Woo, H.-G.; Hong, L.-Y.; Kim, S.-Y.; Park, S.-H.; Song, S.-J.; Ham, H.-S. *Bull. Korean Chem. Soc.* **1995**, *16(8)*, 774. (c) Woo, H.-G.; Hong, L.-Y.; Park, S.-H.; Ham, H.-S. Unpublished results.
10. Walsh, R. *Acc. Chem. Res.* **1981**, *14*, 246.
11. Huheey, J. E.; Keiter, E. A.; Keiter, R. L. *Inorganic Chemistry: Principles of Structure and Reactivity*, 4th ed.; Harper Collins College Publishers: New York, 1993; p A25.
12. Brandrup, J.; Immergut, E. H. *Polymer Handbook*, 3rd ed.; Wiley: New York, 1989.
13. Ballestri, M.; Chatgilialoglu, C.; Guerra, M.; Guerrini, A.; Lucarini, M.; Seconi, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 421 and references theirin.
14. Hsiao, Y.-L.; Waymouth, R. M. *J. Am. Chem. Soc.* **1994**, *116*, 9779.

The Binding Energy of HIV-1 Protease Inhibitor

Jaejin Ka, Sang-Hyun Park, and Hojing Kim

Department of Chemistry, Seoul National University, Seoul 151-742, Korea
Received September 4, 1995

The potential energies of HIV-1 protease, inhibitor, and their complex have been calculated by molecular mechanics and the "binding energy", defined as the difference between the potential energy of complex and the sum of potential energies of HIV-1 protease and its inhibitor, has been compared to the free energy in inhibition reaction. The trend in these binding energies seems to agree with that in free energies.

Introduction

The etiological agent of AIDS has been identified as human immunodeficiency virus (HIV), and two genetically distinct subtypes, HIV-1 and HIV-2, have been characterized.[1,2] There are many kinds of possibilities in blocking the normal multiplication of HIV-1. One of those methods is to block the normal action of HIV-1 protease so that HIV-1 protease can not effectively make regulatory proteins, structural proteins and maturation proteins from their precursor, polyprotein. HIV-1 protease is consisted of two monomer units each of which has completely the same amino acid sequence of 99 residues as shown in Figure 1 and has C_2 symmetric axis within it as shown in Figure 2.

Several types of HIV-1 protease inhibitors are suggested and investigated by many authors. One of the classical strategies for designing enzyme inhibitors relies on incorporating a transition-state mimic into substrate analogues. Refinements of this strategies, in which nonhydrolyzable dipeptide isosteres were substituted for the scissile amide bond in an appropriate sequence context, proved highly successful for producing potent renin inhibitor.[3] Figure 3 shows inhibitors which is used in our calculation.

Many researchers have tried to develop new inhibitors with better activities. The proper descriptor of activity may relieve an effort of the development but the current methods do not give satisfactory results for the development of new inhibitor, especially which has substantially different structure from the existing inhibitors. So we tried to define the physical quantity which is computable simply, applicable to various inhibitors and predictable qualitative trends of activity and investigated the quantity with available data set.

PRO GLN ILE THR LEU TRP GLN ARG PRO LEU VAL THR ILE
LYS ILE GLY GLY GLN LEU LYS GLU ALA LEU LEU **ASP** THR
GLY ALA ASP ASP THR VAL LEU GLU GLU MET ASN LEU PRO
GLY ARG TRP LYS PRO LYS MET ILE GLY GLY ILE GLY GLY
PHE ILE LYS VAL ARG GLN TYR ASP GLN ILE LEU ILE GLU
ILE CYS GLY HIS LYS ALA ILE GLY THR VAL LEU VAL GLY
PRO THR PRO VAL ASN ILE ILE GLY ARG ASN LEU LEU THRz
GLN ILE GLY CYS THR LEU ASN PHE

Figure 1. Amino acid sequence of HIV-1 protease which was derived from bacterial expression. The six underlined residues are different amino acids from corresponding synthetic protease. The 25th residue, ASP in boldface is very important in the catalytic hydrolyzation of polyprotein.

Methods

Inhibition constant, K_i, for the equilibrium between the enzyme and inhibitor, is defined as follows:

$$EI \Leftrightarrow E + I \quad (1)$$

$$K_i = \frac{[E][I]}{[EI]} \quad (2)$$

The inhibitor which has smaller K_i bind more tightly to its target. Since K_i is obtained through biological assays, *in vitro* experiments, the synthesis of inhibitor is prerequisite to its test. Therefore, even crude, the prediction of the potency of artificial inhibitor, is preferable in alleviating efforts of syntheses. In present work, we tried to find the physical quantities which describe the inhibition constant properly and which can be calculated by theoretical computation even before synthesis.

