MO overlap matrix \mathcal{M} within the occupied orbitals. Only some of the representatives are given below.

A. The diagonal matrix

$$\Xi(a-r) \equiv \langle \Phi_{a;r} | \hat{H}_0 | \Phi_{a;r} \rangle - E_0^0 \langle \Phi_{a;r} | \Phi_{a;r} \rangle = F_{rr} - \langle \underline{a} | \hat{F} | \underline{a} \rangle \quad (\text{A5})$$

For $a < b, r < s$

$$\begin{aligned} \Xi((a \rightarrow r)^S (b-s)^S) &= R(ab \rightarrow rs; S)^2 \\ &\times \left[\begin{array}{c} (F_{ss} + F_{rr}) \left(\mathbf{1} - \frac{1}{2} \mathcal{R}_{ab}^2 \right) + 2\mathcal{M}_{sr} F_{rs} \left(\mathcal{R}_{ab}^2 - \frac{1}{2} \right) \\ - \left(\langle \underline{a} | h | \underline{a} \rangle + \langle \underline{b} | f | \underline{b} \rangle \right) \left(1 - \frac{1}{2} \mathcal{M}_{rs}^2 \right) + 2 \langle \underline{b} | f | \underline{a} \rangle \mathcal{R}_{ab} \left(\frac{1}{2} - \mathcal{M}_{rs}^2 \right) \end{array} \right] \end{aligned} \quad (\text{A6})$$

$$\begin{aligned} \Xi((a \rightarrow r)^T (b-s)^T) &= R(ab \rightarrow rs; T)^2 \\ &\times \left[\begin{array}{c} (F_{ss} + F_{rr}) \left\{ 1 + \frac{1}{2} \mathcal{R}_{ab}^2 \right\} + 2F_{rs} \mathcal{M}_{sr} \left\{ \frac{1}{2} \mathbf{1} + \mathcal{R}_{ab}^2 \right\} \\ - \left\{ \left(\langle \underline{a} | h | \underline{a} \rangle + \langle \underline{b} | h | \underline{b} \rangle \right) \left(1 + \frac{1}{2} \mathcal{M}_{sr}^2 \right) + 2 \langle \underline{b} | h | \underline{a} \rangle \mathcal{R}_{ab} \left(\frac{1}{2} + \mathcal{M}_{sr}^2 \right) \right\} \end{array} \right] \end{aligned} \quad (\text{A7})$$

where

$$\begin{aligned} R(ab \rightarrow rs; S) &= \left[\left(1 - \frac{\mathcal{M}_{rs}^2}{2} \right) + \mathcal{R}_{ab}^2 \left(\mathcal{M}_{rs}^2 - \frac{1}{2} \right) \right]^{-\frac{1}{2}} \\ R(ab \rightarrow rs; T) &= \left[\left(1 + \frac{\mathcal{M}_{rs}^2}{2} \right) + \mathcal{R}_{ab}^2 \left(\mathcal{M}_{rs}^2 + \frac{1}{2} \right) \right]^{-\frac{1}{2}} \end{aligned}$$

B. The off-diagonal matrix elements of the homogeneous term

$$[\mathbf{W}]_{K,L} \equiv W[K|L]$$

$$W[a-r|b-s] = \mathcal{M}_{rs} \langle \underline{b} | F | \underline{a} \rangle \quad (\text{A8})$$

For $c < b, t < u$

$$\begin{aligned} W[(a \rightarrow r)^S (b-s)^S | (c \rightarrow t)^S (d-u)^S] & \\ &= \frac{1}{2} R(ab \rightarrow rs; S) R(cd \rightarrow tu; S) \\ &\times \left[\begin{array}{c} \mathcal{R}_{db} \mathcal{R}_{ca} \{ 2F_{su} \mathcal{M}_{rt} + 2F_{rt} \mathcal{M}_{su} - F_{ru} \mathcal{M}_{st} - F_{st} \mathcal{M}_{ru} \} \\ + \mathcal{R}_{cb} \mathcal{R}_{da} \{ 2F_{ru} \mathcal{M}_{st} + 2F_{st} \mathcal{M}_{ru} - F_{su} \mathcal{M}_{rt} - F_{rt} \mathcal{M}_{su} \} \\ - \{ \mathcal{R}_{ca} F_{\underline{b}d} + \mathcal{R}_{db} F_{\underline{a}c} \} (2\mathcal{M}_{rt} \mathcal{M}_{su} - \mathcal{M}_{ru} \mathcal{M}_{st}) \\ - \{ \mathcal{R}_{cb} F_{\underline{a}d} + \mathcal{R}_{da} F_{\underline{b}c} \} (\mathcal{M}_{rt} \mathcal{M}_{su} - 2\mathcal{M}_{ru} \mathcal{M}_{st}) \end{array} \right] \end{aligned} \quad (\text{A9})$$

C. The coupling terms of the closed shell reference function Φ_0 with the excited CSFs.

$$\langle \Phi(a \rightarrow r) | \widehat{V} | \Phi_0 \rangle = \sqrt{2} \langle r | \widehat{F} | \underline{a} \rangle \quad (\text{A10})$$

