THEORETICAL CALCULATIONS AND DRUG DESIGN

P. Gund and H. B. Schlegel

Merck, Sharp & Dohme Research Laboratories
Rahway, New Jersey 07065

INTRODUCTION

It has been more than twenty years since the Pullmans undertook a systematic quantum mechanical analysis of biologically important molecules.[1] Yet only recently have such calculations become well accepted by practicing medicinal chemists. There are several reasons for this long delay. Technological progress was required in the areas of high-speed computers and quantum mechanical theory in order to permit sufficiently accurate calculations to be made on molecules large enough to be biologically interesting. Time was also required to establish a dialogue between theoreticians and experimentalists. Practicing chemists had to learn the new language of frontier orbitals, symmetry allowed reactions, and charge transfer effects. Theoreticians needed the feedback of well-designed experiments to test and refine the calculational results.

At Merck, we have involved medicinal chemists in theoretical calculations in a fairly painless way. An interactive graphics system was developed for creating and viewing three-dimensional molecular structures and for performing semi-empirical calculations. Based on their initial results with this system, Merck scientists requested a staff theoretician for consultation and to perform more rigorous studies. The remainder of this article highlights the development of theoretical capabilities at Merck and applications of these capabilities in two classes of drugs.

THE MERCK MOLECULAR MODELING SYSTEM (MMMS)

MMMS was created to help medicinal chemists correlate molecular shape with biological activity.[2] The system uses DEC GT43 display terminals, which communicate with the corporate IBM 370/168 computer over telephone lines. MMMS has several molecule input programs for the creation of three-dimensional molecular structures from crystal coordinates, from standard bond lengths and angles, or from two-dimensional structural diagrams. Other programs display the structure stereoscopically as a stick figure or as a space-filling model and allow molecules to be superimposed. Classical and quantum mechanical programs are used to compute relative energies, find stable conformations, and optimize geometries. Finally, calculations of physical properties (dipole moments, electrostatic contour maps, conformational energy maps, etc.) are routinely available.

Using MMMS, the behavior of a molecule can be described by a heirarchy of calculational methods, ranging from simple classical mechanical models to accurate and complex ab initio quantum theory. Drug molecules tend to be so large that exhaustive studies by quantum mechanical methods, especially at a

0077-8923/81/0367-0510 $1.75/0

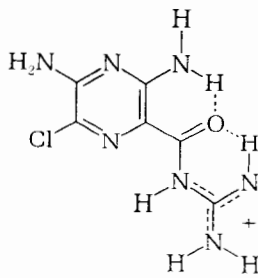
rigorous level, are prohibiti
methods are much less co
containing many heteroator
there are many approxima
practice, most problems mu
one starts with the least sop
extends the study to increasi
structures) for testing and c
agree, considerable confiden

The chemists soon found
range of theoretical calculati
programs. On discovering th
these studies, they requested
chemist as an internal consu
were necessary to calibrate t
of effort has been spent re
mechanical programs[3] into M
upgrade these programs to
large drug molecules. With M
directly to obtain reliable cl
energies. More detailed stud

A

In this example of the use
spectroscopists, and quantu
structural problem. Experim
mechanism; higher quality
with semiempirical results.

Amiloride is a diuretic ag
sodium while sparing potassi
form existed in the first co
existed as an equilibrium mi



(1)

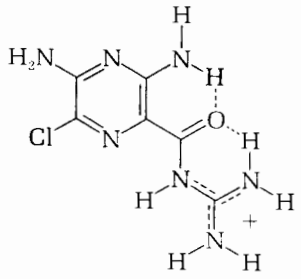
rigorous level, are prohibitively expensive. On the other hand, while classical methods are much less costly, reliable empirical parameters for structures containing many heteroatoms are not yet available. Between these extremes there are many approximate methods of varying accuracy and expense. In practice, most problems must be approached from a variety of levels. Usually, one starts with the least sophisticated method that can model the problem and extends the study to increasingly accurate calculations (often on simplified model structures) for testing and calibration. If the results of the various approaches agree, considerable confidence can be placed in the outcome of the study.

