Theoret. Chim. Acta (Berl.) 60, 41–51 (1981)

THEORETICA CHIMICA ACTA

Molecular Orbital Computations of the Conformational Energy of Ethyl Methyl Phosphate

Camillo Tosi

Istituto Ricerche "G. Donegani", Via G. Fauser 4, 28100 Novara, Italy

Giovanni Lipari

Laboratory of Chemical Physics, National Institute of Arthritis, Metabolism and Digestive Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

As part of a research aimed at an improvement of the description of conformations of polynucleotide chains, we carried out *ab initio* SCF-MO computations on 24 conformations of ethyl methyl phosphate (EMP), a molecule which can be taken as a model for the ω'-rotation of polynucleotides. Quantum-mechanical energies were then reproduced, with good accuracy (standard deviation of fitted from *ab initio* values 0.12 kcal mole^{-1}), through a classical potential function consisting of a "6–12" Lennard-Jones plus a Coulomb term, and a contribution with twofold periodicity accounting for the anomeric effects in the phosphate group.

The binding energies were analyzed using Clementi's concept of Bond Energy Analysis. It was found that the variation of the binding energy as a function of conformation depends on pair atom-atom interactions distributed all over the molecule. This trend in the conformational energy is absent in the classic representation, where changes in binding energy only depend on interactions between non-bonded atom pairs alone. The dependence on interatomic distance of the quantum-mechanical energies, however, is remarkably similar to the one described by classical functions.

Key words: Ethyl methyl phosphate, conformational analysis of ~ – Intramolecular energy, decomposition of ~.

1. Introduction

The conformational behaviour of polynucleotide chains was recently described [1, 2] with a potential function whose parameters were determined, through a

0040-5744/81/0060/0041/$02.20

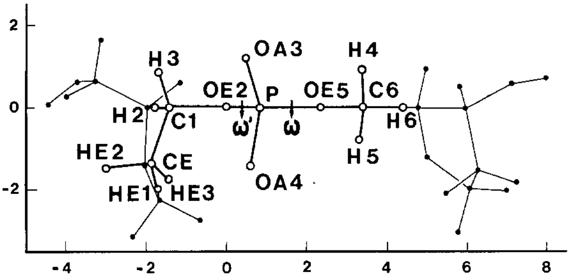

Fig. 1. View of SPS (thin lines) and EMP (heavy lines) in the all-*anti* conformation. The zig-zag chain lies in the *x*, *z*-plane; *x*- and *y*-coordinates are in Å

fitting procedure, from a number of energies obtained with an all-electron quantum-mechanical computation on a low-molecular weight model of polydeoxyribonucleotides, the sugar-phosphate-sugar (SPS) fragment shown in Fig. 1. To our knowledge, this is one of the largest systems for which the possibility of fitting the conformational profiles resulting from *ab initio* computations through an analytical potential has been investigated so far. On account of the complexity of the problem, the potential is the result of a compromise between two opposite requirements, numerical accuracy on one side and conceptual simplicity on the other. Therefore, it cannot be excluded that different expressions will prove better than those already found. That being the case, we asked two general questions [3]: first, how to obtain a potential with a closer agreement with *ab initio* results?, and second, to what extent does this investigation contribute to a better understanding of the origin of the barriers to internal rotation in polynucleotides?

One of the ways which could be reasonably followed in an effort to answer such questions is the study of smaller molecules, representative of well-defined parts of the SPS fragment. The subsystem we focused initially our attention on was the phosphodiester moiety, which is the least hindered portion of the chain. The molecule chosen for this purpose was the dimethylphosphate (DMP) anion $CH_3OPO_2^-OCH_3$ [4]. The potential derived from *ab initio* computations is fairly good from the point of view of the numerical accuracy: the standard deviation of the fitted energies from *ab initio* energies is as low as 0.28 kcal mole^{-1}. However, when applied to SPS, it reproduces the quantum-mechanical energies for the ω-rotation satisfactorily, but fails to reproduce those for the ω'-rotation with desirable accuracy. We suggested a possible explanation of this behaviour [4] by noting that the two rotations about the phosphorus-ester oxygen bonds in DMP could not be a good model for the ω'-rotation of polynucleotides, since this rotation is strongly affected in SPS by interactions of the phosphate group with the furanose ring, which have no counterpart in DMP. If this explanation is correct, a more suitable molecule for mimicking the ω'-rotation in SPS should be the ethyl methyl phosphate anion $C_2H_5OPO_2^-OCH_3$ (hereinafter abbreviated EMP), where the ethyl group simulates the first ring of SPS as regards the above-mentioned interactions. In this paper we present a detailed conformational study of EMP with molecular orbital methods, and we analyse the results of these computations in terms of a "classical" potential function.