For $a < b, r < s$

$$\langle (a \rightarrow r)^S (b - s)^S | \widehat{V} | \Phi_0 \rangle = R(ab; rs)^S (\{2 \langle rs | g | \underline{ab} \rangle - \langle sr | g | \underline{ab} \rangle\}) \quad (\text{A11})$$

$$\langle (a \rightarrow r)^T (b - s)^T | \widehat{V} | \Phi_0 \rangle = \sqrt{3} R(ab; rs)^T \langle sr | g | \underline{ab} \rangle \quad (\text{A12})$$

V. REFERENCES

-
- [1] Chem. Rev. Edited by B. Berharad, P. Hobza (2000)
 - [2] Chem. Rev. Edited by A. W. Castleman, P. Hobza (1994)
 - [3] F. B. van Duijneveldt, G. C. J. M. van Duijneveldt-van deRijdt, J. H. van Lenthe, Chem. Rev. 94, 1873 (1994)
 - [4] S. F. Boys, F. Bernardi, Mol. Phys. 19, 553 (1970)
 - [5] H. B. Jansen, P. Ross, Chem. Phys. Letters, 3, 140 (1969)
 - [6] G. Chalasinski, M. M. Szczesniak, Chem. Rev. 4227 (2000)
 - [7] S. Iwata, in preparation.
 - [8] B. Jeziorski, K. Szalewicz, "Handbook of Molecular Physics and Quantum Chemistry", Ed. by S. Wilson, John Wiley & Sons, London, chapter 9, 233 (2003)
 - [9] I. Mayer, P. Valiron, J. Chem. Phys. 109, 3360 (1998)
 - [10] P. Pulay, S. Saebo, Theor. chim. Acta, 69, 357 (1986)
 - [11] M. Schültz, G. Rauhut, H. -J. Werner, J. Phys. Chem. A102, 5997 (1998)
 - [12] M. Schültz, H. -J. Werner, R. Lindh, F. R. Manby, J. Chem. Phys. 121, 737 (2004)
 - [13] E. Gianinetti, M. Raimondi, E. Tornaghi, Int. J. Quantum, Chem. 60, 157 (1996)
 - [14] A. Famulari, M. Raimondi, M. Sironi, E. Gianinetti, Chem Phys. 232, 275 (1998)
 - [15] G. Calderoni, F. Cargnoni, A. Famulari, M. Raimondi, J. Phys. Chem. A106, 5521 (2002)

- [16] A. Hamza, Á. Vibók, G. J. Halász, I. Mayer, *Theor. Chem. Acc.* 107, 38 (2001)
- [17] T. Nagata, O. Takahashi, K. Saito, S. Iwata, *J. Chem. Phys.* 115, 3553 (2001)
- [18] T. Nagata, S. Iwata, *J. Chem. Phys.* 120, 3555 (2004). Paper I of this series.
- [19] S. Iwata, *J. Theor. Comp. Chem.* in press (2006)
- [20] E. Gianinetti, I. V. Vandoni, A. Famulari, M. Raimondi, *Adv. Quantum Chem.* 31, 251 (1999)
- [21] M. Head-Gordon, P. E. Maslen, C. A. White, *J. Chem. Phys.* 108, 616 (1998)
- [22] S. B. Huh, J. S. Lee, *Chem. Phys. Letters*, 369, 466 (2003)
- [23] N. Okai, H. Ishikawa, K. Fuke, a poster presentation (2P032) at Symposiums of Molecular Spectroscopies and Structures, Tokyo (2005).

Figure Caption

Figure 1. Comparison of the potential energy curves of water dimer and HF dimer in the SCF levels of approximation. The basis set used is aug-cc-pVDZ. The curves of {LP SCF +S} and {LP SCF +S[deloc exclud]} are those of the single excitations from LP SCF wave function with and without the partially delocalized excited orbitals. a) (H₂O)₂ b) (HF)₂.

Figure 2. Same with Figure 1 except that the basis set is cc-p-VDZ.. a) (H₂O)₂ b) (HF)₂.

Figure 3. Comparison of the potential energy curves of water dimer and HF dimer in the MP2 levels of approximation. The basis set used is aug-cc-pVDZ. The curves of {S, P, D}, {S, P, D, Dex}, and {S, P, D, Dex, CT&AR, CT&CR} are compared with those of the MP2 and CP corrected MP2. Here, S, P, D, Dex, CT&AR and CT&CR are the single, pair(PAIR), dispersion(Disp), dispersion-exchange(DispEx), anion reorganization, and cation reorganization types of the excitations defined in the table of text. a) (H₂O)₂ b) (HF)₂.

Figure 4. Same with Figure3 except that the basis set is cc-p-VDZ.

