MOLECULES IN NATURAL SCIENCE AND MEDICINE

An Encomium for Linus Pauling

ZVONIMIR B. MAKSIĆ B.Sc., Ph.D.
Professor of Theoretical Chemistry, Faculty of Sciences and Mathematics
University of Zagreb, Croatia, Yugoslavia

MIRJANA ECKERT-MAKSIĆ B.Sc., M.Sc., Ph.D.
Department of Organic Chemistry and Biochemistry
The Rudjer Bosković Institute, Zagreb, Croatia, Yugoslavia

ELLIS HORWOOD
NEW YORK LONDON TORONTO SYDNEY TOKYO SINGAPORE

Table of contents

reprinted from:
J. Köhler, R. Kikuno and A. Plückthun, Non-covalent interactions in biologically relevant complexes. Importance of electrostatics and the atomic charge concept. In: Molecules in natural science and medicine. An Encomium for Linus Pauling, Z. B. Maksic and M. Eckert-Maksic, Eds. (Ellis Horwood, New York, 1991) pp. 441-465.

23

Non-covalent interactions in biologically relevant complexes. Importance of electrostatics and the atomic charge concept

J. Köhler, R. Kikuno and A. Plückthun
Ludwig-Maximilians-Universität, Munich, FRG

INTRODUCTION

Few scientists have influenced the field of biophysics from so many different sides as has Linus Pauling. His interest has left out hardly any important topic, and many of his ideas and concepts are now called 'classic'. His views on the particular importance of the three-dimensional structures and the underlying physical forces as the basis of biochemical processes have influenced the development of the whole field. The enormous progress in this field demonstrates the fertility of his concepts. Linus Pauling has contributed the classic book on the nature of the chemical bond [1] and he has been especially aware of the nature of forces between large molecules of biological interest. His concepts about the secondary structure of proteins [2, 3] and of molecular architecture for biological catalysis [4] are of particular importance. The analysis of electrostatic forces in proteins [5, 6] and the study of their aqueous environment [7-9] might lead to a deeper understanding of protein structure, function and dynamics [10].

Over the last two decades, several empirical force field methods [11-32] have been developed, mainly for the application of molecular mechanics and molecular dynamics simulations of biological molecules [33] *in vacuo*, in solution or in the crystalline state. Typical examples are the well known program packages AMBER, CHARMM, DISCOVER, GROMOS and X-PLOR (see listing at the end). An overview of biomolecular applications has been given [32].

Here, we show what an analysis of the electrostatic interactions in molecular complexes during a molecular dynamics trajectory reveals. The system investigated is the antibody/antigen complex of McPC603/phosphorylcholine and two related antibodies, M167 and T15, which also bind to phosphorylcholine, but are derived from different light chain genes.

THE MOLECULAR DYNAMICS METHOD

The general idea of a molecular dynamics simulation [34] is to solve Newton's equations of motion for a system of interacting atoms. Currently, time scales from femto- to nanoseconds can be reached for biological molecules. The system can be investigated *in vacuo*, in solution or in the crystalline state. The interacting particles may then be only the solute or the solute plus solvent (and possibly ions). At the start of the simulation, initial velocities are assigned to the atoms corresponding to a Maxwell-Boltzmann distribution at a given temperature. All atoms then move according to the forces they create between themselves as defined by the interaction energies from the force field.

Newton's equations are integrated in small time intervals (usually 0.5 to 2 femtoseconds), and at each time step, the forces, velocities and positions of each atom are calculated to give a so-called trajectory.

$$\frac{d^2\bar{r}_i(t)}{dt^2} = \frac{\bar{F}_i(\bar{r}_1, \ldots, \bar{r}_N)}{m_i} \qquad i = 1, \ldots, N$$

Different algorithms may have advantages for specific applications. These algorithms [35-40] are either computationally very fast or very accurate [41]. The simulations can be quite time consuming: a trajectory of one picosecond of a protein with 235 amino acids in about 4000 water molecules and 20 ions (about 17 000 atoms in total) takes about 1 h of Cray YMP cpu time. If the system contains only one protein molecule, one should have trajectories of up to 100 or 200 ps to sample enough conformational space so that the statistical error becomes small enough.

In the current molecular dynamics approaches [42-45], the atoms are considered to be particles with mass, van der Waals radius and electrical atomic charge. Thus, electrons are not considered explicitly. The force field is a conservative one, meaning that the forces are dependent upon the distances between the atoms only. The potential is a pair potential, making it computationally fast.

The principal form of the current potential energy may be written as

$$\begin{aligned} V(\bar{r}_1, \ldots, \bar{r}_N) = & \sum_{l=1}^{N_b} \frac{1}{2} K_{b_l}(b_l - b_{0_l})^2 + \sum_{l=1}^{N_\theta} \frac{1}{2} K_{\theta_l}(\theta_l - \theta_{0_l})^2 \\ & + \sum_{l=1}^{N_\xi} \frac{1}{2} K_{\xi_l}(\xi_l - \xi_{0_l})^2 + \sum_{l=1}^{N_\phi} K_{\phi_l}(1 + \cos(n_l\phi_l - \delta_l)) \\ & + \sum_{i<j}^{N} \left(\frac{C_{12}(ij)}{r_{ij}}\right)^{12} - \left(\frac{C_6(ij)}{r_{ij}}\right)^{6} + \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}} \end{aligned}$$

The first term concerns the covalent bonds between two atoms: as soon as the atoms move to the distance b_l and away from their equilibrium distance b_{0l}, a force with the force constant K_{b_l} drives them back. The second term describes the bond angles. The fourth and fifth terms model the dihedral angles, which can undergo limited (i.e. in rings) or full 360° rotations, respectively. In the last three terms, the interactions between atoms that are not covalently bound to each other are described. The repulsion between atoms has a $1/r^{12}$ dependency, the attractive (or dispersion) part is proportional to $1/r^6$, and together they model the van der Waals forces in a

Lennard-Jones-type potential. The last term contains the electrostatic interactions between all atom pairs. It is actually the first term of the infinite multipole expansion. This means that all terms of higher order (dipoles, quadrupoles, etc. and their interactions) are not taken into account explicitly. The force field is an 'effective' one, meaning that all parameters such as the force constants, Lennard-Jones parameters, atomic charges, etc. are scaled such that they reproduce the experimental observables of molecular ensembles well.

There are also somewhat different terms in use. Moreover, additional terms may be introduced: for example, Morse potentials can be taken for the covalent bonds, in order to have a better description of the bond-lengths, or cross-terms can be applied as in the DISCOVER force field [46] to model the coupling between two terms (used for calculating vibrational spectra). Some force fields contain extra hydrogen bond terms of van der Waals (Lennard-Jones)-type (the AMBER force field [20, 21]).

The computation of the non-bonded interactions in the above formula between an atom and all its neighbors takes about 80% of the total simulation time, mostly because the search algorithms for neighbors are time consuming.

WHAT CAN BE LEARNED FROM MOLECULAR DYNAMICS TRAJECTORIES?

The method at the current state of the art is designed mainly for trajectories on the picosecond time scale. A few examples that have been studied in detail may serve to illustrate some of the characteristics: the agreement between the neutron diffraction structures of cyclodextrin crystals and the averaged simulated structures at room temperature (293 K) is rather high [47, 48]. The experimental positions of non-hydrogen glucose backbone atoms were reproduced within 0.34 Å. This value is smaller than the root-mean-square (rms) atomic fluctuation of 0.41 Å as derived from the crystallographic B-factors. The simulation showed an overall rms atomic fluctuation of 0.49 Å, slightly larger than the experimental value.

At low temperature (120 K) [48] the experimental positions were reproduced to within 0.46 Å. This value is larger than the rms atomic fluctuation of 0.19 Å (experimental) or 0.22 Å (simulation). Both experiment and simulation showed a reduction by a factor 2 of the atomic mobility when the temperature was lowered from 293 to 120 K. The larger deviation of the atomic positions from the experimental ones at low temperature might be due to the fact that the empirical interatomic potential function has been designed as an effective force field at room temperature. The flexibilities of molecules may be understood from the rms deviations during the trajectories, giving an impression of how extended the conformational space is at a given temperature.

On the picosecond time scale, the dynamic behavior of hydrogen bond phenomena occurs [49-51] and can be observed in the simulations. Forming and breaking of two-, three- or multiple-center hydrogen bonds [52] as well as flip-flop hydrogen bonds [53, 54] were observed in trajectories of cyclodextrin crystals, with an accuracy of about 70 to 80% agreement compared to neutron diffraction data [55, 56].

From a simulation of an α-cyclodextrin molecule in solution [57] one can see that the molecule clearly explores a larger part of the conformational space in solution than in the crystal as determined in the neutron diffraction structure.

CALCULATION OF DIFFERENCES IN FREE ENERGY WITH MOLECULAR DYNAMICS

The possibility of calculating differences in free energies between related molecular systems has given the molecular dynamics method a broad range of applications in the field of protein-ligand interaction [58] as a defined theoretical method. The principles of this method have been described by several groups [31], and computational details have been discussed for ions in solution [59, 60], or applications for biological molecules [58, 61].

The methodological concepts for the calculation of differences in free energies with the molecular dynamics method have been explained in review articles by Kollman and van Gunsteren [62], and by several other authors [29, 31, 63-65]. Some applications of this method and a somewhat different approach to calculate free reaction energies have been reported [58]. (For a general overview of biochemical thermodynamics, see ref. 66.)

The calculated differences between the free energies of two systems can be directly compared to the experimentally determined binding constants using the following thermodynamic cycle:

$$\begin{array}{ccc} & \Delta G_1 & \\ Enzyme + Substrate_1 & \rightleftharpoons & Enzyme - Substrate_1 - Complex \\ \Delta G_3 \Downarrow & & \Delta G_4 \Downarrow \\ Enzyme + Substrate_2 & \rightleftharpoons & Enzyme - Substrate_2 - Complex \\ & \Delta G_2 & \end{array}$$

In this example, the values for ΔG_1 and ΔG_2 may be obtained from experiment (binding constants K_1 and K_2 of the two different substrates; $\Delta G = RT \ln K$). Values for ΔG_3 and ΔG_4 come from molecular dynamics calculations. Since the G-function is a state-function, the sum of all ΔG must be zero. The way in which the vertical processes are simulated is without physical reality: the force field parameters are changed from the ones that belong to substrate$_1$ into the ones that belong to substrate$_2$ during the simulation. This can be done [31] by redefining the Hamiltonian $H(p, r)$ in potential energy $V(r)$ and momentum p space as a function of the coupling parameter λ. Now the starting state A is described by $H(p, r, \lambda_A)$ and the final state B by $H(p, r, \lambda_B)$:

$$H(p, r, \lambda) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i(\lambda)} + V(r_1, \ldots, r_N, \lambda)$$

During such a reversible process, the molecules adopt many configurations for which the enthalphic and entropic contributions to the total energy are calculated. The value for the free energy is evaluated by integrating:

$$\Delta F_{BA} = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial H(p, r, \lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda$$

For an isobaric ensemble, a similar formula is valid [67]. The evaluation of the entropy is also based on the ensemble average:

$$S = -\frac{\partial F}{\partial T}$$

so that finally the difference $\Delta G = \Delta H - T\Delta S$ in free energy between states A and B can be obtained from the simulation. This fact makes molecular dynamics simulations a very powerful tool for comparing theoretical and experimental data, although it is computationally very intensive.

The calculated data are in good agreement with experimental data as long as the two structures are closely related, or at least do not differ in their absolute charge values. Here again, the importance of the charges becomes obvious. An example for a calculation of the relative change in binding free energy of a protein-inhibitor complex is the case of the enzyme thermolysin with a pair of phosphonamidate and phosphonate ester inhibitors. The calculated difference was 4.21 ± 0.54 kcal/mol, in nice agreement with the experimental value of 4.1 kcal/mol [68].

BIOCHEMICAL RELEVANCE: ANTIBODY/ANTIGEN COMPLEXES

Antibodies are a class of proteins with which most of the problems relevant in protein-ligand interactions can be investigated. Even chemical rate accelerations can be observed and studied with antibodies. Since antibodies can in principle be made against any substance, this offers exciting prospects of developing new catalysts.

Once again, this concept can be traced back to Linus Pauling. In an article in *Chemical and Engineering News* [4], he explained his picture of enzyme catalysis. It was clear to him that the surface of the enzyme must be complementary not to the substrate molecule itself but rather to a 'strained configuration', corresponding to the activated complex. This way, part of the intrinsic binding energy of the substrate can be used to bring the substrate closer to the transition state. At this time, of course, no three-dimensional structure of an enzyme was known, but today, this concept has been verified by crystallography [69, 70]. In 1969, Jencks [71] then first proposed raising antibodies against analogs of the transition state to test whether they can cause catalysis. This has been tried several times soon thereafter, but true success only came after the invention of monoclonal antibodies, allowing one to reach the high fairly molar concentration of a pure antibody necessary to observe significant rate accelerations. A whole number of reactions have now been catalysed with this concept [72-75].

An example shall illustrate this concept. The antibody McPC603 binds phosphorylcholine, and its three-dimensional structure is known [76, 77]. During the hydrolysis of an appropriate ester, the intermediate is tetrahedral [78a, b], and therefore structurally similar to the tetrahedral arrangement of the phosphate (Fig. 1). Hammond's postulate suggests that the two transition states leading to the tetrahedral intermediate and away from it will be very similar in structure to the intermediate. Indeed, this antibody McPC603 [79-81] as well as the related antibodies T15 and M167 [82, 83] all catalyze the hydrolysis of p-NO_2-phenyl-choline-carbonate, albeit moderately.

Unfortunately, the structures of the antibodies T15 and M167 are not known.