The chemists soon found that MMMS made it extremely easy to perform a range of theoretical calculations with the classical mechanical and CNDO-INDO programs. On discovering that they needed advice in interpreting the results of these studies, they requested (and received) the services of a resident quantum chemist as an internal consultant. It became clear that more rigorous calculations were necessary to calibrate the semiempirical results. Consequently, a great deal of effort has been spent recently to incorporate the latest ab initio quantum mechanical programs[3] into MMMS. Considerable additional work was needed to upgrade these programs to handle special problems encountered in modeling large drug molecules. With MMMS, a chemist can carry out ab initio calculations directly to obtain reliable charge distributions, molecular orbitals, and relative energies. More detailed studies are performed as a service.

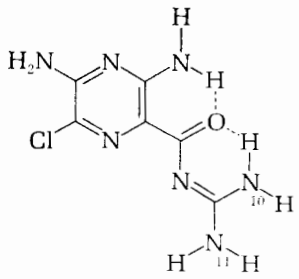
Amiloride Conformation

In this example of the use of MMMS, cooperation among medicinal chemists, spectroscopists, and quantum chemists was crucial to sorting out a thorny structural problem. Experiments were devised to test a theoretically derived mechanism; higher quality calculations were used to resolve problems found with semiempirical results.

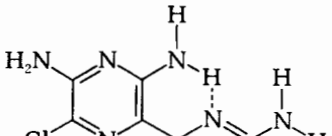
Amiloride is a diuretic agent possessing the desirable property of eliminating sodium while sparing potassium. A CNDO/2 study suggested that the protonated form existed in the first conformation shown below, **(1)**, while the free base existed as an equilibrium mixture of two forms, (**2**) and (**3**).[4]



(1)

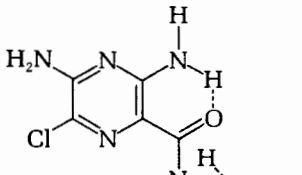


(2)

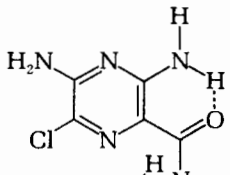


(3)

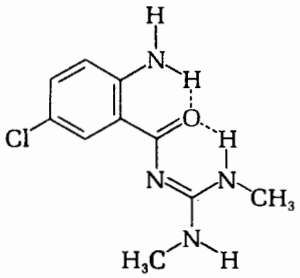
It was observed that the two terminal nitrogens (N_{10} and N_{11}) of the guanidine moiety were equivalent in the ^{15}N-nuclear magnetic resonance (nmr) spectrum and that substituents on these nitrogens were equivalent by ^{1}H- and ^{13}C-nmr. This indicated that amiloride was undergoing dynamic conformational changes. CNDO/2 calculations suggested that equilibration by simple C=N double bond rotations (via (**4**)) was relatively high in energy (23 kcal) and should be slow on the nmr time scale, while an alternative double rotation (via (**5**)) was more favorable, presumably due to an attractive intramolecular nonbonded interaction in (**5**).



(4)



(5)



(6)

The carbocyclic analogue (**6**) was then synthesized to test the mechanistic proposal and was found to equilibrate just as easily, although, in this case, the transition state corresponding to (**5**) was calculated to be high in energy. Thus, the CNDO/2 calculations apparently over-emphasized both the barrier of C=N

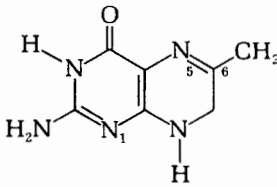
rotation in (**4**) and the nonb
are known from other studie

In an extension of this w
CNDO/2 calculated proton
However, in studies with
obtained. A more careful st
several model systems. Com
idinium ion provided a moc
upon protonation. Based on
on an extensive series of
optimal geometries, afforded
tion energies and experimer

DIHYDROFOL

The mode of action of a
and anticancer drugs involv
Some years ago, this enzym
design project at Merck.