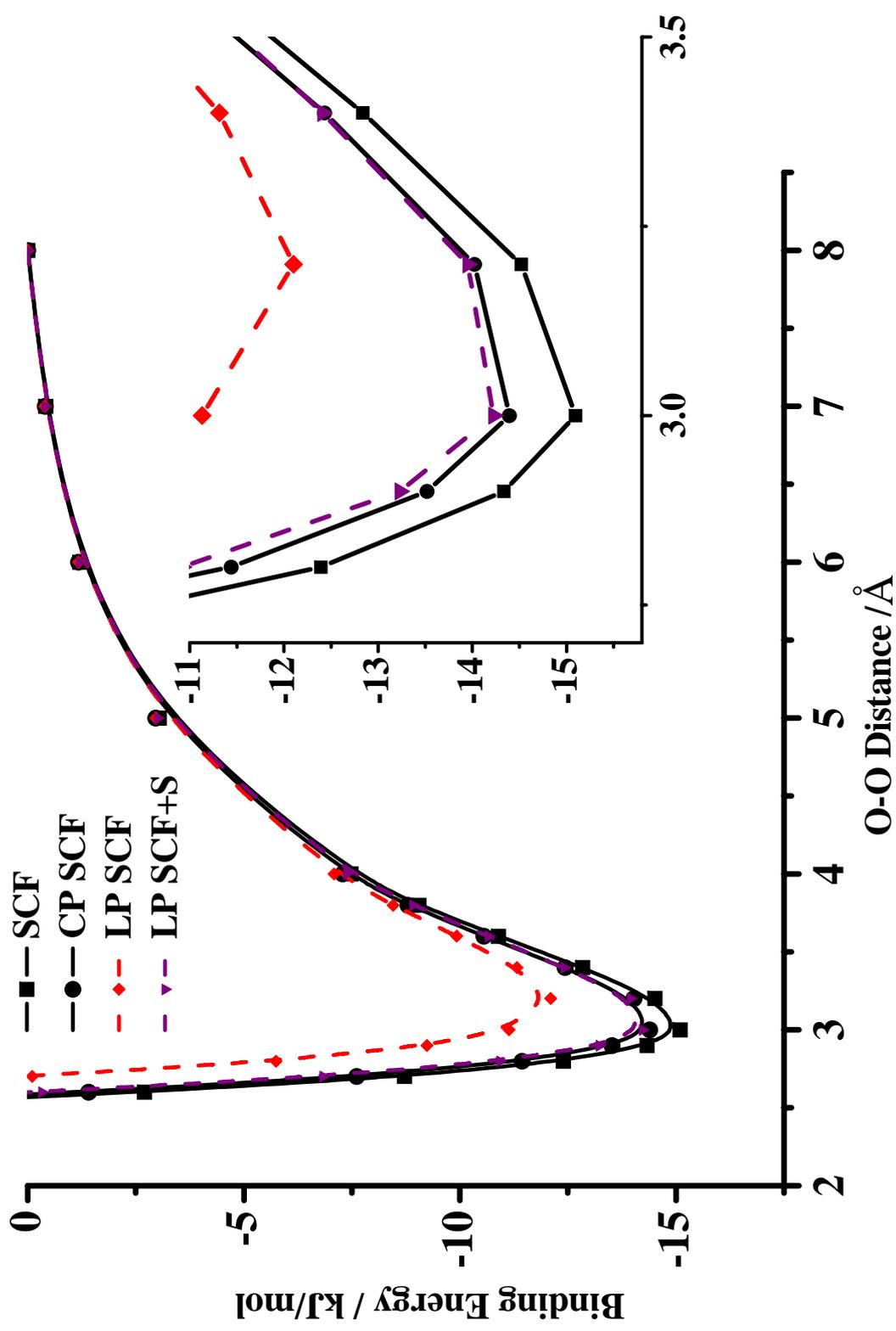


Fig. 1 a) I & N

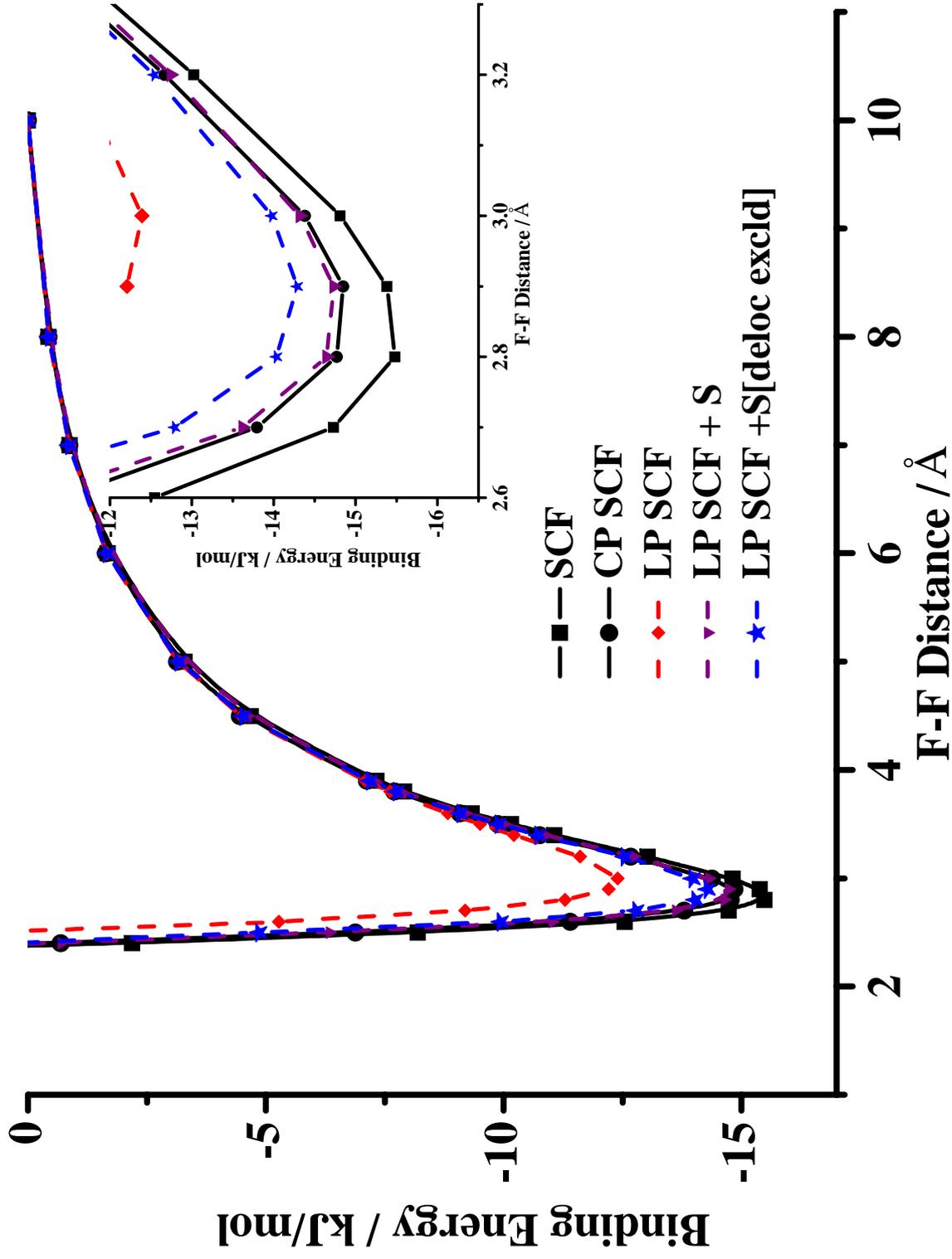