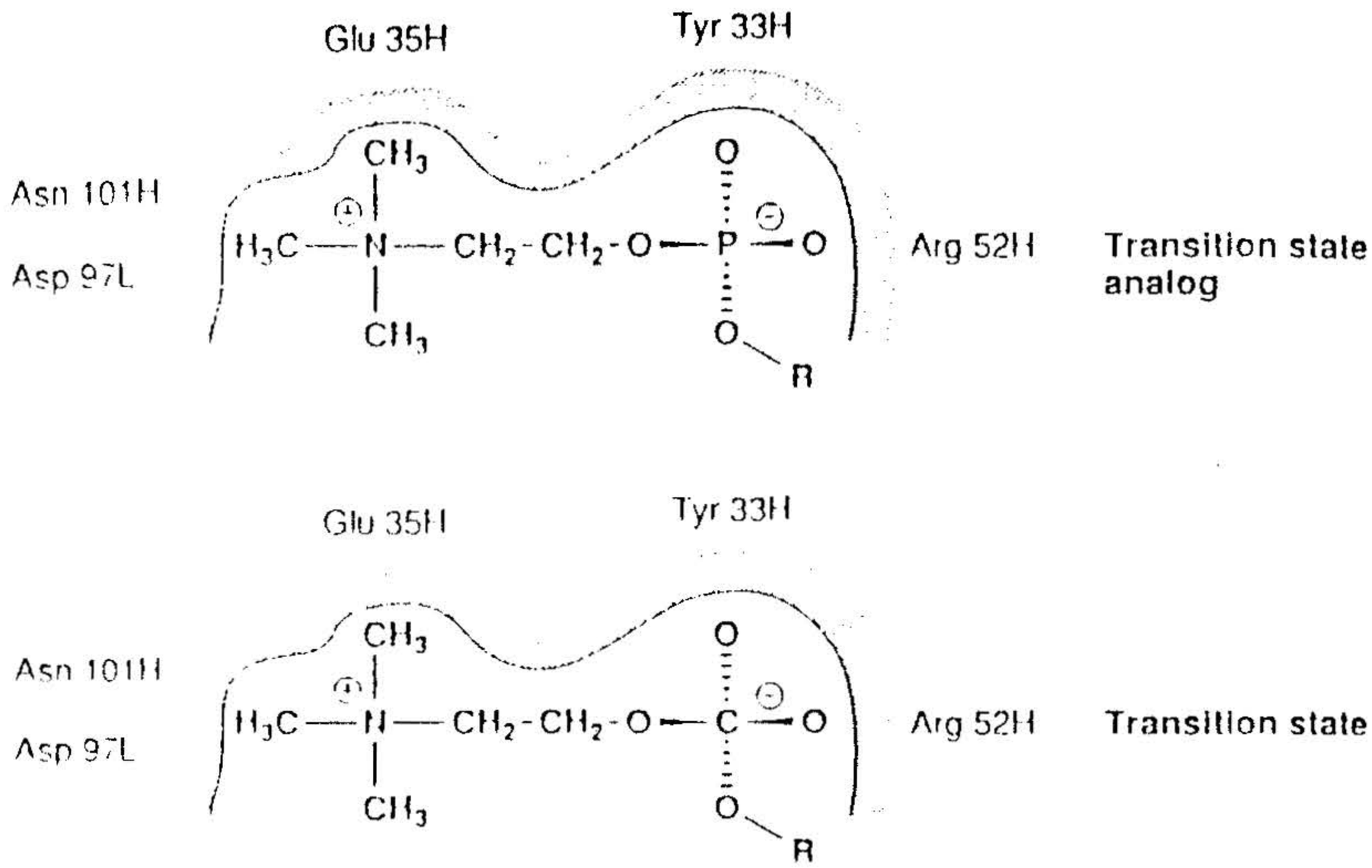


Fig. 1. Schematic representation of the analogy of transition state binding and binding of the transition state analog. Residues from the antibody McPC603 are indicated schematically.

Recently, progress in the production of engineered antibodies from bacteria [84-86] has allowed the investigation of the effect of contributions from single amino acids [79, 80] on binding and catalysis, as well as renewed attempts of crystallizing antigen binding fragments of the antibodies T15 and M167.

The desire for a theoeretical understanding of this experimentally well-characterized system has been the impetus of some of the work described here. Binding constants between similar antigens or mutants differ only by a few kilocalories, and thus demand rather accurate methods for any useful predictions.

The first step was to model the structures of T15 and M167 starting from the known structure of McPC603 by replacing the different amino acids. This is complicated by the fact that the loops making up the binding pocket have different lengths. Models constructed with a distance matrix of the Brookhaven-Protein-Data-Base were used as the starting structures for the molecular dynamics simulations. The calculated systems contained one molecule of either McPC603, T15 or M167 (the Fv-fragments only, which is the heterodimer of the variable domains of the heavy and light chain, together about 235 amino acids), and some 4000 water molecules and ions (Fig. 2). After energy minimization, a trajectory of 30 ps at 300 K was calculated for each system (Figs 3(a), (b), (c)). The agreement between the crystal structure and the calculated trajectory of McPC603 can be seen in Figs 4(a), (b), where the B-factors, rms positional fluctuations and structural elements of Cα atoms [87] of the Fv-fragment are displayed. The atomic fluctuations in the crystal are significantly lower than in the simulated solution structure. There is a general agreement between experiment and simulation concerning the flexibility of certain areas (peaks in Figs 4(a), (b))

Whereas the trajectory of McPC603 equilibrated, the ones for T15 and M167 did not. This can be seen from the analysis of the contributions of single amino acids to

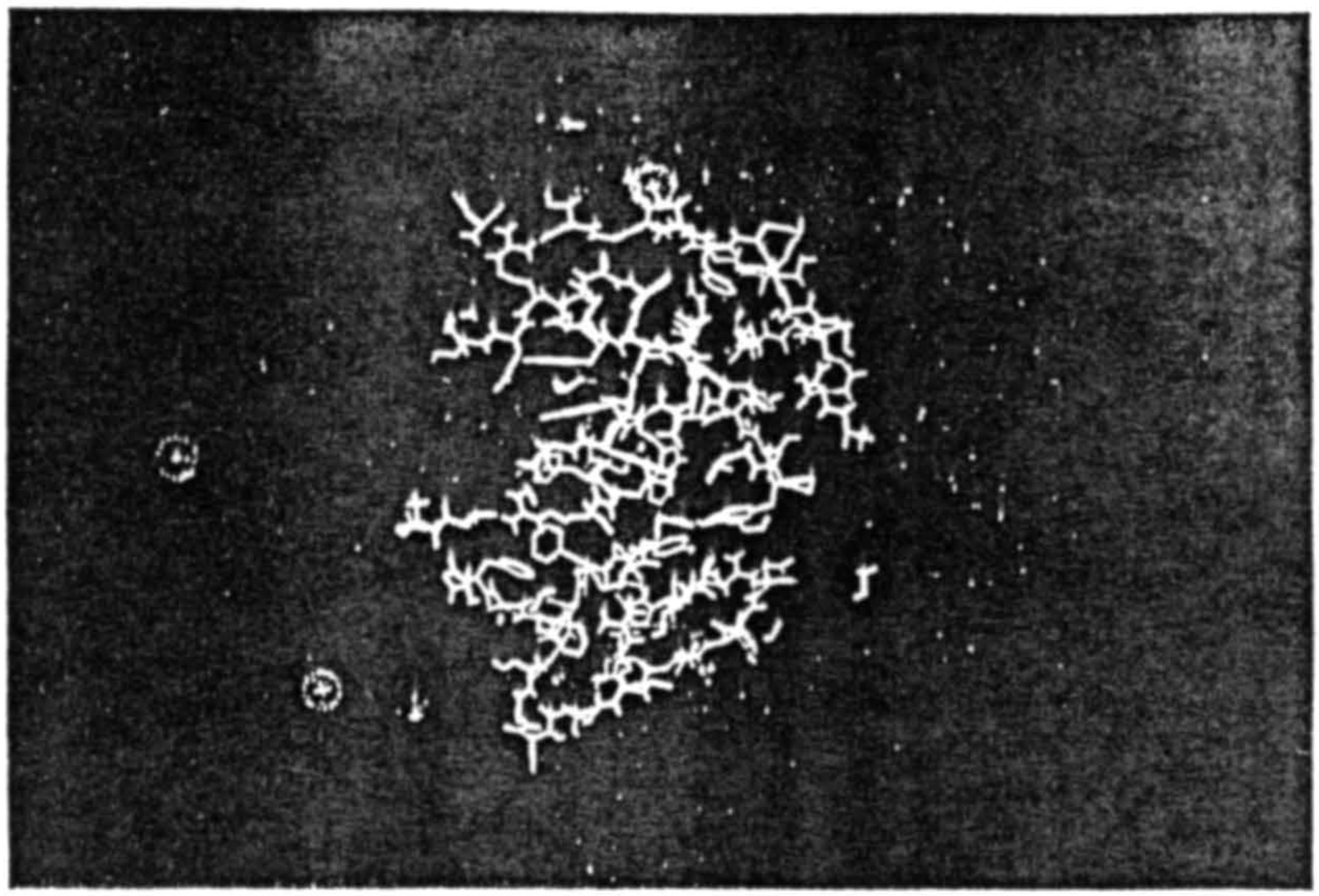


Fig. 2. — The Fv-fragment (of McPC603) and the bound antigen phosphorylcholine in a shell (6 Å) of water molecules including Na^+ and Cl^- ions. This system was simulated over 30 ps without any constraints (GROMOS87).

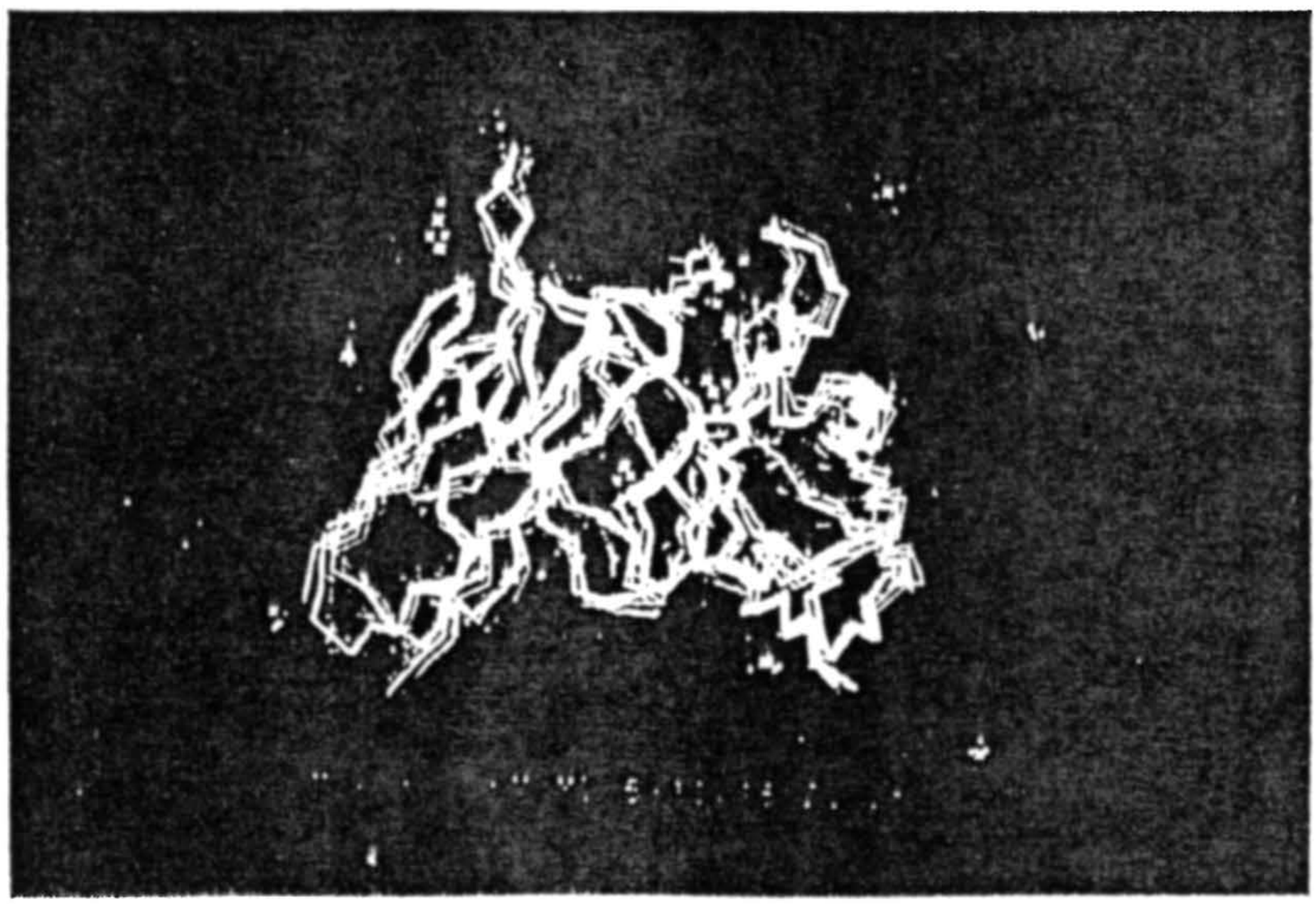
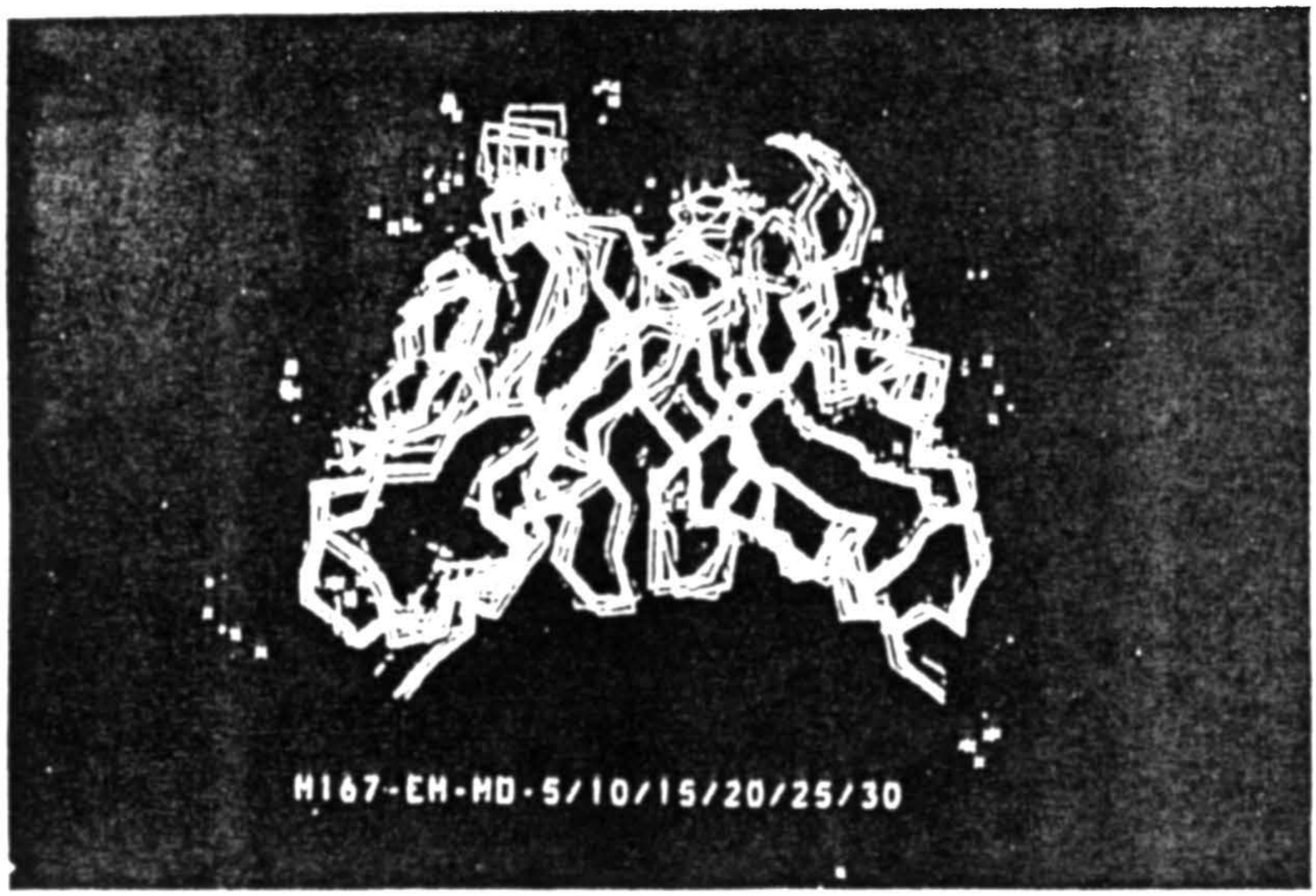