One question raised duri
tion site of dihydrofolic acid
tion gave a substantially mo
charge density at N_1) and th
the lowest unoccupied mol
attack occurred during enzy
was confirmed by ultravio
during reduction rather than



(7)

H_2N

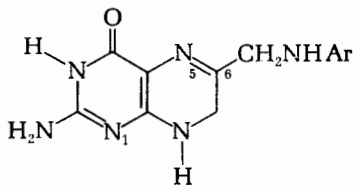
rotation in (**4**) and the nonbonded attractive interaction in (**5**); both shortcomings are known from other studies.[4]

In an extension of this work, a preliminary correlation was obtained between CNDO/2 calculated protonation energies and experimentally measured pK_as.[5] However, in studies with larger numbers of derivatives, no correlation was obtained. A more careful study using ab initio methods was then undertaken on several model systems. Complete geometry optimization of guanidine and guanidinium ion provided a model for the structural changes that occur in amiloride upon protonation. Based on these findings, a new set of CNDO/2 computations on an extensive series of amiloride derivatives, using the ab initio-derived optimal geometries, afforded a practical correlation between calculated protonation energies and experimentally determined pK_as.[5]

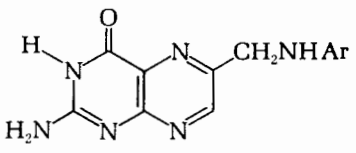
Dihydrofolate Reductase (DHFR) Inhibitors

The mode of action of a group of clinically useful antibiotics, antiparasitics, and anticancer drugs involves the inhibition of dihydrofolate reductase (DHFR). Some years ago, this enzyme was chosen as a target system for a "rational" drug design project at Merck.

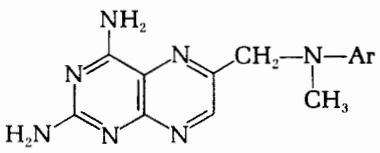
One question raised during these studies was, What is the preferred protonation site of dihydrofolic acid (**7**)? CNDO/2 calculations suggested that N_5 protonation gave a substantially more stable cation that N_1 protonation (despite higher charge density at N_1) and that the resulting N_5-protonated species had ~50% of the lowest unoccupied molecular orbital amplitude at C_6, where hydride-like attack occurred during enzyme-promoted reduction.[6] Preferential N_5 protonation was confirmed by ultraviolet studies, although protonation appears to occur during reduction rather than upon binding to the enzyme.[7]



(**7**)



(**8**)



(**9**)

Interestingly, both folic acid (**8**) and methotrexate (**9**) protonate preferentially at N_1 according to CNDO/2 calculations, which were confirmed both by ab initio calculations and by experiment. This seems to have significance for the biological action; recent studies show that dihydrofolate must lie on the enzyme active site with the pteridine ring "flipped" relative to the orientation of methotrexate in the bound enzyme crystal structure.[8]

The x-ray structure of the enzyme-inhibitor complex provided the starting point for further theoretical investigations into the nature of the methotrexate-DHFR interaction. Even with the fastest methods, only an extremely simplified model can be examined. Specifically, attention was focused on the three most important interactions of the enzyme with the pteridine ring of N_1 protonated methotrexate, as shown in FIGURE 1.[9] Ab initio calculations indicate that the salt bridge between the protonated N_1 of methotrexate and the carboxylate ion of Asp-27 is strongest. Furthermore, this primarily electrostatic interaction raises the energy of the highest occupied molecular orbital (HOMO) and, to a lesser extent,

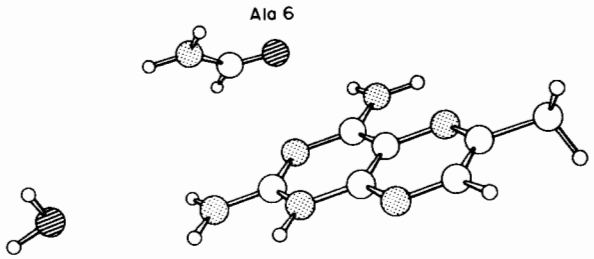


FIGURE 1. A model for the dihydrofolate reductase binding site of methotrexate.

the lowest unoccupied molecular orbital (LUMO), thus reducing the HOMO-LUMO gap. This accounts semiquantitatively for the hitherto unexplained shift to longer wavelengths of the lowest frequency ultraviolet absorbance between N_1 protonated methotrexate in free solution and dihydrofolate reductase.[9]

Examination of a Kendrew Model of the DHFR enzyme-methotrexate complex has already resulted in the "rational" design of novel inhibitors with significantly higher binding potencies.[10] In order to expand our capabilities in this area, a project is underway to develop macromolecular modeling capabilities. A protein crystallography group is being set up to isolate and determine the molecular structure of interesting target enzymes. We expect to obtain improved computer graphics equipment to aid the chemists in using these and other available enzyme structures for the design of effective new inhibitors. Theoretical calculation of drug-receptor binding energies will undoubtedly be crucial to the success of such studies.