Fig. 1 b) I & N

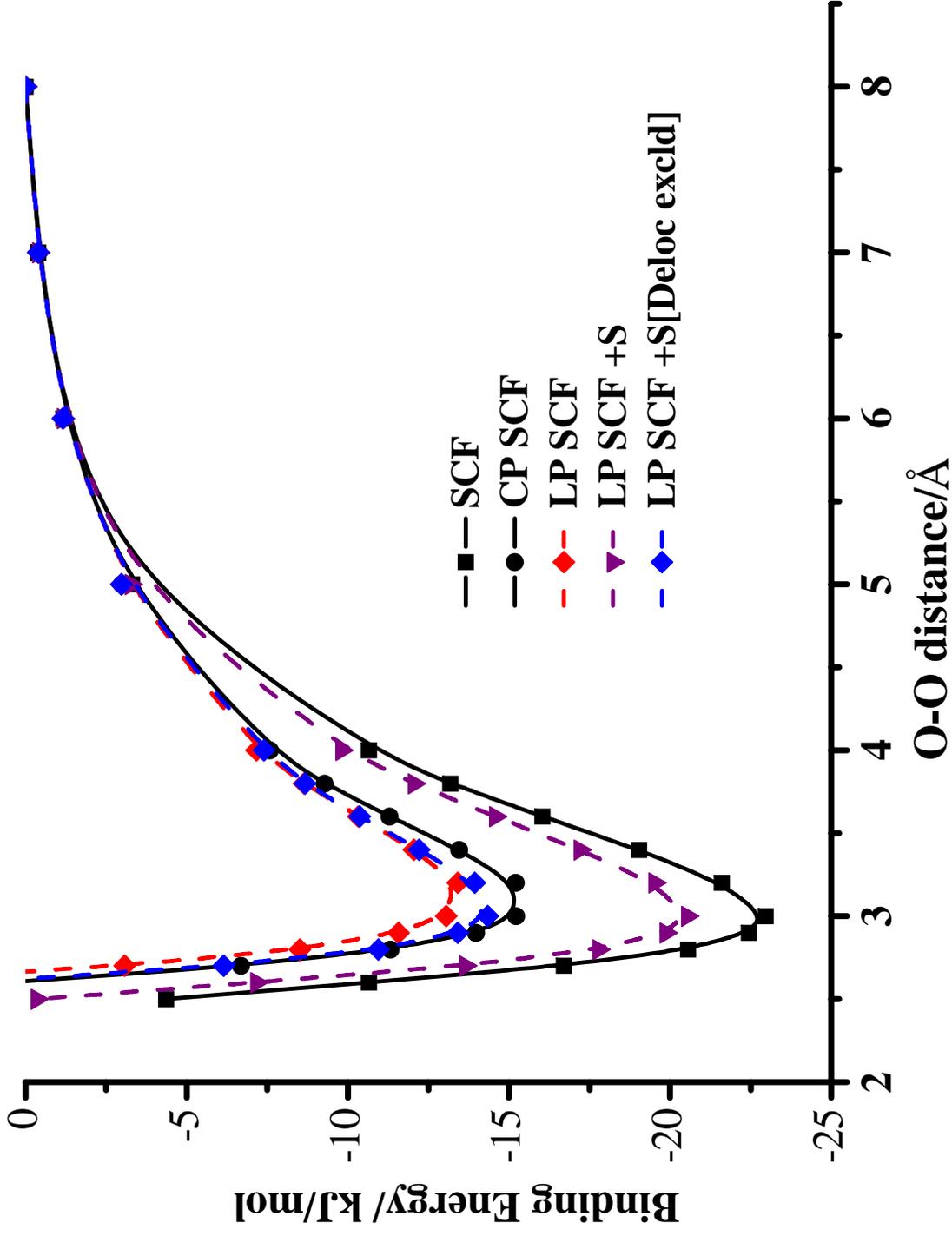