Fig. 3. — The backbone atoms of the Fv-fragments of McPC603 (A).

B



C

M167 (B) and T15 (C) during the molecular dynamics simulations of 30 ps each. In dark blue the starting structures including the side chain atoms are shown.

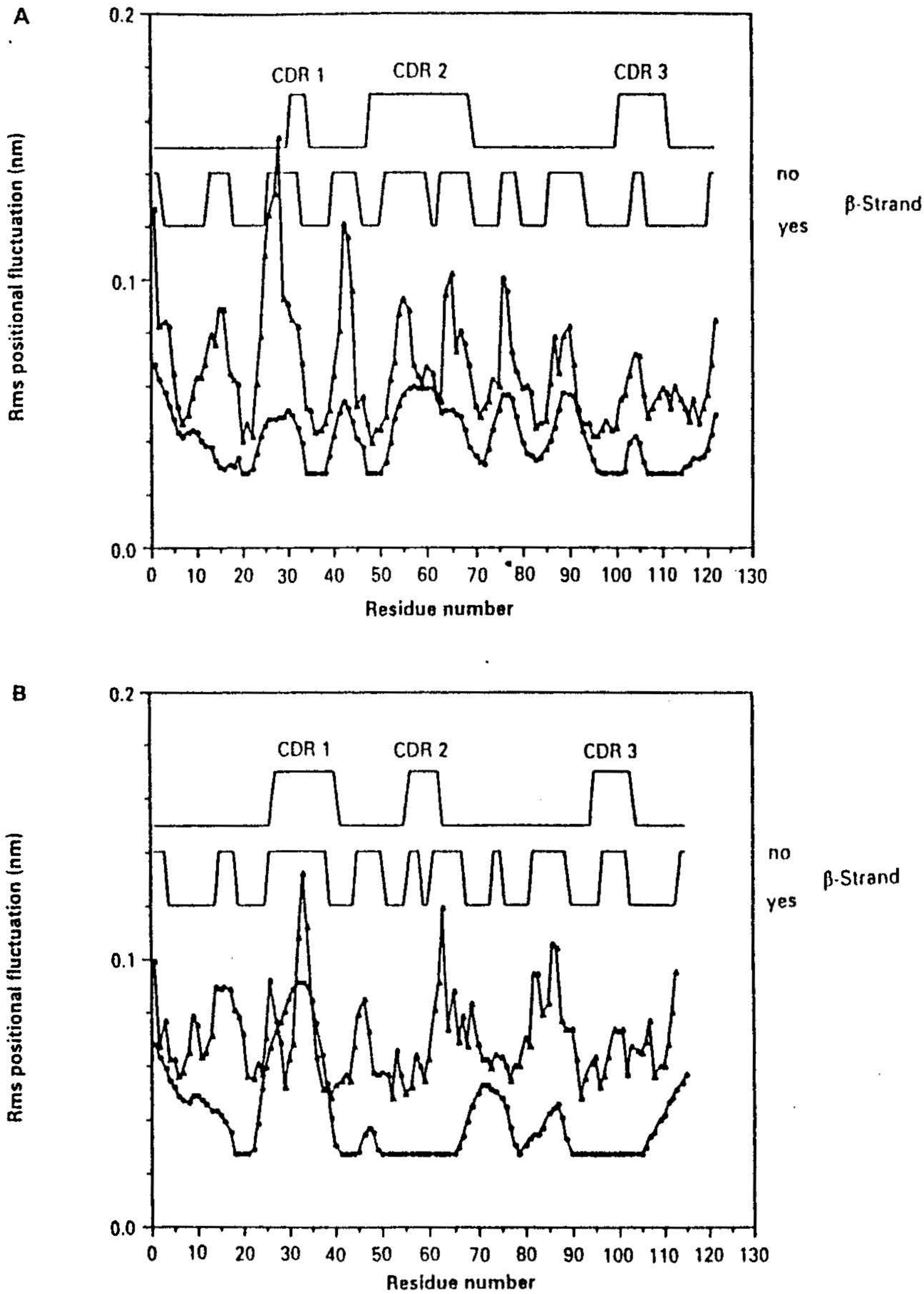
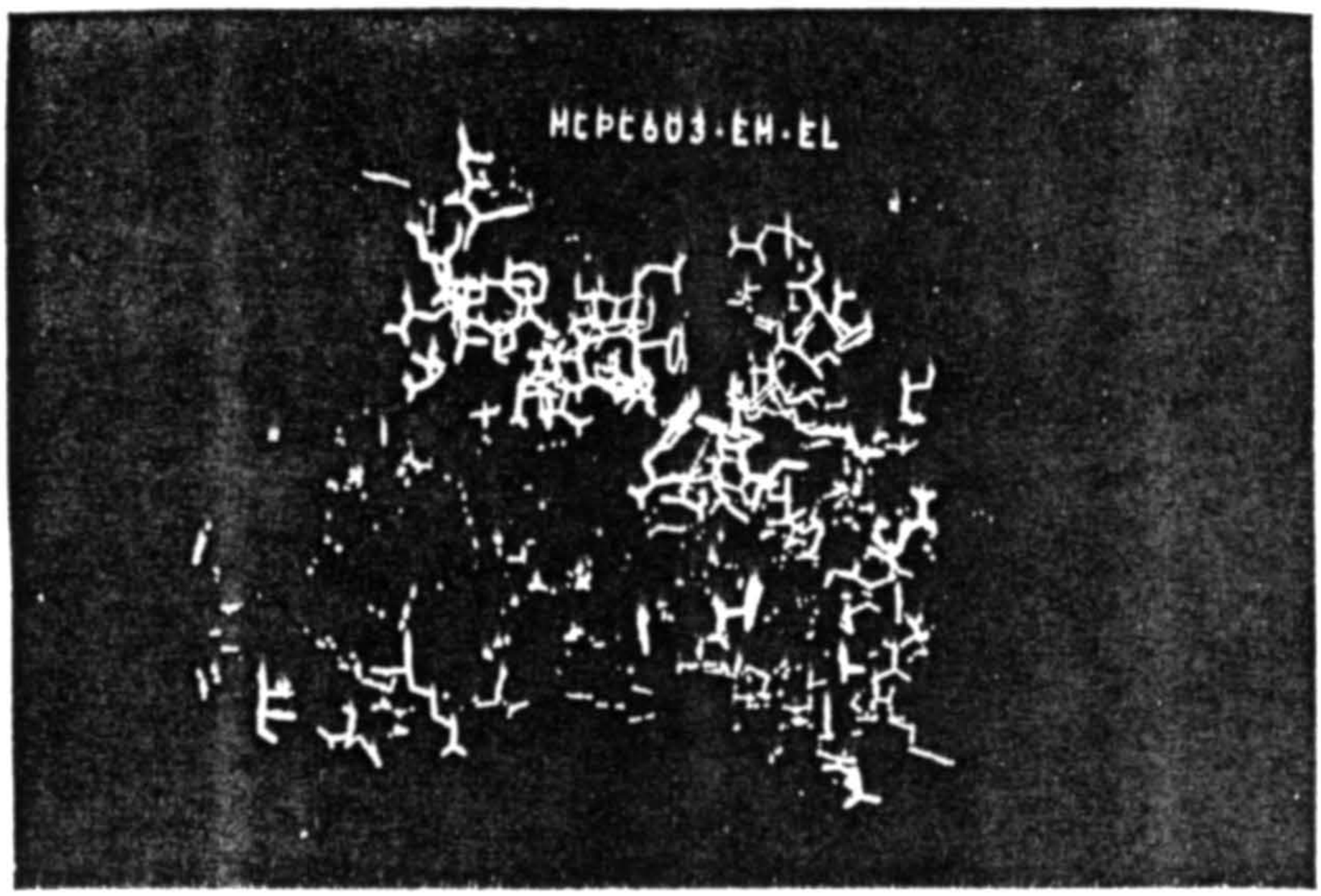


Fig. 4.—Rms fluctuations derived from crystallographic B-factors of Cα atoms (open circles) and rms positional fluctuations from a 30 ps trajectory in solution at 297 K (triangles) are shown. Both refer to the McPC603-phosphorylcholine complex [76], and (A) is the analysis for the heavy chain, (B) for the light chain. The β-strands are indicated (defined as in ref. 87), as are the complementarily determining regions (CDR) of each domain, defined by genetic variability. These are the antigen binding loops.

A



B

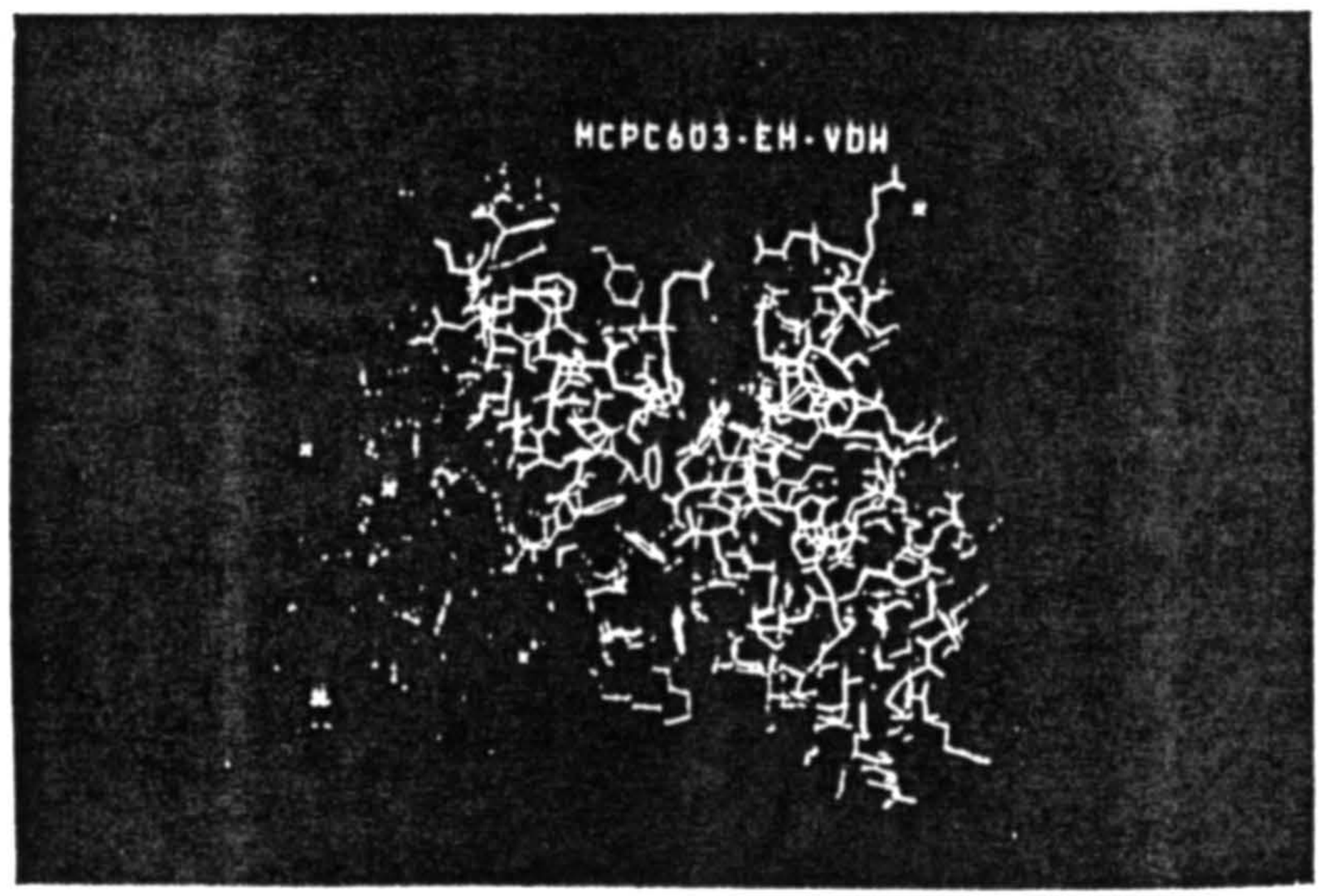


Fig. 5. — The energy-minimized structure of McPC603 with all amino acids colored according to their interaction energies towards phosphorylcholine. Red: strongest interaction calculated with the program PROELW; yellow, cyan, white, increasingly weaker interactions. (A) electrostatic, (B) van der Waals, (C) sum of (A) and (B).