2. Results of Quantum-Mechanical Computations

The geometry chosen for EMP is superimposable to the one of SPS described in Ref. [1]; it is highlighted in Fig. 1. The atom called CE in EMP has the same position as the atom called C2 in SPS (it is defined by a dihedral angle of 69.2° between the planes CE-C1-OE2 and C1-OE2-P), and the atom called HE3 here corresponds to the hydrogen of SPS which is replaced by a hydroxyl group in ribose systems.

Quantum-mechanical computations for conformations corresponding to 24 rotations (in steps of 15°) about OE2-P were made with the *ab initio* method IBMOL-6 (see Tables VI–IX of Ref. [1] for a complete description of the basis set, and Ref. [4] for a discussion on the influence of its size on the results). The total energies are shown in Table 1. Subtracting from each of them the sum (−755.19776 Hartree) of the self-energies of the sixteen atoms of EMP (a term which, in a first approximation, may be assumed independent of conformation), we have the so-called "binding energy" of EMP, the conformation-dependent term we try to reproduce with an analytical potential function. For the sake of comparison, the energy barriers computed for SPS [1, 2] are given in Table 1.

For each conformation, we also evaluated the net charges NC according to Mulliken's formalism of Electron Population Analysis [5], and the molecular orbital valency state energies MOVS, according to Clementi's formalism of Bond Energy Analysis (BEA) [6]. The average values of NC and MOVS in the 24 conformations taken into account, together with their variances, are given in Table 2. It can be seen that the higher the variance, the larger the variation of

Table 1. Results of *ab initio* computations on EMP and comparison with the results on SPS[a]

ω' (deg)	Energy, Hartree	Energy barrier, kcal mole^{-1}		ω' (deg)	Energy, Hartree	Energy barrier, kcal mole^{-1}	
		EMP	SPS			EMP	SPS
0	−755.95282	0.53	0.25	180	−755.94466	5.65	6.63
15	−755.95248	0.74		195	−755.94672	4.36	
30	−755.95134	1.46	1.59	210	−755.94908	2.87	2.95
45	−755.95099	1.68	2.19	225	−755.95148	1.37	
60	−755.95181	1.16	1.85	240	−755.95338	0.18	0.15
75	−755.95285	0.51		255	−755.95364	0.01	
90	−755.95366	0.00	0.00	270	−755.95261	0.66	1.51
105	−755.95219	0.92	0.41	285	−755.95209	0.99	
120	−755.94742	3.92	3.42	300	−755.95149	1.36	2.55
135	−755.94145	7.66		315	−755.95107	1.63	
150	−755.93951	8.88	9.78	330	−755.95148	1.37	1.74
165	−755.94182	7.43		345	−755.95224	0.89	

[a] See Table I of Ref. [2].

Table 2. Net charges (NC) and molecular orbital valency state energies (MOVS) in EMP

Atom	NC Average	Variance	MOVS Average	Variance
CE	−0.574	0.0047	1.0934	0.0128
HE1	0.166	0.0075	0.2604	0.0029
HE2	0.146	0.0057	0.2629	0.0022
HE3	0.230	0.0156	0.3071	0.0058
C1	−0.209	0.0056	1.2506	0.0053
H2	0.162	0.0020	0.2650	0.0007
H3	0.179	0.0089	0.2756	0.0034
OE2	−0.618	0.0055	0.4492	0.0166
P	1.730	0.0053	1.9483	0.0077
OA3	−0.860	0.0068	0.3203	0.0049
OA4	−0.838	0.0076	0.2796	0.0068
OE5	−0.618	0.0035	0.4429	0.0092
C6	−0.413	0.0017	1.2046	0.0009
H4	0.170	0.0020	0.2626	0.0009
H5	0.184	0.0021	0.2747	0.0009
H6	0.161	0.0007	0.2732	0.0003

the environment of an atom during the ω'-rotation: thus, e.g., HE3 exhibits a considerable polarization, the charge on it reaching a maximum (0.258) at $\omega' = 30°$ when its distance from OA4 is in a minimum (1.97 Å), and a minimum (0.203) at $\omega' = 195°$ when its distance from OA4 is near a maximum (4.46 Å).