Studies to date suggest
provide useful insights into
are most useful when cal
systems. With the appropria
convinced of the utility of t
applications are expected.

The authors thank Dr. J.
version of MMMS; Drs. J. E
who made many contributi
ties; and a large number of
ly, supported it, and suggest

1. PULLMAN, B. & A. PULLMA
2. GUND, P., J. D. ANDOSE, J.
3. HEHRE, W. J., W. A. LA
Gaussian 70, Program 2
A. WHITESIDE, R. KRISH
R. KAHN & J. A. POPLE.
4. SMITH, R. L., D. W. COCH
191-201; SMITH, R. L.,
Amiloride and Epitheli
A. Scriabine, Eds.: 21-3
5. BOCK, M., G. M. SMITH &
6. GUND, P., M. POE & K. HO
7. SUBRAMANIAN, M. & B. T.
8. FONTECILLA-CAMPS, J. C.,
L. KUSLIUK. 1979. J. Am.
BOLIN, D. J. FILMAN, F
PASTORE, L. T. PLANTE,
9. SCHLEGEL, H. B., M. POE &
10. POE, M., J. K. WU, J. R. FL
Submitted.

A.-M. SAPSE. (*John J*
comment on the difference
geometry and rotational ba
set to the optimization of t

Conclusion

Studies to date suggest that theoretical calculations on drug molecules can provide useful insights into their biological activity. Semiempirical calculations are most useful when calibrated with higher quality calculations on model systems. With the appropriate programs now installed, and with Merck chemists convinced of the utility of these calculations, an increasing number of successful applications are expected.

Acknowledgments

The authors thank Dr. J. D. Andose, who designed and implemented the first version of MMMS; Drs. J. B. Rhodes, G. M. Smith, J. Wendoloski, and E. Fluder, who made many contributions to MMMS and extended its theoretical capabilities; and a large number of Merck chemists, who have used the system effectively, supported it, and suggested improvements.

References

1. Pullman, B. & A. Pullman. 1963. Quantum Biochemistry. Interscience. London.
2. Gund, P., J. D. Andose, J. B. Rhodes & G. M. Smith. 1980. Science. **208:** 1425–31.
3. Hehre, W. J., W. A. Lathan, R. Ditchfield, M. D. Newton & J. A. Pople. 1973. Gaussian 70, Program 236. Quantum Chemistry Program Exchange; Binkley, J. S., R. A. Whiteside, R. Krishnan, R. Seeger, D. J. Defrees, H. B. Schlegel, S. Topiol, L. R. Kahn & J. A. Pople. 1981. Gaussian 80, Program 406. Ibid.
4. Smith, R. L., D. W. Cochran, P. Gund & E. J. Cragoe, Jr. 1979. J. Am. Chem. Soc. **101:** 191–201; Smith, R. L., D. W. Cochran, E. J. Cragoe, Jr. & P. Gund. 1979. *In* Amiloride and Epithelial Sodium Transport. A. W. Cuthbert, G. M. Fanelli, Jr. and A. Scriabine, Eds.: 21–33. Urban and Schwartzenberg. Baltimore, Md.
5. Bock, M., G. M. Smith & H. B. Schlegel. 1981. J. Org. Chem. In press.
6. Gund, P., M. Poe & K. Hoogsteen. 1977. Mol. Pharmacol. **13:** 1111–5.
7. Subramanian, M. & B. T. Kaufman. 1978. Proc. Nat. Acad. Sci. USA **75:** 3201–5.
8. Fontecilla-Camps, J. C., C. E. Bugg, C. Temple, Jr., J. D. Rose, J. A. Montgomery & R. L. Kusliuk. 1979. J. Am. Chem. Soc. **101:** 6114–5; Matthews, D. A., R. A. Alden, J. T. Bolin, D. J. Filman, F. T. Freer, R. Hamlin, W. G. J. Hol, R. L. Kisliuk, E. J. Pastore, L. T. Plante, N. Xuong & J. Kraut. 1978. J. Biol. Chem. **253:** 6946–54.
9. Schlegel, H. B., M. Poe & K. Hoogsteen. 1981. Mol. Pharmacol. In press.
10. Poe, M., J. K. Wu, J. R. Florance, J. P. Springer & K. Hoogsteen. 1981. J. Biol. Chem. Submitted.