C

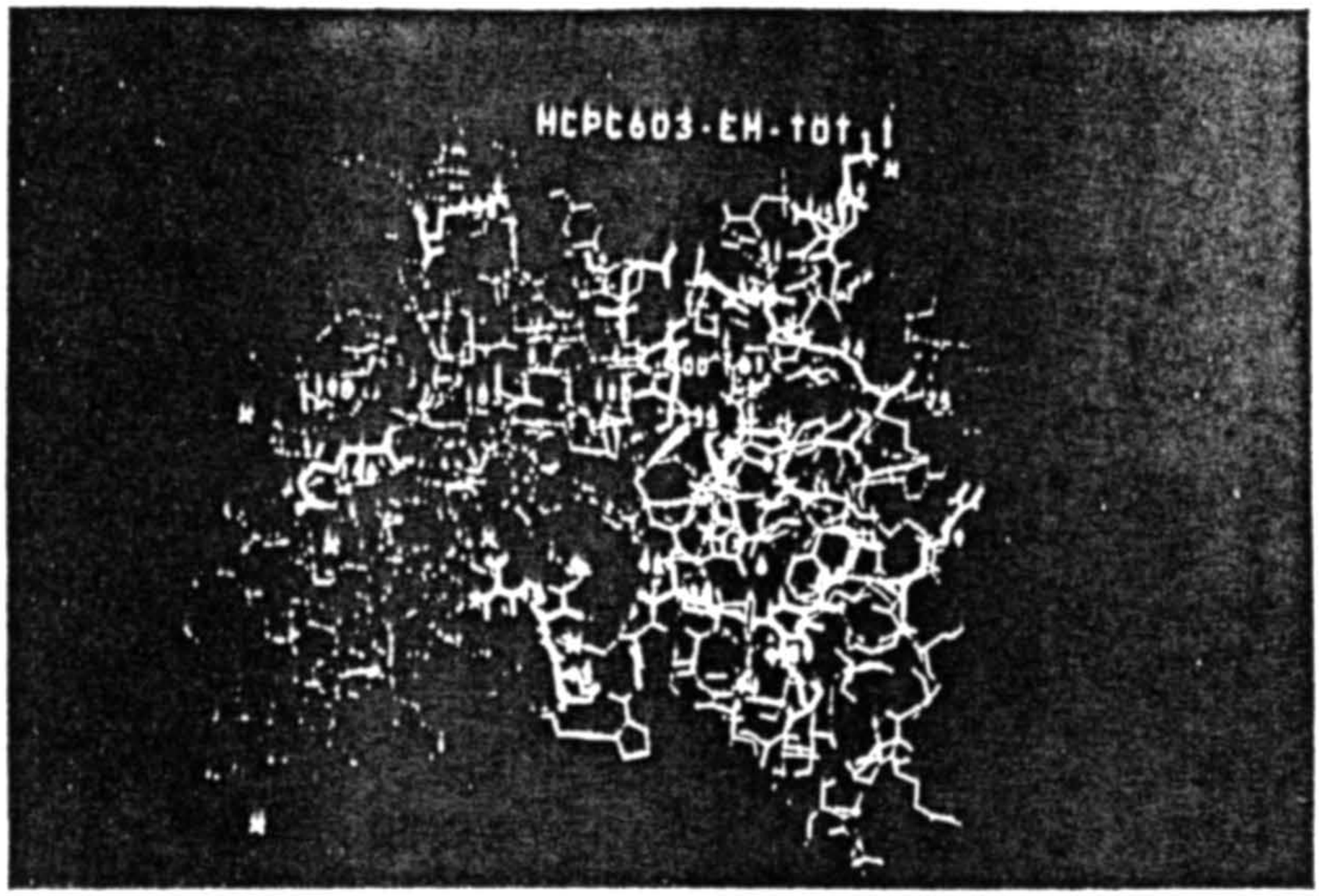
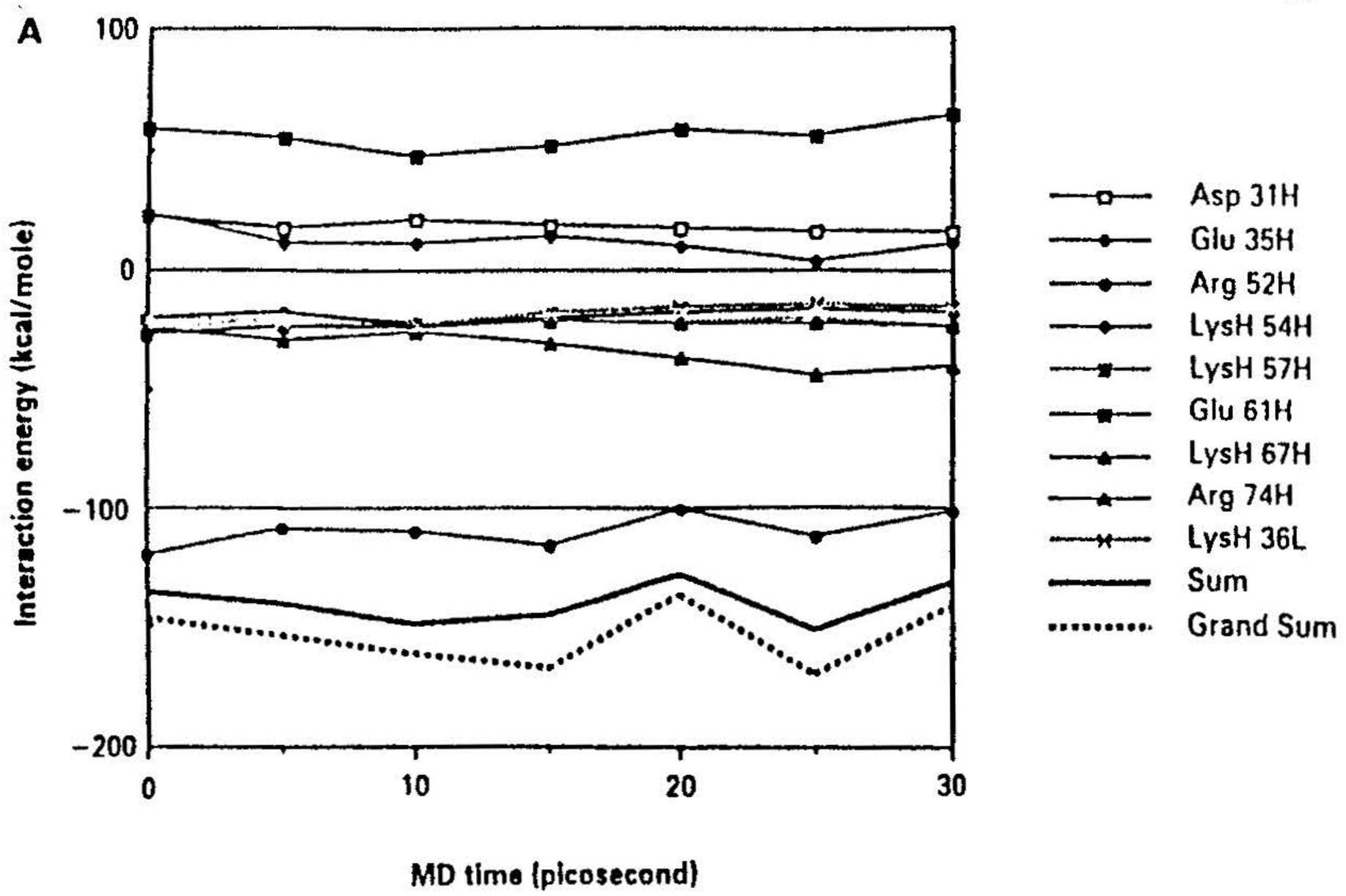


Fig. 5 *Contd.*

electrostatic and van der Waals energies and the sum of both (Figs 5(a), (b), (c)) over time [88]. The fluctuations of these electrostatic energies (Figs 6(a), (b), (c)) and van der Waals energies (Figs 7(a), (b), (c)) or the sum of both (Figs 8(a), (b), (c)) between amino acids of the antibodies and phosphorylcholine in the simulation of the modeled structures of M167 and T15 are much larger than in McPC603, which could be started from its known crystal structure. It is clear, therefore, that convergence is fairly slow under these conditions. It is hoped that the continued trajectories will converge and their results might be compared to experimental structures of M167 and T15 in the near future.

It should be noted that the comparison of potential energies, like electrostatic and van der Waals energies, must be used with care for any interpretation of substrate binding, as the real binding reaction contains entropic terms. Instead, free energy calculations have to be performed on derivatives of the protein and the ligand [89]. These calculated differences in free energies can be compared directly to the experimental binding constants.

Since in the case of the antibodies McPC603, M167 and T15 the binding constants to phosphorylcholine and its derivatives have been measured to be in the narrow range from 3 to 7 kcal/mol, a very high accuracy of the atomic charges used in the simulations is needed. This is of even higher importance since the substrates are molecules with net charges, and some have zwitterionic structures. In order to obtain atomic charges of reasonably equal quality for all derivatives we will use the SCALCHA approach [90, 91] in the future. The method performs a projection of quantum mechanically or semiempirically determined charges onto atomic charges as used in current molecular dynamics force fields.



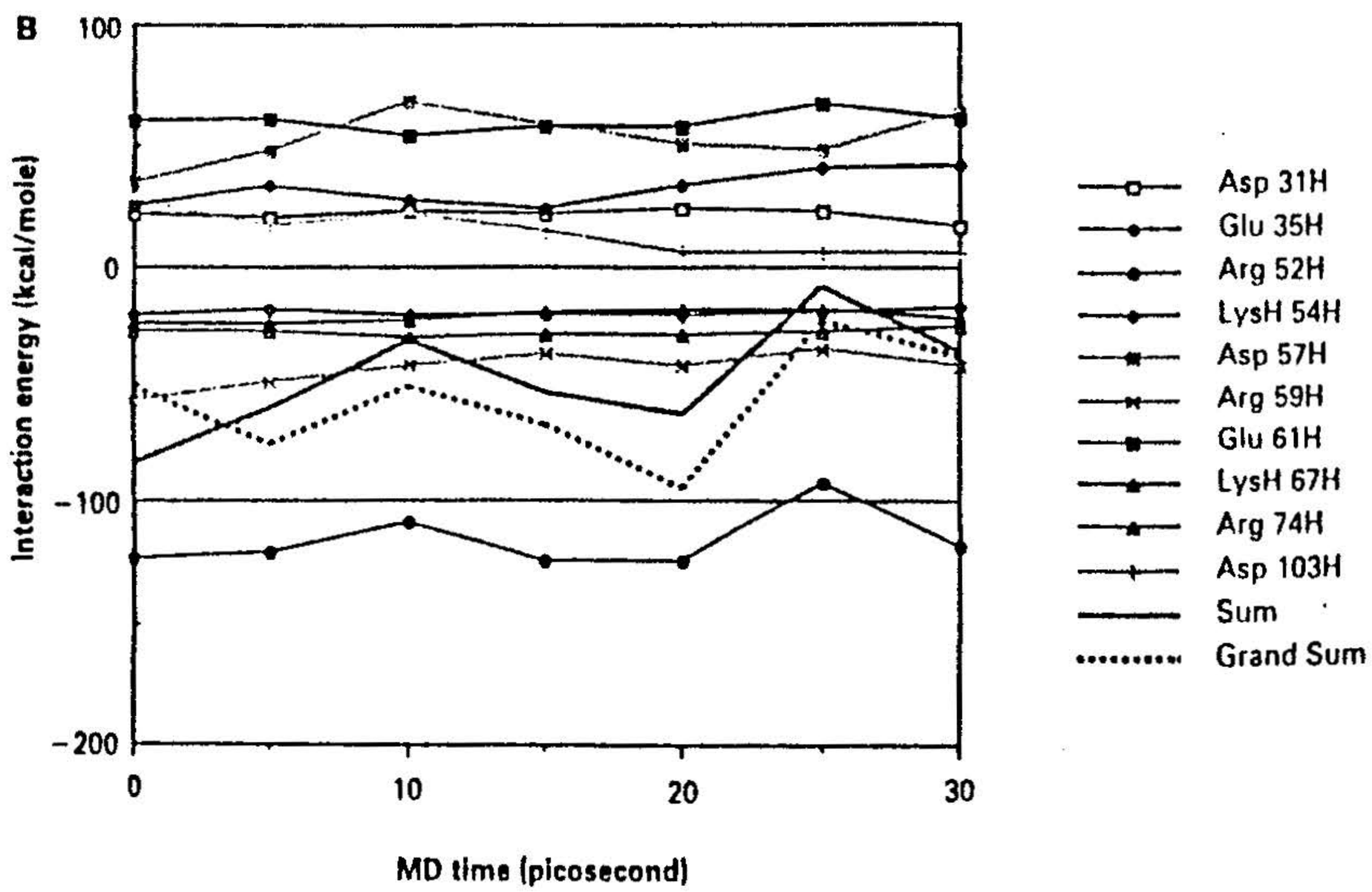


Fig. 6. — The electrostatic interaction energy (kcal/mol) for the three antibodies McPC603 (A), M167 (B) and T15 (C), each bound to (deprotonated) phosphorylcholine calculated from the structures in the trajectories at 5 ps intervals. The amino acids with the highest interaction energies with phosphorylcholine are shown. The sum is the sum of these amino acids. The grand sum is the sum over all amino acids in the whole calculated system. LysH denotes the protonated form of Lys in the GROMOS force field.