The BEA technique allows to decompose the total energy of a molecule into bonding and nonbonding pairs of interacting atoms. Since any molecular orbital can be expanded into a linear combination of functions centered on the nuclei, the total energy can be written as the sum of one-, two-, three- and four-centre energies. The partition of these terms into only two-center contributions requires a number of manipulations which are arbitrary and basis-set dependent; the net result is a symmetric matrix in which the off-diagonal terms represent all pair-wise interactions within the molecule [7]. EMP has 16 atoms, so the total number of pairs is $\binom{16}{2} = 120$. It is expedient to define four classes of interactions: 15 interactions correspond to the ordinary chemical bonds (type B, or *bonded*); 26 correspond to atoms separated by two bonds (type G, or *geminal*); of the remaining 79 interactions, 30 correspond to pairs of atoms either both lying before or after the rotated bond, or one of which lies on the rotation axis (type NBI, or *nonbonded inactive*) and 49 result from 7 atoms before the rotated bond (the ethyl group) interacting with 7 atoms after the rotated bond (OA3, OA4, OE5 and the methyl group) (type NBA, or *nonbonded active*). The sum of the contributions from the 120 off-diagonal terms of a BEA-matrix equals the binding energy. The single contributions of B, G, NBI and NBA terms to each of the 24 matrices corresponding to the complete rotation of ω' are shown in Fig. 2. The inspection of their trend reveals that, according to the quantum-

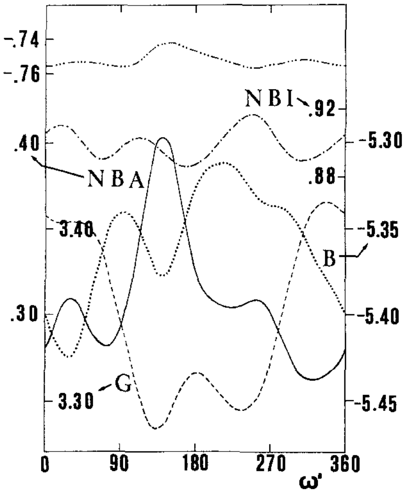

Fig. 2. The decomposition of *ab initio* binding energy (*top curve*) into bonded (*B*), geminal (*G*), nonbonded inactive (NBI) and nonbonded active (NBA) contributions. All energies are in Hartree

mechanical computation, the *overall* variation of the binding energy depends on pair interactions *distributed all over the molecule.* The total fluctuation of the binding energy on rotating ω' is much smaller than the changes in its four components. None of the curves for B, G, NBI and NBA contributions is similar to the curve for the binding energy drawn on top of Fig. 2 (the curve we want to reproduce with the analytical potential). The minima of the binding-energy curve do not correspond to a steady point of any of the contributing curves, but arise from the compensation of contrasting effects, as shown in Table 3. When the binding energy is reproduced with a classical potential function, one implicitly assumes that *only the* NBA *terms contribute to it* (or, put in another way, that B, G and NBI curves are straight-lines parallel with the ω'-axis); should NBA terms not suffice alone to an accurate description of quantum-mechanical energy, other terms may be added (e.g. a torsional barrier accounting for the distortion of bond orbitals during a rotation). Classical theory considers the energy of a molecule as though the atoms whose reciprocal distance is not affected by internal rotation were completely insensitive to the redistribution of internal energy produced by the rotation. Our analysis of the BEA-matrices indicates that the

Table 3. First derivatives (kcal mole^{-1} deg^{-1}) of the curves for bonded (B), geminal (G), nonbonded inactive (NBI) and nonbonded active (NBA) interactions in the steady points of the binding-energy curve

ω', deg	B	G	NBI	NBA
0	−0.69	−0.19	0.25	0.63
45	1.13	−0.12	−0.31	−0.70
90	0.06	−1.14	0.26	0.82
150	0.64	0.68	−0.24	−1.08
255	−0.25	0.50	−0.25	0.00
315	−0.63	0.56	0.13	−0.06

decomposition of binding energy done on a quantum-mechanical basis is not similar to the decomposition done on a classical basis.