Fig. 2 a) I & N vdz (H₂O)₂

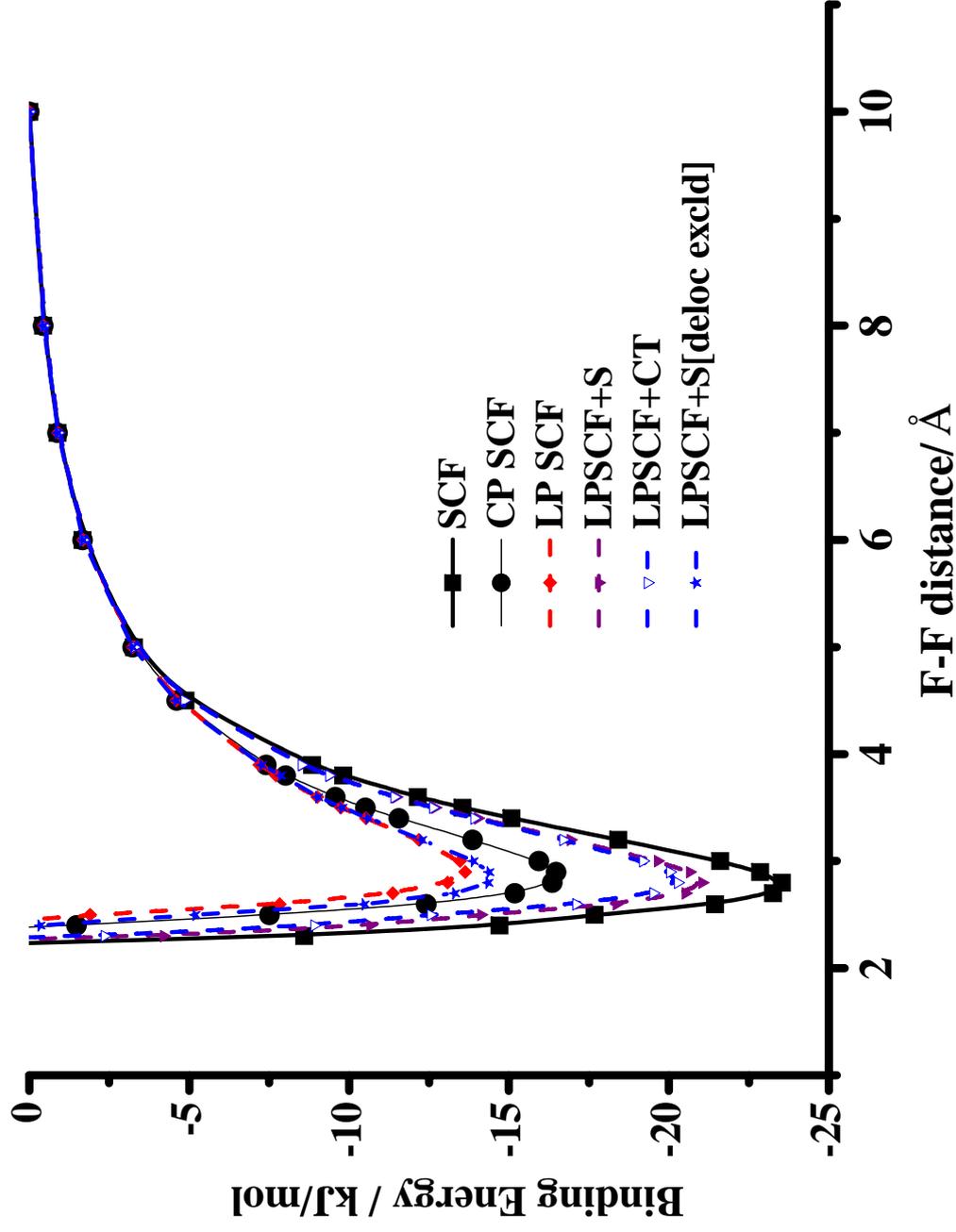


Fig. 2 b) I & N