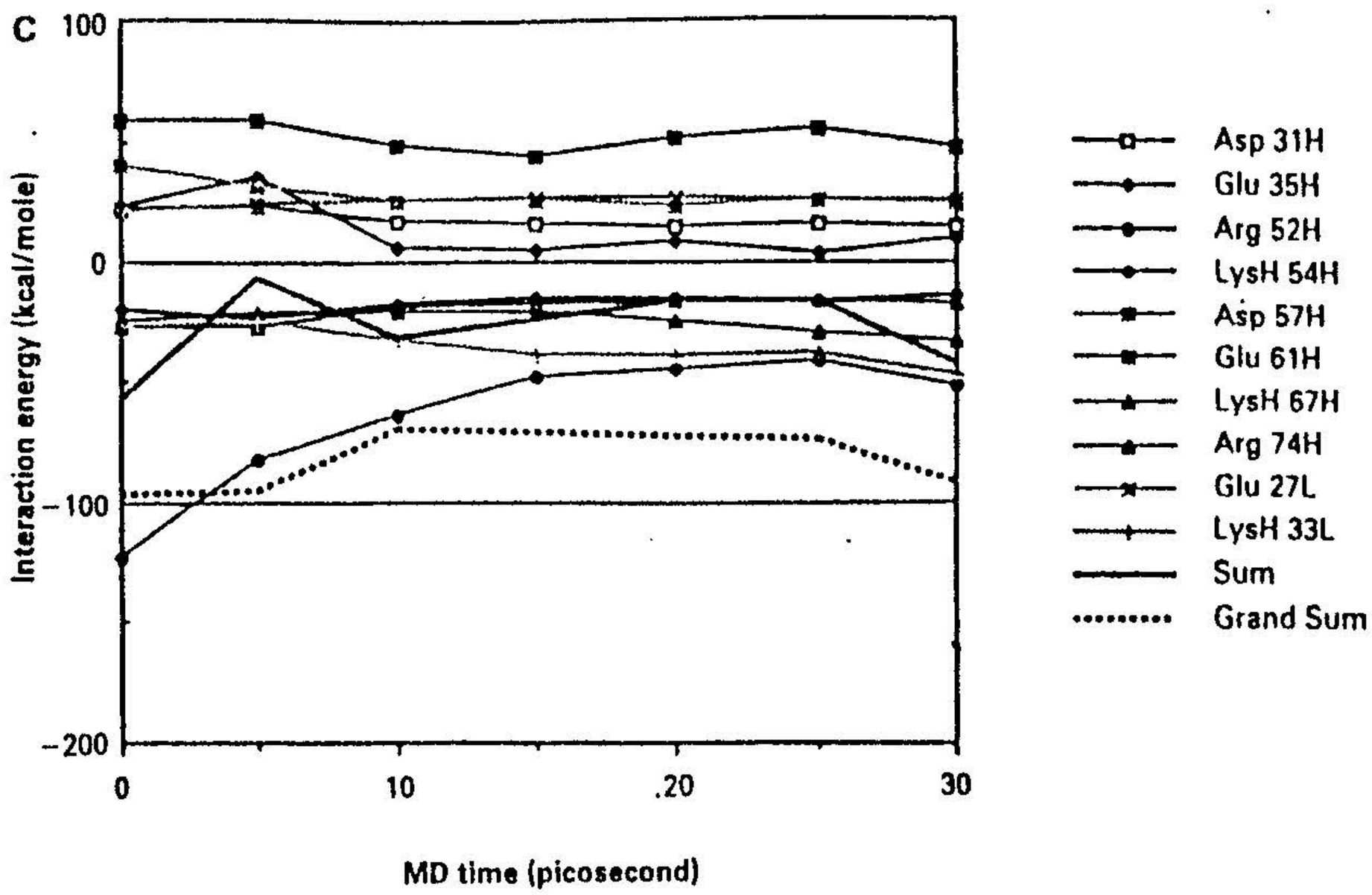


Fig. 6 *Contd.*

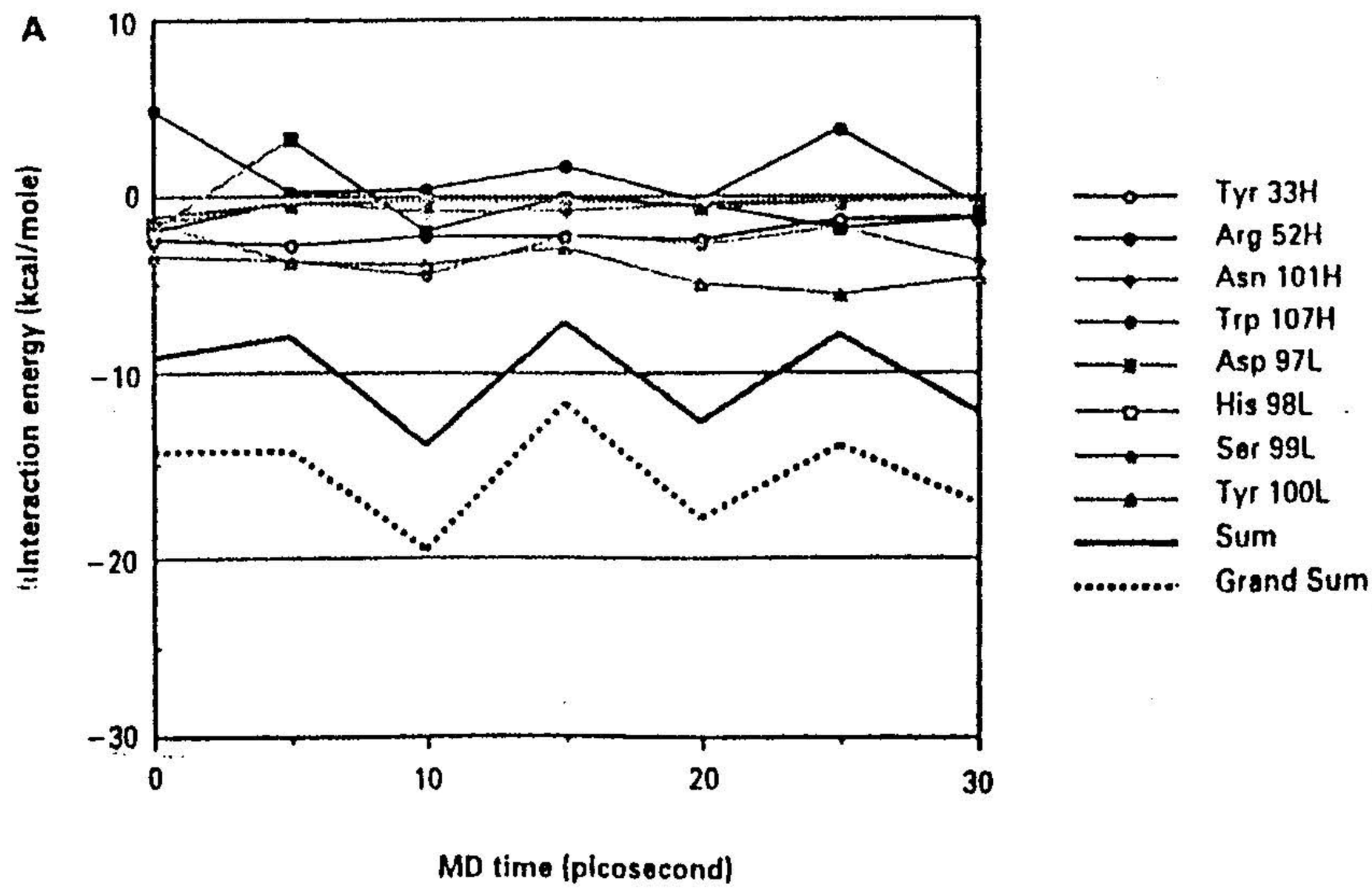
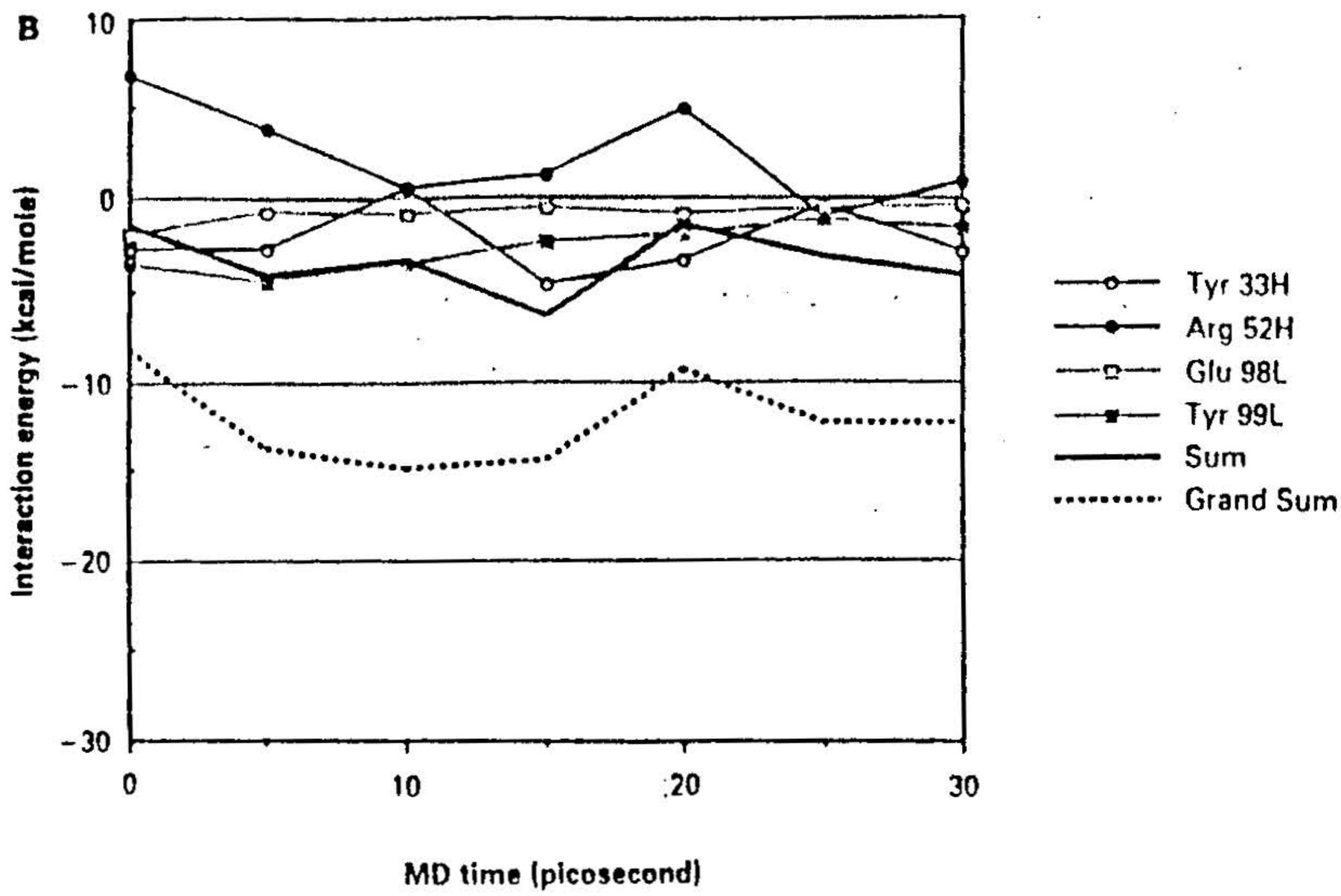


Fig. 7.—Van der Waals interaction energy (kcal/mol) for the three antibodies McPC603 (A). All parameters are as Fig. 6.