3. Derivation of an Analytical Potential

The first question to be answered when deriving a classical potential is its form, i.e. the number and kind of terms to introduce in order to have a physically meaningful and numerically accurate fit. It was suggested [8] that the MOVS' represent an objective criterion for assigning nonbonded atoms to different groups in a pair-wise interaction potential (e.g. a Lennard-Jones or a Buckingham expression). It is now widely recognized that assignments based on the atomic number, or even the hybridization state, are inadequate. As shown in recent works [9, 10], the simultaneous consideration of both NC and MOVS seems more enlightening than the use of either separately. In a plot of MOVS *vs.* NC, the points for the various atoms are clustered in different regions, depending on their *topological* characteristics. In the case of EMP, Fig. 3 shows that there are three distinct regions for the carbon atoms:

C1 (bound to H, H, C, O)
C6 (bound to H, H, H, O)
CE (bound to H, H, H, C),

two regions for the oxygen atoms:

OE (bound to P and C)
OA (bound to P only),

and two regions for the hydrogen atoms: HE3 (bound to a variable extent to oxygen *via* a intramolecular hydrogen bond) and all the other H's (whose distance from oxygen is never smaller than 2.3 Å).

If we assume that, from the point of view of the nonbonded interactions, the behaviour of an atom is determined by its belonging to one of such regions, we should consider 8 kinds of interacting atoms in EMP: this would give rise, in principle, to 36 different atom pairs (of which, though, it is easily seen that only 16 would bring a contribution to the *classical* conformational energy).

Now, it is advisable to limit the parameterization to a more manageable size, since the number of observations of the dependent variable (the *ab initio* energy)

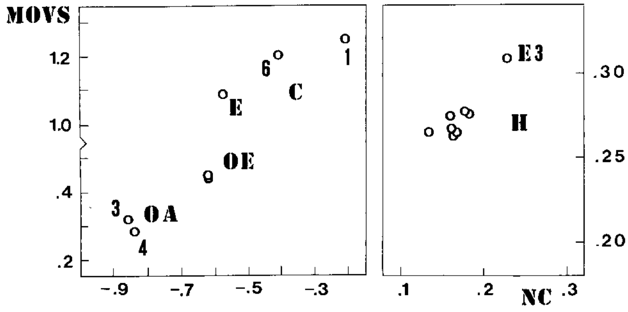

Fig. 3. Molecular orbital valency state energies *vs.* net charges for EMP

Table 4. Groups of nonbonded interactions in EMP

Groups	No. of interactions	Distance of interacting atoms, Å	
		Minimum	Maximum
H···H	15	2.990	7.729
C···H	11	2.903	6.784
C···C	2	3.866	5.773
OA···H	10	1.971	5.694
OA···C	4	2.761	4.674
OE···H	5	1.808	5.640
OE···C	2	2.589	4.614

is small, and further *ab initio* computations, apart from their cost, would add essentially no new information. Using this approach, and under the reasonable assumption that the inclusion of a coulombic term will mostly account for the differences between hydrogen and carbon atoms, a tentative subdivision for EMP is into 5 kinds of atoms: H, C, OA, OE, and P. With this subdivision, the 49 atom-atom interactions active in EMP when ω' is rotated may be assigned to 7 groups, as shown in Table 4.