Discussion

A.-M. Sapse. (*John Jay College—CUNY, New York*): I would like to comment on the difference between guanidine and guanidinium ion in terms of geometry and rotational barriers. We applied Gaussian 70 with the 6-31G basis set to the optimization of the geometry of several substituted guanidines, amino

and methyl substituted, and compared them to the same guanidinium ions. There are some differences in bond lengths and angles but the major difference is in the rotational barrier. We find that, when the barrier includes breaking hydrogen bonds, it's really high in the substituted guanidinium molecule. We find barriers on the order of 19 to 20 kcal/mol. When the barrier does not involve breaking hydrogen bonds, we get values as low as 10 to 12 kcal/mol. In the guanidine, the nonprotonated species, we find quite high barriers for the rotation around the double bond when electronegative substituents, such as F, are on the bond. We even find energies as high as 40 kcal/mol with optimization.

P. GUND: Very nice. Is that published?

SAPSE: It's been submitted.

GUND: It is interesting in that regard that these hydrogen bonds do seem to make a difference, even though they're in a six-membered ring and those are very nonlinear hydrogen bonds. Dr. Hoogsteen at Merck thought that you can't invoke them as hydrogen bonds because they're too nonlinear, and yet radius calculations and, apparently, your calculations show that they are real hydrogen bonds.

R. CRAMER (*Smith Kline & French, Philadelphia, Pa.*): I've gotten the impression that quantum chemists are critical of calculations that other quantum chemists make. How is it that organic chemists at Merck do quantum chemistry? Does this work out well in practice?

GUND: Well, I didn't mention this in my presentation, but after you create a structure with COORD and examine it to make sure it's the right structure, you type CNDO and get back the output of this semiempirical calculation. A lot of chemists are using the program and many have talked to me, Dr. B. Schlegel, or Dr. G. Smith and we've tried to help them to be reasonably cautious about their conclusions. There have been occasional arguments on differences in charge density in the third decimal place; we try to discourage that. But there have been a number of cases where it's been successful. Somebody may misinterpret a result, but that's the chance you take in making these programs available.

W. LIPSCOMB (*Harvard University, Cambridge, Mass.*): The barriers involving nitrogen are somewhat sensitive to the basis set and often require more than 2s2p basis functions, in particular, polarization functions, to come out quantitatively right. I think that's fairly well established from simpler molecules, and it might apply to your molecules as well, at least to the quantitative results.

GUND: Dr. Schlegel counted basis sets and decided he could add *d*-orbitals for some of our model calculations. We look forward to seeing if that lowers the barrier (of **(4)**).

H. SCHERAGA (*Cornell University, Ithaca, N.Y.*): You said that the amiloride equilibrium geometry was confirmed by x-ray study. Did the x-ray structure show that the hydrogen bonding arrangement was, indeed, intramolecular, as it was in your calculation, or were there intermolecular hydrogen bonds?

GUND: The x-ray structure was for the protonated molecule, and there were

intramolecular hydrogen bo
real hydrogen bonds are the

SCHERAGA: So that confo
bonding network?

GUND: I haven't seen th
intermolecular hydrogen bo

J. SAVAGE (*University of*
developed at the Universit
contouring of electron densi
knobs and match the stick m
It runs on the MMSX built b

LIPSCOMB: I think there a
are similar to MMMS.

GUND: Yes, they're bec
commercially available. For
of MMSX that will be made

intramolecular hydrogen bonds. There's also evidence from high dilution IR that real hydrogen bonds are there.

SCHERAGA: So that conformation didn't lead to an intermolecular hydrogen bonding network?

GUND: I haven't seen the packing, so it could well be that there are also intermolecular hydrogen bonds.

J. SAVAGE (*University of Alberta, Edmonton*): There is a program called M3, developed at the University of Alberta, which is capable of doing real time contouring of electron densities and also places stick models inside. You can turn knobs and match the stick models at the density maps. The program is available. It runs on the MMSX built by the St. Louis group.

LIPSCOMB: I think there are a number of molecular modeling systems now that are similar to MMMS.

GUND: Yes, they're becoming fairly common; in fact, they are becoming commercially available. For instance, I believe that Dr. G. Marshall has a version of MMSX that will be made commercially available.