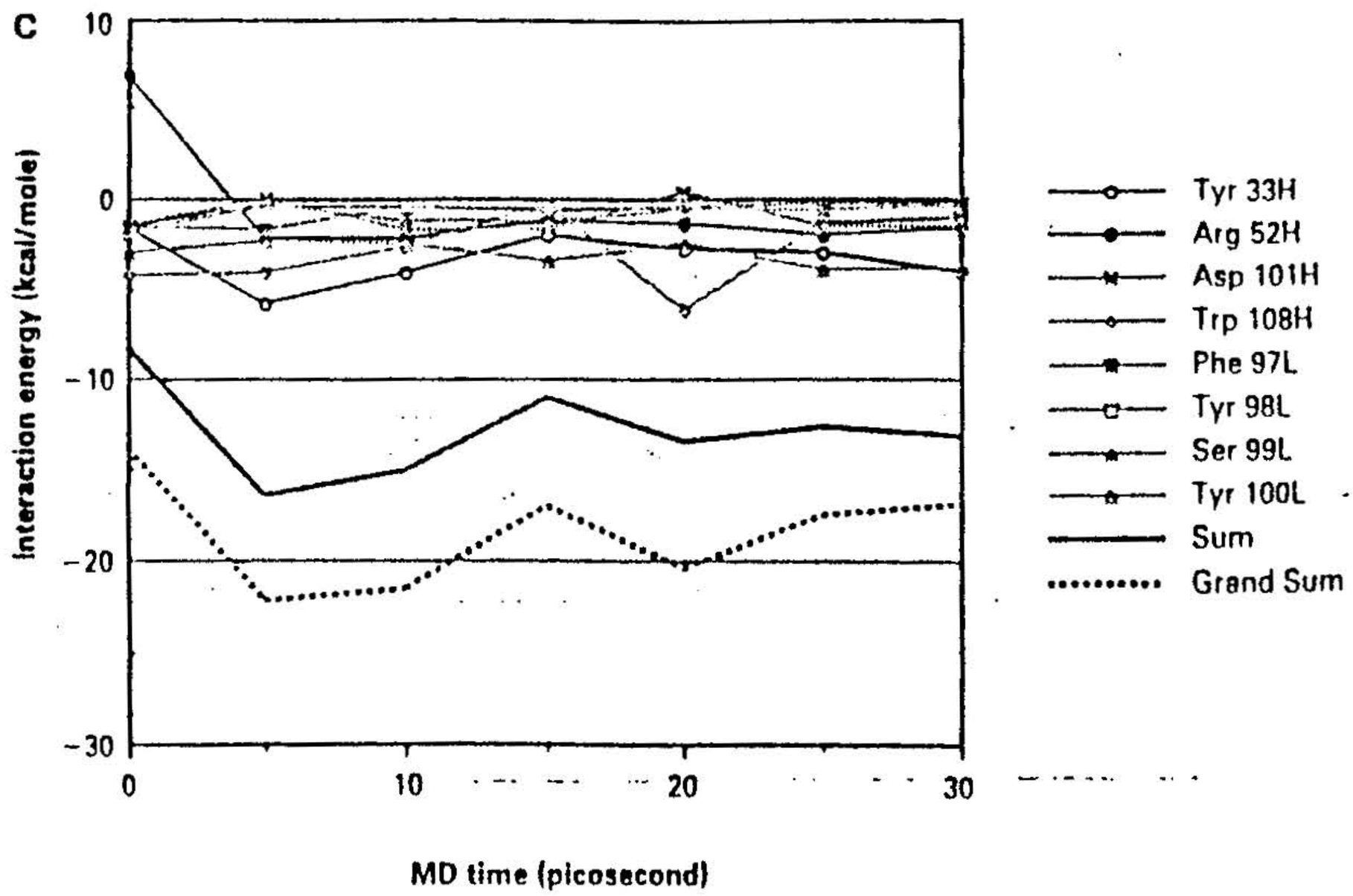
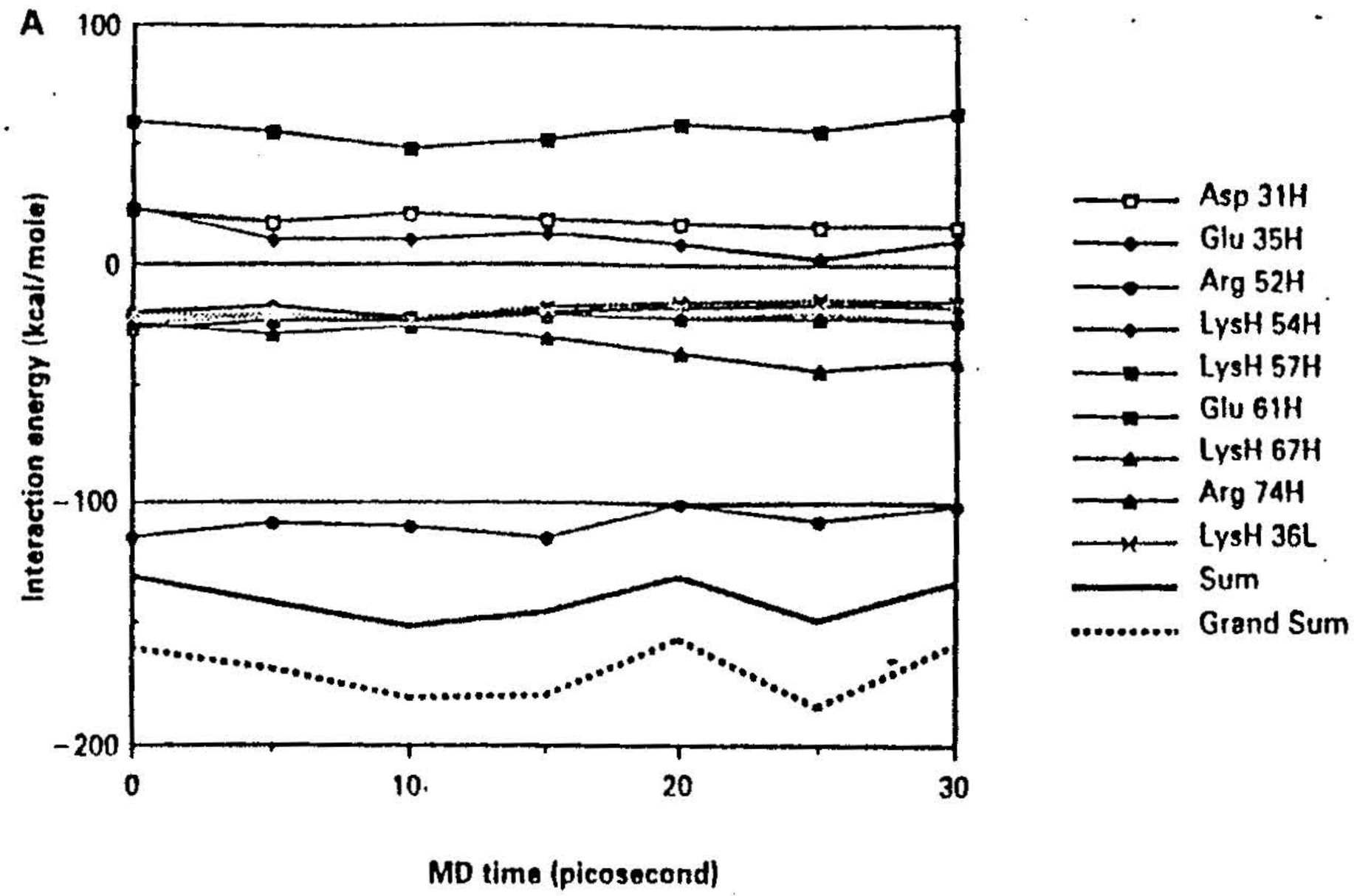


Fig. 7 *Contd.* — M167 (B) and T15 (C), each bound to (deprotonated) phosphorylcholine for the structures from the trajectories at 5 ps intervals. All parameters are as in Fig. 6.



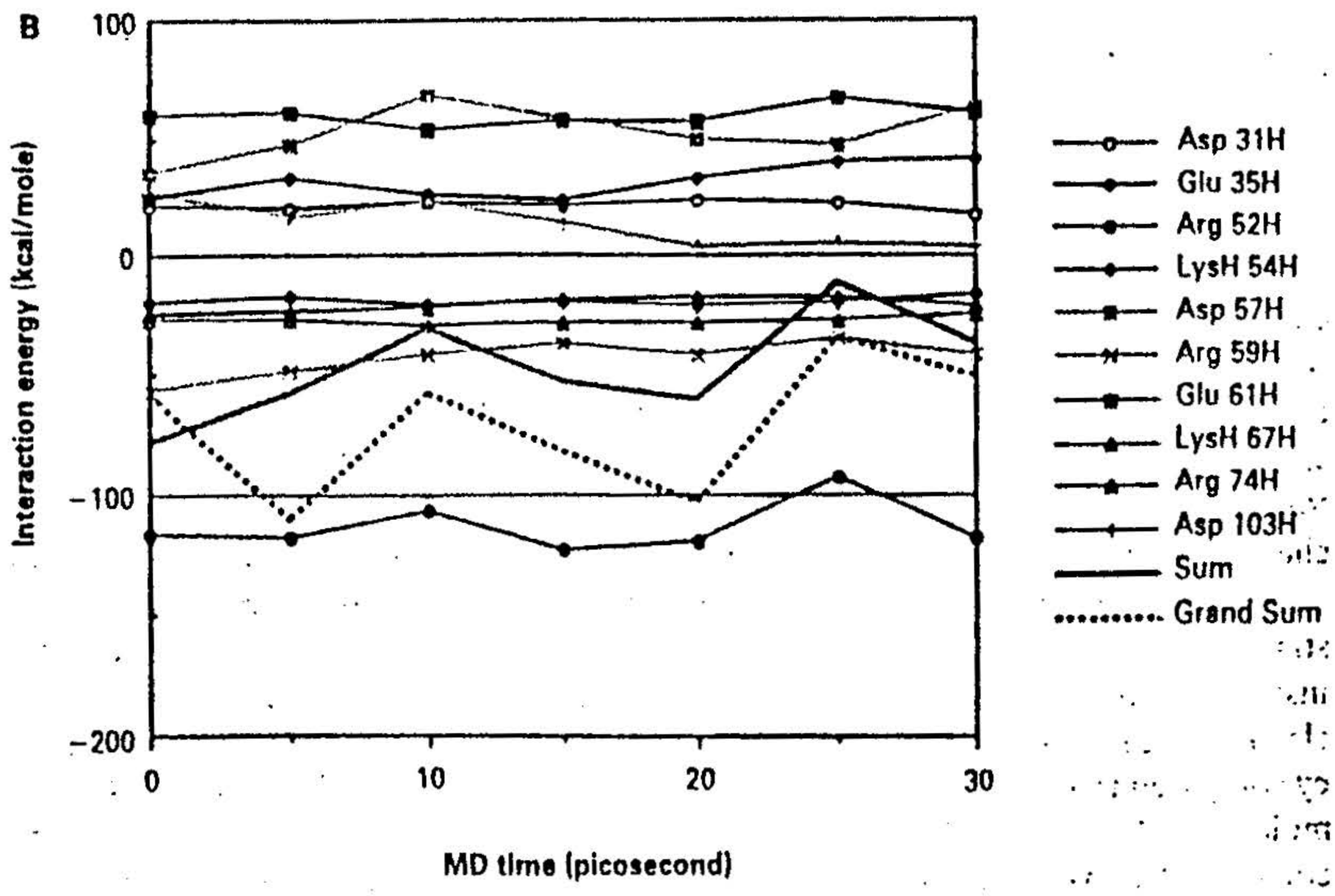


Fig. 8.—Total interaction energy (kcal/mol), which is the sum of the van der Waals and electrostatic energies for the three antibodies McPC603 (A), M167 (B) and T15. All parameters are as in Figs 6 and 7.

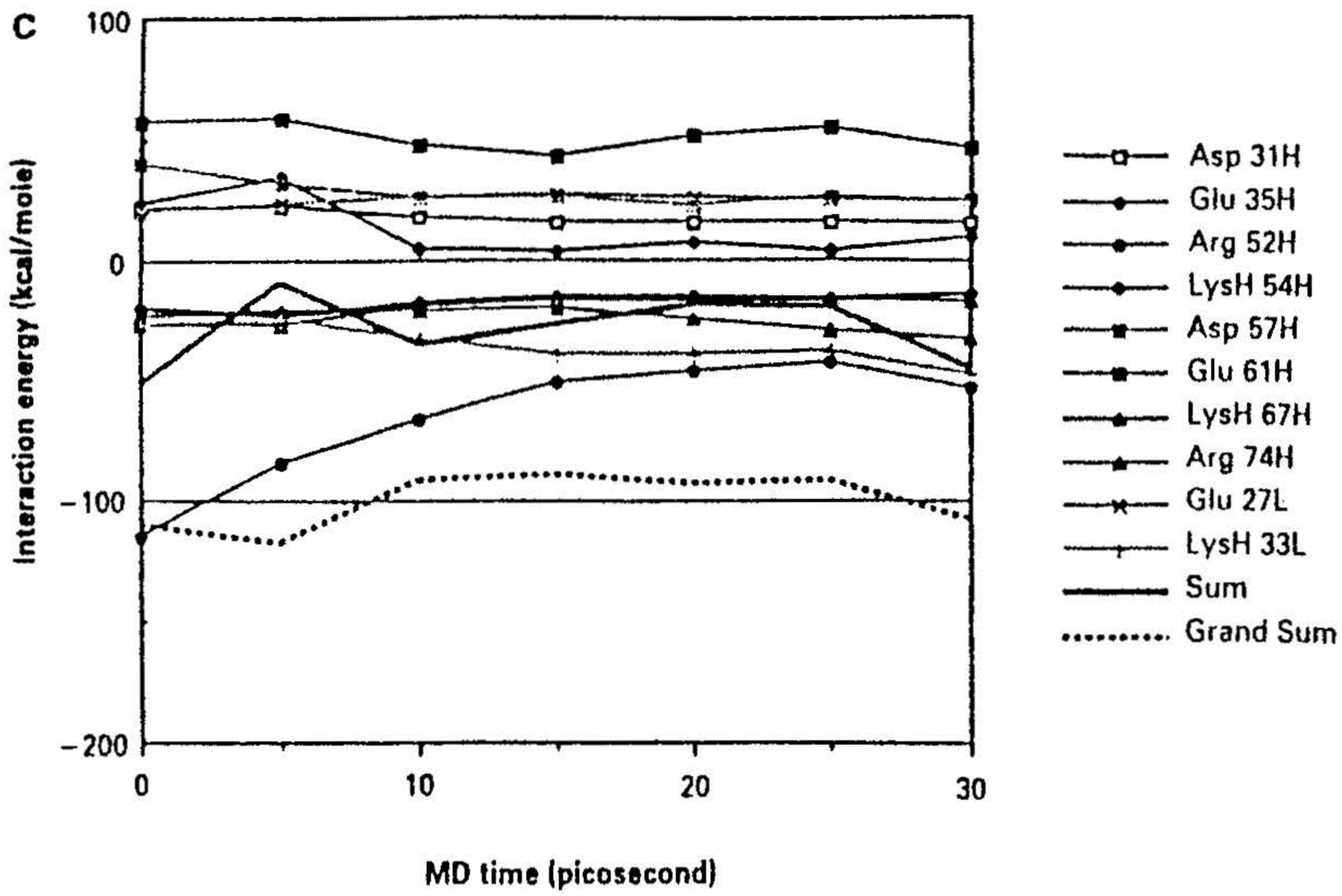


Fig. 8 *Contd.* — Total interaction energy (kcal/mol), which is the sum of the van der Waals and electrostatic energies for the three antibodies McPC603. (C), each bound to (deprotonated) phosphorylcholine for the structures from the trajectories at 5 ps intervals. All parameters are as in Figs 6 and 7.