We therefore try to reproduce *ab initio* energies with the following expression, similar to the one used for DMP [4]:

$$E_t = K_0 + \tfrac{1}{2} V_0' (1 - \cos 3\omega') + \tfrac{1}{2} V_0'' (1 + \cos 2\omega') + \sum_{i>j} \left(-\frac{A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \frac{C_{ij}}{r_{ij}} \right),$$

where K_0 is a constant which includes all interaction energies other than NBA's, and the summation in the latter term is extended over the 49 NBA interactions. $C_{ij} = 332.17\ Q_i \cdot Q_j/\varepsilon$, with 332.17 = conversion factor to kcal mole^{-1}, ε = dielectric constant and Q_i = *average* values of NC_i for the various kinds of atoms ($Q_P = 1.73$, $Q_{OA} = -0.850$, $Q_{OE} = -0.619$, $Q_C = -0.397$, $Q_H = 0.175$). The nonbonded and coulombic parameters which minimize the expression $\sum_{\omega'} (\text{B.E.} - E_t)^2$, determined with the Chandler's technique [11], are given in Table 5. The best values of K_0, V_0' and V_0'' are -443.23, 0.0 and 1.856 kcal-mole^{-1}, resp.. The optimized dielectric constant (incorporated into the term C in the last column of Table 5) is $\varepsilon = 2.53$. The numerical accuracy of the fit is good: the standard deviation is 0.12 kcal mole^{-1} (0.17 kcal mole^{-1} if energy barriers, instead of absolute energies, are considered). Fig. 4 shows the curve of the analytical energy for the ω'-rotation in EMP with its components, and the "points" (*ab initio* energies) to which it is fitted. We note that the potential with twofold periodicity, which reflects the anomeric effects in the phosphate group, is essential to obtain a good fit; without it, we did not succeed in reproducing the minima at 90° and 255° correctly, and hence we could not avoid a consistent loss of accuracy. The similarity (see Fig. 4) of the coulombic

Table 5. Parameters of the potential which gives the best fit to *ab initio* energies for the ω'-rotation of EMP[a]

Group	A_{ij}, kcal mole^{-1} Å^6	B_{ij}, kcal mole^{-1} Å^{12}	C_{ij},[b] kcal mole^{-1} Å
H···H	0.98678+002	0.41344+003	0.40192+001
C···H	0.95915+003	0.13046+007	−0.91178+001
C···C	0.10404+005	0.00000	0.20684+002
OA···H	0.21022+003	0.56107+004	−0.19522+002
OA···C	0.36339+004	0.15209+007	0.44287+002
OE···H	0.00000	0.17117+004	−0.14216+002
OE···C	0.46395+003	0.52235+006	0.32251+002

[a] +00*n* means multiplication by 10^n.

[b] $C_{ij} = 332.17\ Q_i Q_j/\varepsilon$.

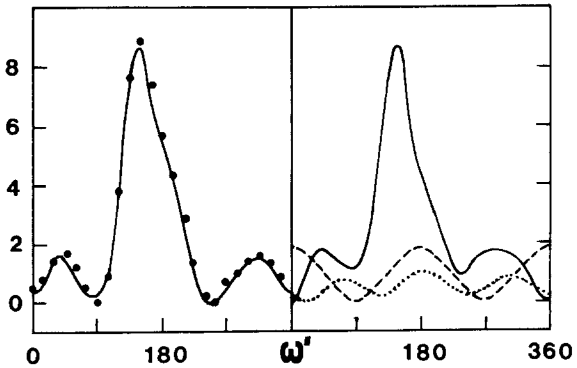

Fig. 4. The energy curve for the ω'-rotation in EMP, constructed with the analytical potential which fits the *ab initio* energies (*full circles, left*), and its decomposition into van der Waals (*solid line*), Coulomb (*dotted line*) and anomeric (*dashed line*) contributions (*right*)

contribution to a cosine function with 120°-periodicity suggests that a comparable fit could be obtained by incorporating the electrostatic contribution into the term $\frac{1}{2}V_0'(1-\cos 3\omega')$: this turns out to be true, but the standard deviation cannot be lowered below 0.16 kcal mole^{-1} (0.24 kcal mole^{-1} for relative energies).