vdz (HF)2

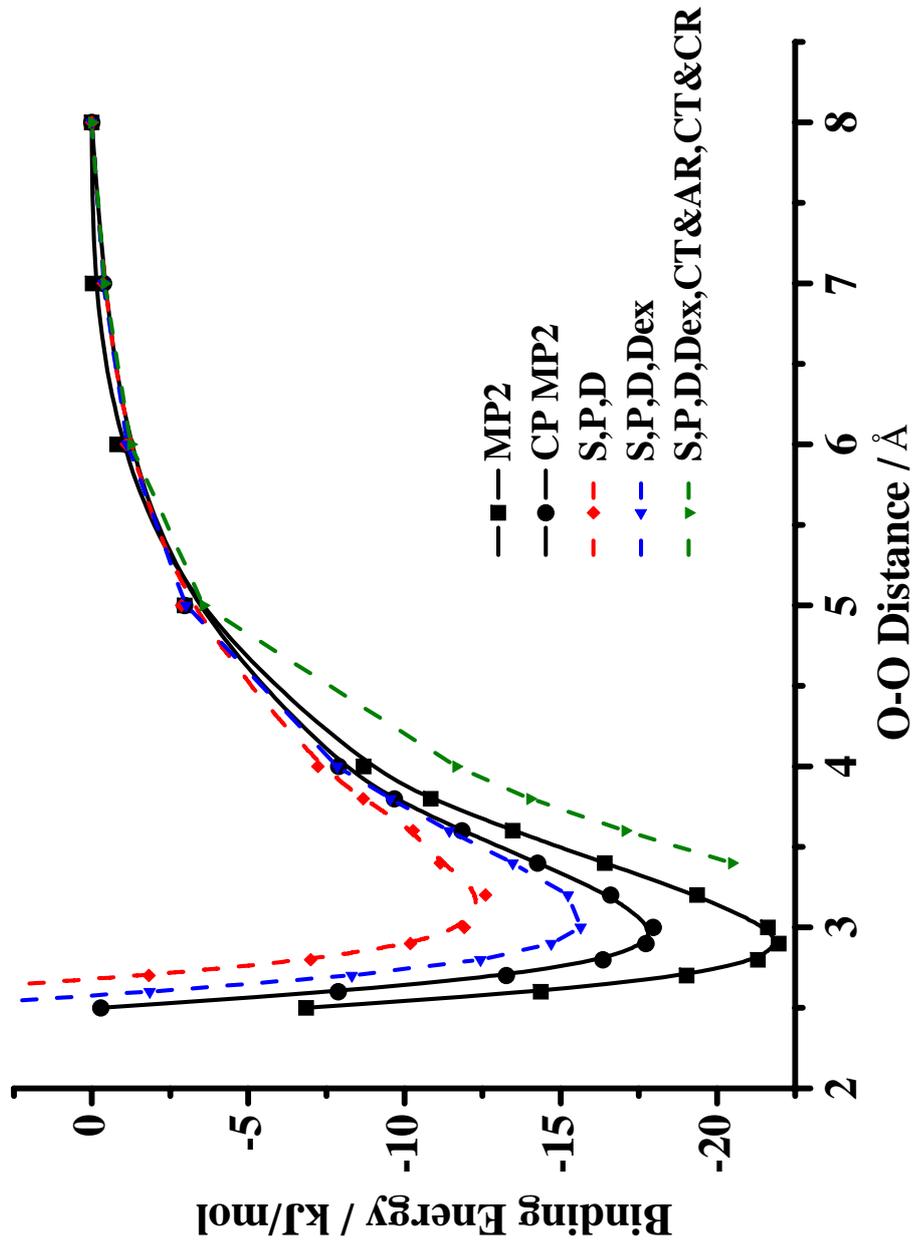


Fig.3 a) I & N