FUTURE DIRECTIONS IN ELECTROSTATICS

A more detailed analysis of electrostatic interactions might be desirable in the future. Problems with the monopole approach become immediately clear if one realizes that the interaction between two charged atoms in molecular dynamics (using a dielectric constant of 1) levels off at a distance of approximately 30 Å. This means that a large cutoff radius of at least 8 Å, but better 12-15 Å, needs to be applied when the non-bonded interactions between an atom and its neighbors are computed. Therefore, the current molecular dynamics approaches use the so-called concept of charge groups. A charge group is a small group of atoms, chemically reasonably selected, that has a total net charge of preferentially zero. This concept is used to avoid creating artificial charges at the rim of the molecule when calculations with the periodic boundary conditions are performed. It also leads to a faster convergence of electrostatic forces in simulations. It has been shown [92] that the use of bond dipoles instead of atomic monopoles leads to a quite good long-range convergence of electrostatic forces, since the dipole interaction energies level off with a $1/r^3$ dependency compared to the convergence of monopole energies with $1/r$. For a set of 24 organic molecules, Williams [92] showed that their HF/6-31G** wavefunction potential could be fitted very well by a multipole distribution up to quadrupoles. If monopoles and dipoles were chosen, the fit was still very good. Moreover, it could be shown that the concept of introducing restricted bond dipole moments leads to a reproduction of the electric potential that was comparable to the pure monopole model.

Another aspect related to electrical charges is the polarizability of atoms and molecules. The effects of anisotropy and anharmonicity in crystallographic refinement and their relation to molecular dynamics have been studied [93]. The importance of including electronic polarization (e.g. electron correlation effects [94]) into model force fields has been discussed by van Duijnen and Rullman [95]. They applied their 'direct reaction field method' to the active site of the protease papain. The polarization influences molecular structures [96] and the overall results are better described by anisotropic atom-atom models. These models reproduce the behavior of π and lone-pair electrons in a better way than pure isotropic models [97]. A comparison of results for the SO_2 molecule from calculated intermolecular potentials of Lennard-Jones type and distributed multipole moments (Monopoles, Dipoles and Quadrupoles) for the electrostatic part leads to the conclusion that the second virial coefficient criterion is a necessary but not sufficient condition for a selection of atomic multipoles [98]. Water, especially with its high dielectric constant, is difficult to model in an efficient and computationally fast way.

High-speed algorithms for the calculation of atomic charges based upon the concept of orbital electronegativities have been proposed by several authors [99-102]. Dipole moments from partial equalization of orbital electronegativities have been derived [103] and a quantification of effective polarizabilities was proposed [104]. The introduction of these slightly modified concepts into the MM2 force field has been tested for 40 compounds and found to be in good agreement with experimental data [105].

As a first suggestion we think that one should clearly distinguish between improvements to analyze static structures and improvements that are related to the calculations during the trajectories. One could develop much more accurate programs including multipole terms etc. for the analysis of static structures obtained from experiment (X-ray) or from simulated trajectories. On the other hand, it might not be necessary to include the same high level of accuracy during the simulation of trajectories, since it has been shown that the most important structural features can be reproduced without. Here it might be more interesting to think about concepts and algorithms such that longer time scales might be accessible. Of course, if there should be enough computing time available, there might be a chance to also include the high-level concepts during the simulation.

CURRENT FRONTIERS OF RESEARCH FOR MOLECULAR DYNAMICS IN GENERAL

1. Attempts to exceed the picosecond time scales

One step foward is the attempt to use stochastic molecular dynamics in order to describe molecular processes that exceed the picosecond time scale [106]. Statistics can also be improved if several molecules are simulated simultaneously. An ingenious implementation has been used for studying the separation of CO from myoglobin [107]. In this work, a time-dependent Hartree approach was implemented into the molecular dynamics algorithm, so that within one simulation a system of one myoglobin molecule and about 60 CO molecules was investigated at one time. This

method solves 60 'normal' trajectories in one, since the CO molecules were defined as having no interactions between themselves.

In another example the time of the molecular process exceeds by far the period that can be simulated nowadays: the question of how proteins fold cannot be answered by simulations yet, because the folding process takes place in the range of seconds or even longer. Some trials on protein folding with the Monte Carlo method [7] have been performed recently [108], however using very much simplified molecular models.

Simulated annealing methods are under development [109-112], to solve three-dimensional structures based on NMR NOE distance constraints [113]. One of the basic problems is that there is no easy way to decide where the global minimum on a multidimensional energy-hypersurface is [114].

2. Improvements in analyzing multiconfigurational space

A combined molecular dynamics and X-ray refinement [115-119] improves on the understanding of the multiconfigurational problems (tyrosine flip etc.) occurring in crystal structures.

3. Attempts to describe polarization

The polarizability of molecules [120], especially the development of a polarizable water model, is of high interest [121-125]. Cooperative effects lead to extended network water structures [50, 126-130], which makes it hard to model water with pure pair potentials unless they are properly scaled to give effective models, as mentioned above.

4. Consistent force fields

A consistent set of atomic properties might be desirable: the need of an overall parametrization procedure for molecular dynamics force fields is becoming more and more important [131, 132]. The extension of parameters also to very heavy atoms has been investigated in the shell structure approach for neutral atoms from hydrogen to uranium and for several mono-charged positive ions from helium to barium and lutetium to radium by means of non-relativistic SCF wavefunctions and the resulting spherically averaged charge densities [133]. This survey is interesting in that a more general charge density approach covering almost the whole periodic system is discussed.

5. Attempts to model chemical reactions

Since in the current force fields no electrons are considered explictly, electron transfer reactions [134] or organic reactions involving covalent bond breaking and forming [135] as well as photochemical processes cannot be described directly [136, 137], although some attempts have been reported in the direction of including quantum mechanics into molecular dynamics [63, 138, 139]. Recently, the interest in non-

equilibrium molecular dynamics has been increasing [140], since the unsolved problems of finding possible structures on a folding pathway are of general interest in macromolecular systems.

6. New very fast computers with different architectures

The first step for speeding up computers was their vectorization. A Brownian dynamics simulation program has been written especially for a vector computer [141], and a polarizable water model can be computed on a vector computer Cyber 205 [123]. The development of very fast computers with new (parallel) architectures [142] and new programming languages [143] might help to calculate molecular systems [144] that are much larger or that need much longer simulation times than can now be handled.

CONCLUSION

Molecular dynamics trajectories display different behavior for molecular structures such as antibody/antigen complexes when started from model-built structures rather than from an X-ray structure.

In the case of McPC603/phosphorylcholine there was an X-ray structure available, whilst none was available for the two related complexes M167 and T15 bound to phosphorylcholine—only their sequences. We replaced individual amino acids and loop structures in the X-ray structure of McPC603 to obtain starting structures for M167 and T15. We performed one trajectory of 30 ps at 293 K for each of the three Fv-fragments in aqueous solution and analyzed the electrostatic interactions between all atoms of the most important amino acids of the proteins with respect to phosphorylcholine binding in the trajectory. The simulated structure of McPC603 was more or less equilibrated after 30 ps. For M167 and T15 we found larger fluctuations of the sensitive electrostatic energies and concluded that after 30 ps these structures were not in equilibrium. Clearly, one has to be careful with model structures built from existing X-ray structures. The change of even few amino acids might cause larger structural shifts of the protein than can be easily predicted today.

ACKNOWLEDGMENTS

Special thanks go to the authors (W.F. van Gunsteren, H.J.C. Berendsen) of the program package GROMOS 87, which we used to perform the molecular dynamics simulations. We thank B. Steipe for fruitful discussions and for providing the data from the protein structure analysis and the secondary structure.

We acknowledge the support of Cray Research Munich (D. Labrenz), the DLR Oberpfaffenhofen (H. Herchenbach, S. Leisen) and the computer center of the MPI Biochemie Martinsried, (W. Steigemann, G. Franz, H. Hanewinkel), for providing the necessary computing facilities, and Professor E. L. Winnacker and Hoechst Japan Ltd. for financial support.

REFERENCES

[1] L. Pauling, *The Nature of the Chemical Bond.* 3rd edn, Cornell University Press, Ithaca, New York (1960).
[2] L. Pauling, *Nature* **161** (1948) 707.
[3] L. Pauling and R.B. Corey, *Proc. Natl. Acad. Sci. USA* **37** (1951) 235.
[4] L. Pauling, *Chem. Eng. News* **24** (1946) 1375.
[5] D. van Belle, I. Couplet, S. Prevost and S. Wodak, *J. Mol. Biol.* **198** (1987).
[6] B. Jayaram, K.A. Sharp and B. Honig, *Biopolymers* **28** (1989) 975.
[7] W.L. Jorgensen, *J. Phys. Chem.* **87** (1983) 5304.
[8] M.M. Karelson, A.R. Katritzky and M.C. Zerner, *Int. J. Quant. Chem.* **20** (1986) 521.
[9] I. Rips, J. Klafter and J. Jortner, *J. Chem. Phys.* **88** (1988) 3246.
[10] W.S. Bennett and R. Huber, *CRC Crit. Rev. Biochem.* **15** (1984) 291.
[11] H.A. Scheraga, *Adv. Phys. Org. Chem.* **6** (1968) 103.
[12] N.L. Allinger and J.T. Sprague, *J. Am. Chem. Soc.* **94** (1972) 5734.
[13] N.L. Allinger and J.T. Sprague, *J. Am. Chem. Soc.* **95** (1973) 3983.
[14] N.L. Allinger and J.C. Graham, *J. Am. Chem. Soc.* **95** (1973) 2523.
[15] A.T. Hagler, E. Huler and S. Lifson, *J. Am. Chem. Soc.* **96** (1974) 5319.
[16] M. Karplus and J.A. McCammon, *CRC Crit. Rev. Biochem.* **9** (1981) 293.
[17] P.K. Weiner and P.A. Kollman, *J. Comp. Chem.* **2** (1981) 287.
[18] J. Hermans, H.J.C. Berendsen W.F. van Gunsteren and J.P.M. Postma, *Biopolymers* **23** (1984) 1513.
[19] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, *J. Comp. Chem.* **4** (1983) 187.
[20] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner, *J. Am. Chem. Soc.* **106** (1984) 765.
[21] S.J. Weiner, P.A. Kollman, D.T. Nguyen and D.A. Case, *J. Comp. Chem.* **7** (1986) 230.
[22] P.A. Kollman, *Acc. Chem. Res.* **18** (1985) 105.
[23] P.A. Kollman, *Methods in Enzymol.* **154** (1987) 430.
[24] D.L. Beveridge and W.L. Jorgensen (Eds), Computer Simulation of Chemical and Biomolecular Systems. *Ann. N.Y. Acad. Sci.* **482** (1986) 1.
[25] M. Karplus and J.A. McCammon, *Sci. Am.* **254** April (1986) 42.
[26] L. Nilson and M. Karplus, *J. Comp. Chem.* **7** (1986) 591.
[27] J.A. McCammon and S.C. Harvey, *Dynamics of Proteins and Nucleic Acids.* Cambridge University Press (1987).
[28] C.L. Brooks, M. Karplus and B.M. Pettitt (Eds), *Proteins: a Theoretical Perspective of Dynamics, Structure and Thermodynamics. Advances in Chemical Physics* LXXI, John Wiley & Sons (1988).
[29] W.F. van Gunsteren, *Protein Engineering* **2** (1988) 5.
[30] W.L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110** (1988) 1657.
[31] W.F. van Gunsteren and P.K. Weiner (Eds), Computer Simulations of Biomolecular Systems. Theoeretical and Experimental Applications. *ESCOM*, Leiden, The Netherlands (1989).
[32] W.F. van Gunsteren and H.J.C. Berendsen, *Angew. Chem.* **29** (1990) 992.
[33] J.M. Goodfellow (Ed.), *Molecular Dynamics. Applications in Molecular Biology.* The Macmillan Press, Basingstoke, UK (1991).

[34] D.A. McQuarrie, *Statistical Mechanics*. Harper & Row, Harper's Chemistry Series, S.A. Rice (Ed.) (1976).
[35] L. Verlet, *Phys. Rev.* **159** (1967) 98.
[36] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice Hall, Englewood Cliffs, N.J. (1971).
[37] W.F. van Gunsteren and H.J.C. Berendsen, *Mol. Phys.* **34** (1977) 1311.
[38] J.P. Ryckaert, G. Ciccotti and H.J.C. Berendsen, *J. Comp. Phys.* **23** (1977) 327.
[39] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola and J.R. Haak, *J. Chem. Phys.* **81** (1984) 3684.
[40] H.J.C. Berendsen and W.F. van Gunsteren, Practical Algorithms for Dynamic Simulations. In: G. Ciccotti and W.G. Hoover. (Eds). *Proc. Enrico Fermi School of Physics*, Varenna, North Holland Physics Publishing (1986) 43.
[41] R.W. Pastor, B.R. Brooks and A. Szabo, *Mol. Phys.* **65** (1988) 1409.
[42] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren and J. Hermans, In: Pullman B, (Ed.). *Intermolecular Forces*. D. Reidel Publishing Company (1981) 331.
[43] H.J.C. Berendsen and W.F. van Gunsteren, Molecular Dynamics Simulations: Techniques and Approaches. Molecular Liquids-Dynamics and Interactions. In: A.J. Barnes *et al.* (Eds). *NATO-ASI Series*, **C135** (1984) 475.
[44] W.F. van Gunsteren and H.J.C. Berendsen, *Molec. Simul.* **1** (1988) 173.
[45] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.* **79** (1983) 926.
[46] A.T. Hagler, J.R. Maple, T.S. Thacher, G.B. Fitzgerald and U. Dinur, Potential Energy Functions for Organic and Biomolecular Systems. In: *Computer Simulation of Biomolecular Systems*, W.F. van Gunsteren and P.K. Weiner (Eds), *ESCOM*, Leiden (1989).
[47] J.E.H. Koehler, W. Saenger and W.F. van Gunsteren, *Eur. Biophys. J.* **15** (1987) 197.
[48] J.E.H. Koehler, W. Saenger and W.F. van Gunsteren, *Eur. Biophys. J.* **15** (1987) 211.
[49] P.A. Kollman and L.C. Allen, *Chem. Rev.* **72** (1972) 283.
[50] P. Barnes, J.L. Finney, J.D. Niocholas and J.E. Quinn, *Nature* **282** (1979) 459.
[51] J.E.H. Koehler, B. Lesyng and W. Saenger, *J. Comp. Chem.* **8** (1987) 1090.
[52] G.A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures*. Springer Verlag, Berlin, Heidelberg, New York (1990).
[53] W. Saenger, *Nature* **279** (1979) 343.
[54] W. Saenger, C. Betzel, B. Hingerty and G.M. Brown, *Nature* **296** (1982) 581.
[55] J.E.H. Koehler, W. Saenger and W.F. van Gunsteren, *Eur. Biophys. J.* **16** (1988) 153.
[56] J.E.H. Koehler, W. Saenger and W.F. van Gunsteren, *J. Biomol. Struc. Dyn.* **6** (1988) 181.
[57] J.E.H. Koehler, W. Saenger and W.F. van Gunsteren, *J. Mol. Biol.* **203** (1988) 241.
[58] A. Warshel, F. Sussman and J.K. Hwang, *J. Mol. Biol.* **201** (1988) 139.
[59] T.P. Straatsma and H.J.C. Berendsen, *J. Chem. Phys.* **89** (1988) 5876.
[60] G. Ciccotti, M. Ferrario, J.T. Hynes and R. Kapral, *J. de Chimie Physique* **85** (1988) 925