4. Classical vs. Quantum-Mechanical Description of Nonbonded Interactions

The curves of energy *vs.* distance for the seven groups of NBA interactions, as derived from the parameter set of Table 5, are shown in Fig. 5. Their trend is the result of a balance between van der Waals and Coulomb forces. Provided they are applied only within the ranges of interatomic distances which are observed in EMP (Table 4), the presence of unacceptable asymptotic behaviours, like that of the C···C curve at small distances, does not lead to any unrealistic predictions. The functional dependence of the energy of nonbonded interactions

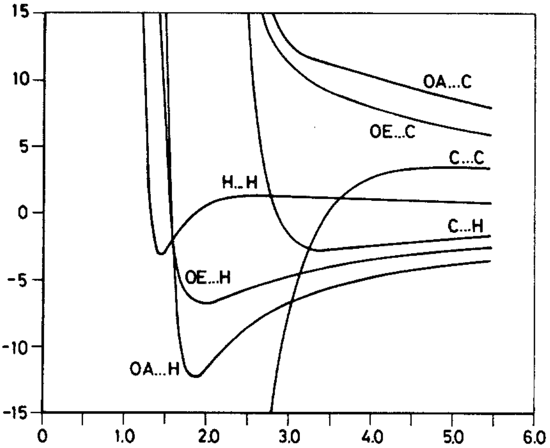

Fig. 5. The curves for the seven groups of interacting atoms, as resulting from the parameters of Table 5

on distance can be also derived from the BEA-matrices. The curves resulting from plots of their off-diagonal terms against nonbonded distances have the seemingly strange property of being not described by single-valued curves (see e.g. Fig. 6). This property can be explained using the following considerations. When the EMP molecule undergoes a complete rotation about the OE2-P bond, the distance of any pair of nonbonded atoms varies from a minimum to a maximum value (e.g. 2.864 Å at $\omega' = 180°$ and 3.857 Å at $\omega' = 0°$ for the pair C1-OE5 of Fig. 6), and takes on all intermediate values twice. There are, therefore, two different ω'-values for which the atoms C1 and OE5 are at the same distance; but the distances between all other pairs of atoms are different (except those having extrema in the same positions as C1-OE5, i.e. 0° and 180°). This implies a difference of the distribution of *ab initio* energies in the two cases and, in particular, of the interaction energy between C1 and OE5.

If we neglect this feature of BEA-curves, and draw smooth lines through the relevant points, we find the plots shown in Fig. 7. The number of these plots reflects the different behaviours produced by the different properties of atoms depicted in Fig. 3: thus, e.g. the scale of repulsions of the oxygen-carbon curves, OA $\cdots$ CE > OE $\cdots$ CE > OA $\cdots$ C1 > OE $\cdots$ C1 reflects the scale of the products

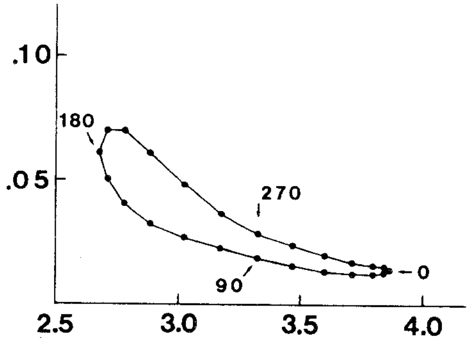

Fig. 6. Interaction energy (Hartree) from Clementi's BEA-formalism *vs.* distance (Å) for the pair OE5-C1. The numbers attached to the curve are ω'-values (deg)

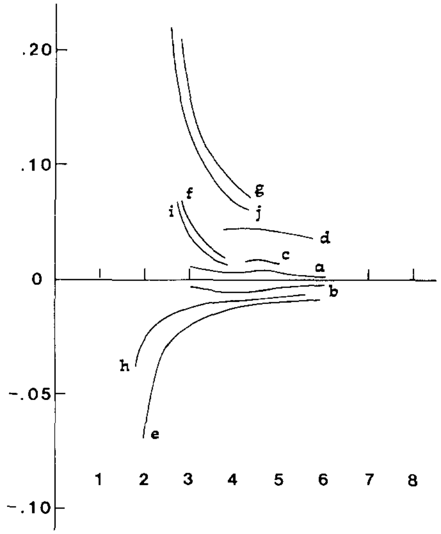

Fig. 7. Nonbonded energy curves (Hartree) from BEA-matrices *vs.* distance (Å): *a*, H ··· H interactions; *b*, C ··· H; *c*, C1 ··· C6; *d*, CE ··· C6; *e*, OA ··· H; *f*, OA ··· C1; *g*, OA ··· CE; *h*, OE ··· H; *i*, OE5 ··· C1; *j*, OE5 ··· CE. The curves are drawn inside the ranges of distances found in EMP