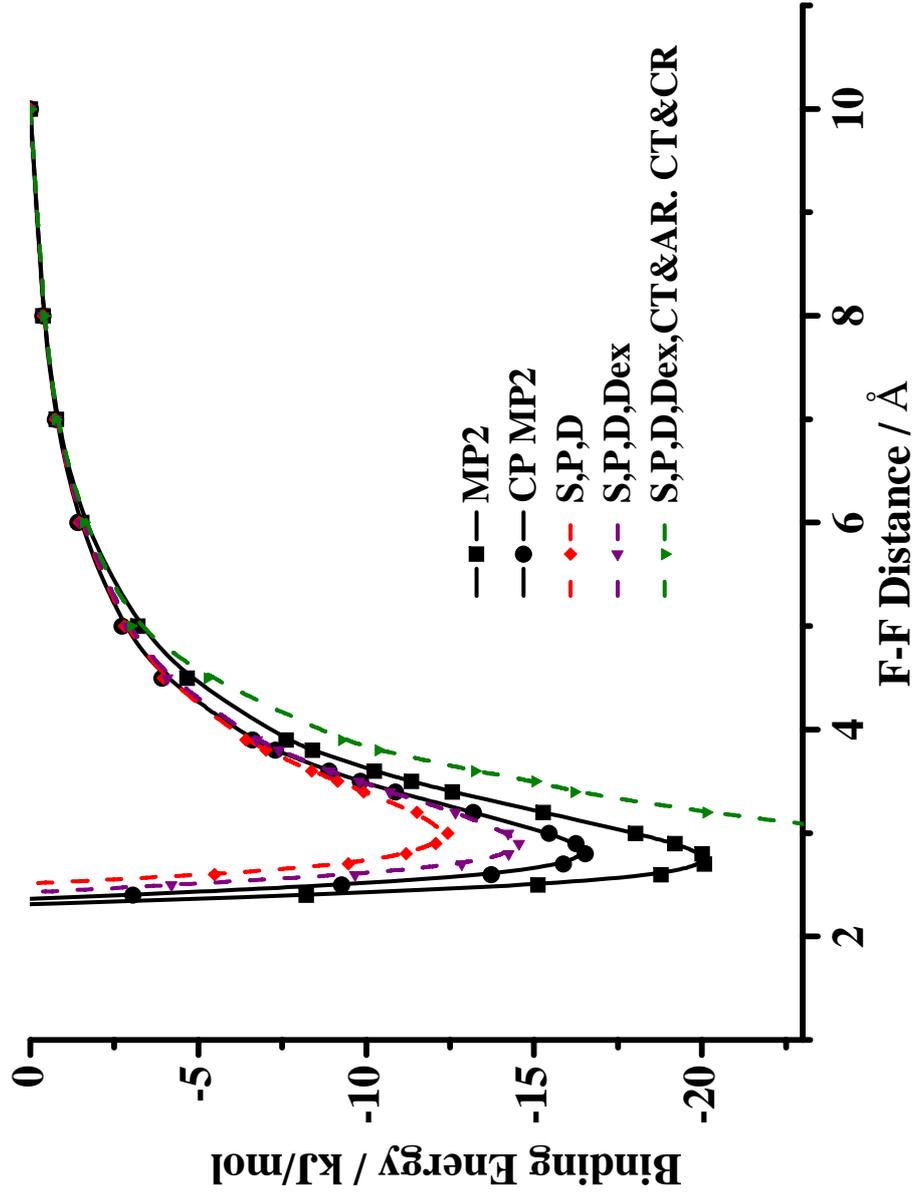


Fig. 3 b) I & N

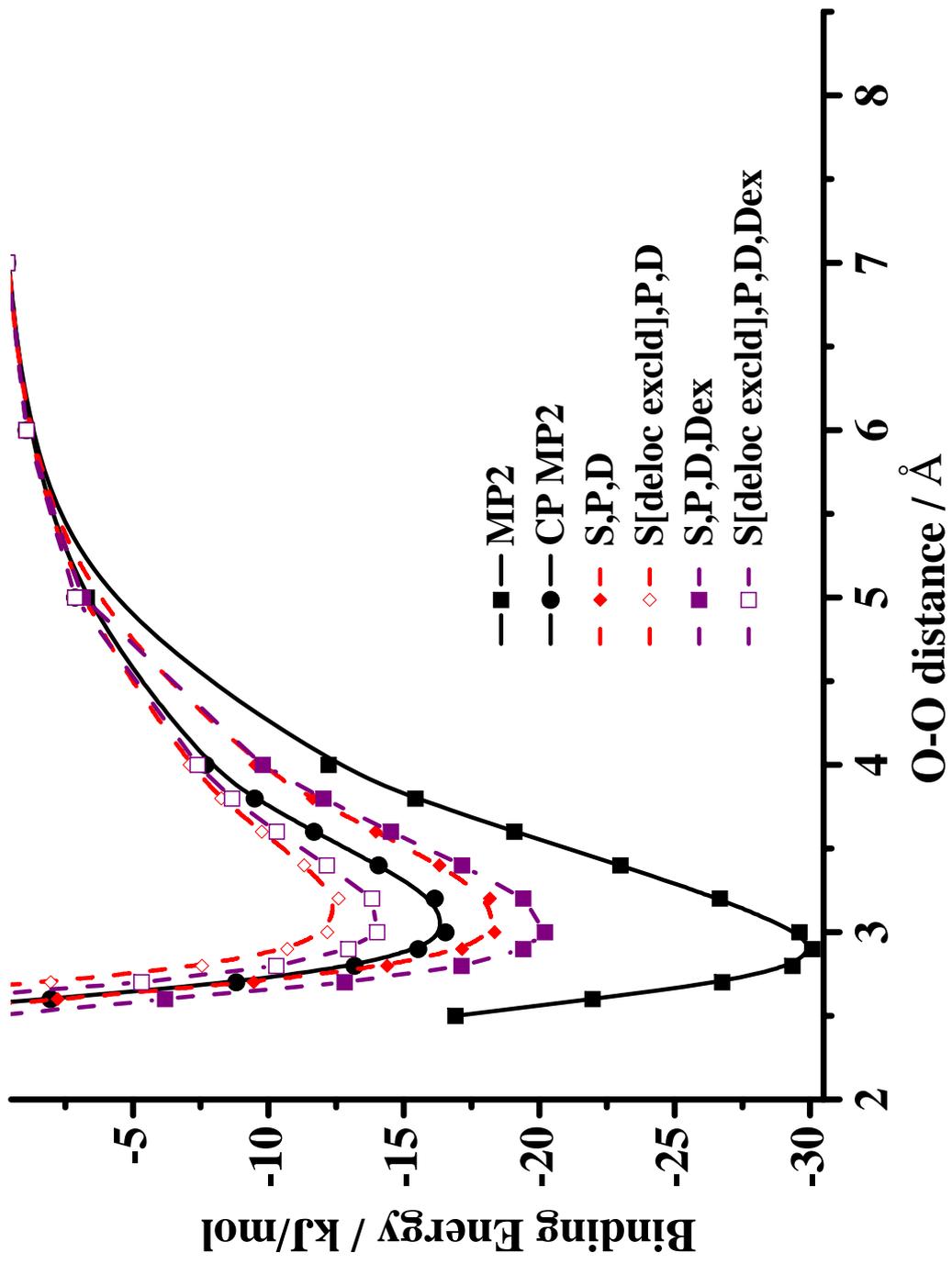