[61] A. Warshel, G. Naray-Szabo, F. Sussman and J.K. Hwang, *Biochemistry* **28** (1989) 3629.
[62] P.A. Kollman and W.F. van Gunsteren, *Methods in Enzymol.* **154** (1987) 430.
[63] P.A. Bash, M.J. Field and M. Karplus, *J. Am. Chem. Soc.* **109** (1987) 8092.
[64] U.C. Singh, F.K. Brown, P.A. Bash and P.A. Kollman, *J. Am. Chem. Soc.* **109** (1987) 1607.
[65] P. Cieplak and P.A. Kollman, *J. Am. Chem. Soc.* **110** (1988) 3734.
[66] M.N. Jones (Ed.). *Biochemical Thermodynamics.* Elsevier, Amsterdam-Oxford-New York-Tokyo (1988).
[67] W.F. van Gunsteren and H.J.C. Berendsen, *J. Comput.-Aided Mol. Design* **1** (1987) 171.
[68] P.A. Bash, U.C. Singh, F.K. Brown, R. Langridge and P.A. Kollman, *Science* **235** (1987) 574.
[69] E. Lolis and G.A. Petsko, *Ann. Rev. Biochem.* **59** (1990) 597.
[70] R. Wolfenden, *Ann. Rev. Biophys. Bioeng.* **5** (1976) 271.
[71] W.P. Jencks, *Catalysis in Chemistry and Enzymology.* McGraw-Hill, New York (1969).
[72] P.G. Schultz, *Angew. Chem. Int. Ed. Engl.* **28** (1989) 1283.
[73] K.M. Shokat and P.G. Schultz, *Ann. Rev. Immunol.* **8** (1990) 335.
[74] R.A. Lerner and S.J. Bencovic, *Bioessays* **9** (1988) 107.
[75] G.M. Blackburn, A.S. Kang, G.A. Kingsbury and D.R. Burton, *Biochem. J.* **262** (1989) 381.
[76] D.M. Segal, E.A. Padlan, G.H. Cohen, S. Rudikoff, M. Potter and D.R. Davies, *Proc. Natl. Acad. Sci. USA* **71** (1974) 4298.
[77] Y. Satow, G.H. Cohen, E.A. Padlan and D.R. Davies, *J. Mol. Biol.* **190** (1986) 593.
[78a] Euranto, E.K., Esterification and Ester Hydrolysis. In: S. Patai (Ed.), The Chemistry of Carboxylic Acids and Esters. The Chemistry of Functional Groups (series), Interscience Publishers, London, 1969.
[78b] U. Holzgrabe and G. Bejeuhr, *Pharmazie in unserer Zeit* **6** (1990) 247.
[79] J. Stadlmüller and A. Plückthun, (1991), in preparation.
[80] R. Glockshuber, J. Stadlmüller and A. Plückthun, *Biochemistry* **30** (1990) in press.
[81] A. Plückthun, R. Glockshuber, A. Skerra and J. Stadmüller, *Behring Inst. Mitt.* **87** (1990) 48.
[82] S.J. Pollack, J.W. Jacobs and P.G. Schultz, *Science* **234** (1986) 1570.
[83] S.J. Pollack and P.G. Schultz, *Cold Spring Harbor Symp. Quant. Biol.* **52** (1987) 97.
[84] A. Skerra and A Plückthun, *Science* **240** (1988) 1038.
[85] A. Plückthun, *Nature* **347** (1990) 497.
[86] A. Plückthun, *Biotechnology* **9** (1991) 545.
[87] W. Kabsch and C. Sander, *Biopolymers* **22** (1983) 2577.
[88] R. Kikuno and J.E.H. Koehler, PROELW program (1990).
[89] J. Novotny, R.E. Bruccoleri and F.A. Saul, *Biochemistry* **28** (1989) 4735.
[90] C. Köhler, Diploma Thesis, Physics Dept. TU Munich, (1990).
[91] J. Köhler, C. Köhler, V. Helms and K. Adelhard, *Croat. Chem. Acta.* (1991) in press.

[92] D.E. Williams, *J. Comp. Chem.* **9** (1988) 745.
[93] J. Kuriyan, G.A. Petsko, R.M. Levy and M. Karplus, *J. Mol. Biol.* **190** (1986) 227
[94] H.J. Werner and W. Meyer, *Molecular Physics* **31** (1979) 855.
[95] P.T. van Duijnen and J.A.C. Rullman, *Int. J. Quant. Chem.* **XXXVIII** (1990) 181
[96] S. Wolfe, *Acc. Chem. Res.* **5** (1972) 102.
[97] S. Price, *Mol. Simulat.* **1** (1988) 135.
[98] M. Pavlovic, F. Sokolic and Z.B. Maksic, *J. Mol. Struc. (Theochem.)* **202** (1989) 265.
[99] J. Gasteiger and M. Marsili, *Tetrahedron* **36** (1980) 3219.
[100] W.J. Mortier, K. van Genechten and J. Gasteiger, *J. Am. Chem. Soc.* **107** (1985) 829.
[101] L. Baumer, G. Sala and G. Sello, *Tetrahedron Comp. Methodol.* **2** (1989) 37.
[102] J. Mullay, *J. Comp. Chem.* **12** (1991) 369.
[103] J. Gasteiger and M.D. Guillen, *J. Chem. Res. (S)*, (1983) 304.
[104] J. Gasteiger and M.G. Hutchings, *J. Chem. Soc. Perkin Trans.* **II** (1984) 559.
[105] L.G. Hammarstroem, T. Liljefors and J. Gasteiger, *J. Comp. Chem.* **9** (1988) 424.
[106] B. Cartling, *J. Chem. Phys.* **91** (1989) 427.
[107] R. Elber and M. Karplus, *J. Am. Chem. Soc.* **112** (1990) 9161.
[108] J. Skolnick and A. Kolinski, *Science* **250** (1990) 1121.
[109] A.T. Bruenger, *J. Mol. Biol.* **203** (1988) 803.
[110] M. Nilges, G.M. Clore and A.M. Gronenborn, *FEBS Lett.* **239** (1988) 129.
[111] M. Nilges, G.M. Clore and A.M. Gronenborn, *FEBS Lett.* **239** (1988) 317.
[112] A.T. Bruenger, *Acta Cryst. A* **45** (1989) 42.
[113] G.M. Clore, A.T. Brunger, M. Karplus and A.M. Gronenborn, *J. Mol. Biol.* **191** (1986) 523.
[114] A. Howard and P.A. Kollman, *J. Med. Chem.* **31** (1988) 1669.
[115] A.T. Bruenger, M. Karplus and G.A. Petsko, *Acta Cryst. A* **45** (1989) 50.
[116] M. Fujinaga, P. Gros and W.F. van Gunsteren, *J. Appl. Cryst.* **22** (1989) 1.
[117] A.E. Torda, R.M. Scheek and W.F. van Gunsteren, *Chem. Phys. Lett.* **157** (1989) 289.
[118] A.E. Torda, R.M. Scheek and W.F. van Gunsteren, *J. Mol. Biol.* **214** (1990) 223.
[119] P. Gros, W.F. van Gunsteren and W.G.J. Hol, *Science* **249** (1990) 1149.
[120] B.T. Thole, *Chem. Phys.* **59** (1981) 341.
[121] F.H. Stillinger and C.W. David, *J. Chem. Phys.* **69** (1978) 1473.
[122] P. Ahlstrom, A. Wallqvist, S. Engstrom and B. Jonsson, *Mol. Phys.* **68** (1989) 563.
[123] J.C. Sauniere, T.P. Lybrand, J.A. McCammon and L.D. Pyle, *Computers Chem.* **13** (1989) 313.
[124] J. Caldwell, L.X. Dang and P.A. Kollman, *J. Am. Chem. Soc.* **112** (1990) 9144.
[125] F.A. Webster, J. Schnitker, M. Friedrichs, R.A. Friesner and P.J. Rossky, *Phys. Rev. Letters* **66** (1991) 3172.
[126] A.H. Narten and H.A. Levy, *J. Chem. Phys.* **55** (1971) 2263.
[127] L.L. Shipman and H.A. Scheraga, *J. Phys. Chem.* **78** (1974) 909.
[128] M. Mezei and D.L. Beveridge, *J. Chem. Phys.* **76** (1982) 593.

[129] R.J. Speedy, J.D. Madura and W.L. Jorgensen, *J. Phys. Chem.* **91** (1987) 909.
[130] F. Franks (Ed.), *Water Science Review 3*, Cambridge University Press (1988).
[131] M.J. Field, P.A. Bash and M. Karplus, *J. Comp. Chem.* **11** (1990) 700.
[132] K. Palmö, L.O. Pietilä and S. Krimm, *J. Comp. Chem.* **12** (1991) 385.
[133] R.P. Sagar, A.C.T. Ku and V.H. Smith, *J. Chem. Phys.* **88** (1988) 4367.
[134] D.A. Zichi, G. Ciccotti, J.T. Hynes and M. Ferrario, *J. Phys. Chem.* **93** (1989) 6261.
[135] W.L. Jorgensen, *Acc. Chem. Res.* **22** (1989) 184.
[136] H. Treutlein, K. Schulten, J. Deisenhofer, H. Michl, A. Bruenger and M. Karplus. In: *The Photosynthetic Bacterial Reaction Center: Structure and Dynamics.* J. Breton and A. Vermeglio (Eds.), Plenum Press, London (1987) 139.
[137] H. Treutlein, K. Schulten, C. Niedermeyer, J. Deisenhofer, H. Michl and D. de Vault. In: *The Photosynthetic Bacterial Reaction Center: Structure and Dynamics.* J. Breton and A. Vermeglio (Eds), Plenum Press, London (1987) 369.
[138] C. Zheng, C.F. Wong, J.A. McCammon and P.G. Wolynes, *Nature* **334** (1988) 726.
[139] R. Kosloff, *J. Phys. Chem.* **92** (1988) 2087.
[140] C. Massobrio, V. Pontikis and G. Ciccotti, *Phys. Rev. B* **39** (1989) 2640.
[141] C. Elvingson, *J. Comp. Chem.* **12** (1991) 71.
[142] P.J. Denning and W.F. Tichy, *Science* **250** (1990) 1217.
[143] A. Burns, *Programming in OCCAM 2.* Addison Wesley (1988).
[144] J.M. Goodfellow and F. Vovelle, *Eur. Biophys. J.* **17** (1989) 167.

Programs

AM1-Program, AMPAC Program Package, v 2.1. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P. (1985). AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **107**, 3902-3909.

AMBER-Program, Version 3.0 (1987). Singh, U.C., Weiner, P.K., Caldwell, J. and Kollman, P.A., Dept. of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143.

CHARMM. Karplus, M., Harvard University, Dept. of Chemistry, 12 Oxford Street, Cambridge, Mass. 02138, USA.

CHELP. Chirlian, L.E. and Francl, M.M. (1985). Quantum Chemistry Program Exchange, QCPE 524

DISCOVER, published by Biosym Technologies, Inc., 9605 Scranton Rd., Suite 101, San Diego, CA 92121.

GAUSSIAN 86. Frisch, M.J., Binkley, J.S., Schelgel, H.B., Raghavachari, K., Melius, C.F., Martin, L., Stewart, J.J.P., Bobrowicz, F.W., Rohlfing, C.M., Kahn, L.R., Defrees, D.J., Seeger, R., Whiteside, R.A., Fox, D.J., Fleuder, E.M. and Pople, J.A. (1984). *Carnegie Mellon Quantum Chemistry Publishing Unit*, Pittsburgh PA.

GROMOS. Biomos, University of Groningen, 9747 AG Groningen, Nijenborgh 16, The Netherlands; Biomos, ETH-Zentrum, 8092 Zürich, Switzerland.

OCCAM2. Product Definition (preliminary). *INMOS, Bristol* and *Prentice Hall.* Hemel Hempstead, England.

INSIGHT. Biosym Technologies, 10065 Barnes Canyon Road, San Diego, CA 92121. PM3-Program. Stewart, J.J.P. (1983). MOPAC Program Package, QCPE Nr. 455.

PROBE. Hagler *et al.* (1991). Biosym Technologies, 10065 Barnes Canyon Road, San Diego, CA 92121.

QUEST. Singh, U.C. and Kollman, P.A. (1984). An Approach to Computing Electrostatic Charges for Molecules. *J. Comp. Chem.* **2**, 129–145.

X-PLOR. Bruenger, A. Howard Hughes Medical Institute, Dept. Mol. Biophys. Biochem., Yale Univ., New Haven, CT 06511, USA.