of their net charges (0.49, 0.35, 0.18, 0.13). The same happens for the two C ··· C interactions. OA ··· H is consistently more attractive than OE ··· H. H ··· H interactions have a variety of forms, which however are always weakly repulsive. C ··· H interactions are weakly attractive. The *qualitative* resemblance between the curves of Figs. 7 and 5 is remarkable.

5. Conclusions

The body of data reported in this paper allows us to derive some conclusions whose validity goes beyond the particular case investigated here.

1. While it is possible, for a certain molecule, to find a classical potential function capable to reproduce quantum-mechanical conformational energies with a good degree of accuracy, the extrapolation to even very similar molecules produces a consistent loss of accuracy. Thus, e.g. if we apply the potential of Ref. [4] to EMP, the standard deviation raises to 7.51 kcal mole^{-1}; better, yet not completely satisfactory results are obtained with the potential of Ref. [2] (1.63 kcal mole^{-1}) and with the potential of Table 5 applied to DMP (1.58 kcal mole^{-1}).

2. As a consequence, the hope that the difficulties encountered with a relatively large molecular system such as SPS will be directly overcome by studying smaller subsystems such as DMP or EMP seems somewhat illusory. The study of subsystems gives very useful guide-lines *for choosing the form of the classical force field* (e.g. it points out the necessity of taking account of the anomeric effects in the phosphate group through a contribution with twofold periodicity – which was not present in the potential derived for SPS [2]), but does not permit a straightforward evaluation of accurate numerical values of the parameters.

3. The differences (Figs. 2 *vs.* 4) and the analogies (Figs. 5 *vs.* 7) of the decomposition of intramolecular energy carried out on quantum-mechanical and classical basis contribute to shed light on the possibility of parameterization of force fields from *ab initio* computations. At least in the case dealt with here – which, as far as we are cognizant, is the first example of a complete investigation of BEA-matrices for an intramolecular case – there is a good qualitative agreement in the functional dependence of nonbonded energy on distance. A check, by now *a posteriori*, of the relative trend of the two sets of energies against interatomic distances might give or withhold support to the physical significance of the parameters found. Much additional evidence, though, is needed in order to be able to accept this statement as generally valid.

Acknowledgement. One of us (G.L.) is partially supported by U.S. Public Health Service Grant HL-21483 awarded to Professor F. R. N. Gurd, and thanks the Foundation Stiftelsen Blanceflor Boncompagni-Ludovisi Född Bildt, for a fellowship.

References

1. Matsuoka, O., Tosi, C., Clementi, E.: Biopolymers **17**, 33 (1978)
2. Tosi, C., Clementi, E., Matsuoka, O.: Biopolymers **17**, 51 (1978)
3. Tosi, C., Pescatori, E.: Gazz. Chim. Ital. **108**, 365 (1978)
4. Lipari, G., Tosi, C.: Theoret. Chim. Acta (Berl.) **50**, 169 (1978)
5. Mulliken, R. S.: J. Chem. Phys. **23**, 1833 (1955)
6. Clementi, E.: Lecture Notes in Chemistry, Vol. 2, pp. 1–107. Berlin–Heidelberg–New York: Springer 1976
7. Corongiu, G., Clementi, E.: Gazz. Chim. Ital. **108**, 273 (1978)
8. Clementi, E., Cavallone, F., Scordamaglia, R.: J. Am. Chem. Soc. **99**, 5531 (1977)
9. Melberg, S., Rasmussen, Kj., Scordamaglia, R., Tosi, C.: Carbohydr. Res. **76**, 23 (1979)
10. Castellani, G., Scordamaglia, R., Tosi, C.: Gazz. Chim. Ital. **110**, 457 (1980)
11. Chandler, J. P.: QCPE, Program No. 66.1, Indiana University, Bloomington (1973)

Received April 1, 1981