Fig. 4 a) I & N

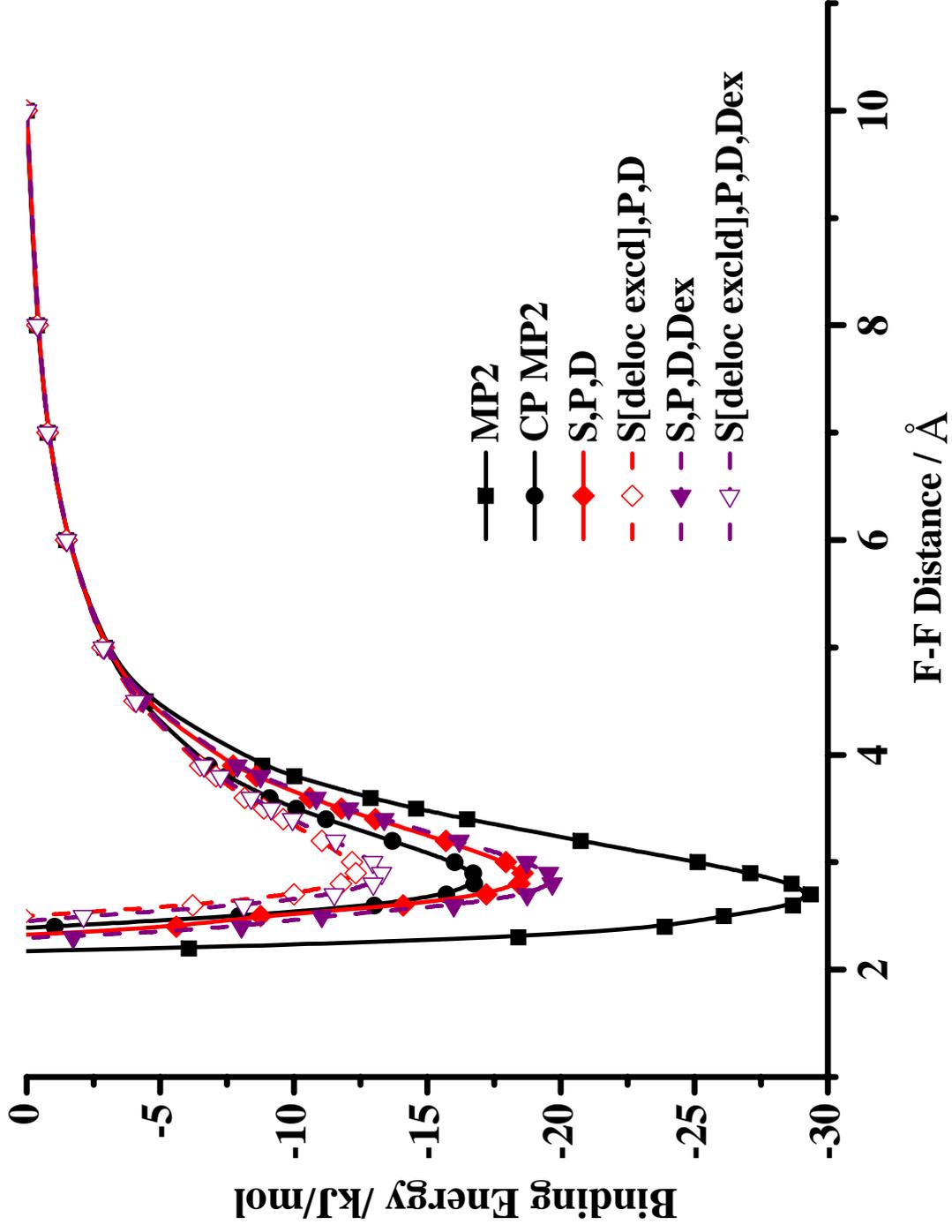


Fig.4 b) I & N