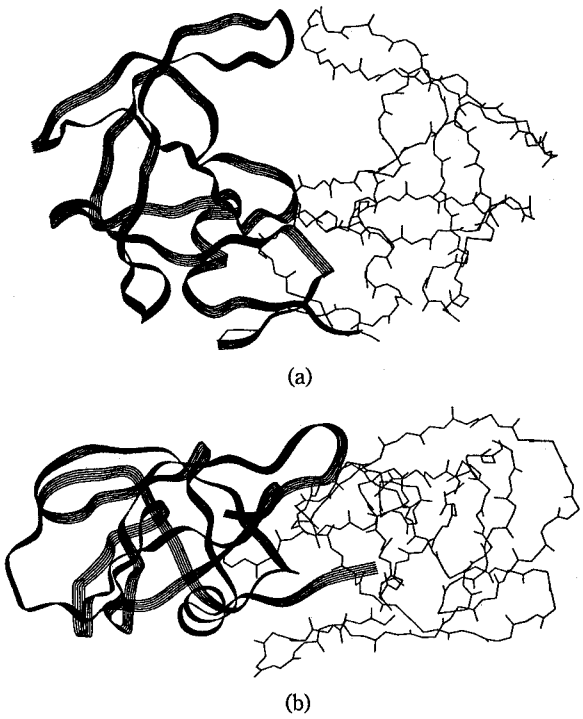

Figure 2. (a) The front view and (b) the top view of HIV-1 protease. One monomer unit of HIV-1 protease dimer is drawn as a ribbon shape and the other monomer unit is drawn as only skeleton shape.

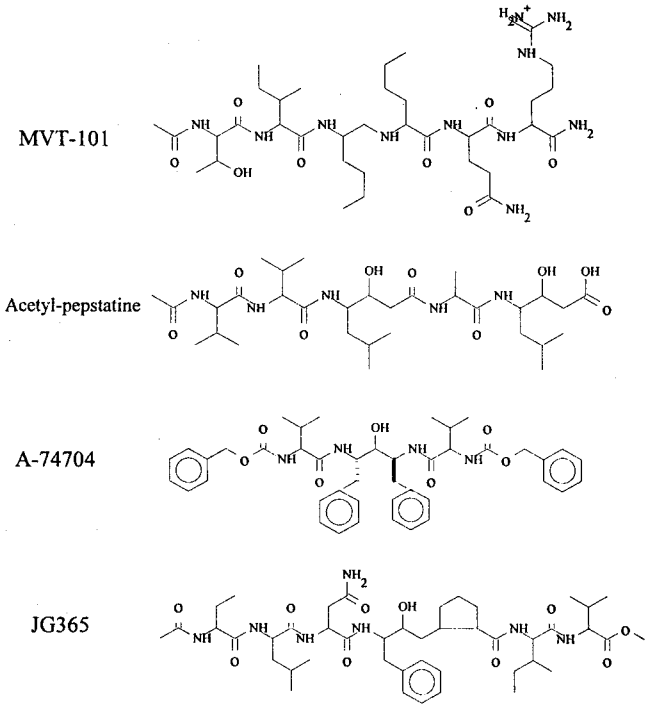

Figure 3. Inhibitors which were used for calculation.

There exists relationship between equilibrium constant K and Gibbs free energy change, $\Delta G°$, as follow.

$$-RT\ln K = \Delta G° \quad (3)$$

And several computer softwares-*Charmm,*[4,5] *Gromos,*[6,7,8] *Amber,*[9,10] *and Discover*[11] *etc.*-are available for the free energy computation. They generate a number of conformations of molecule and obtain an ensemble average.

But there is practical problem in the calculation of free energy for so large system such as enzymes because the computing time will be prohibitively long. The complex of HIV-1 protease and MVT-101 is consisted of 3128+122= 3250 atoms, and its 110ps MD with hydration sphere of 3741 water molecules will require CPU time of 23days 18hrs on IBM work station RS6000 530H. Furthermore the CPU time of free energy calculation is generally longer than that of MD.[12] If we try to get the free energy of eqs (1) and (3) in this way, the CPU time of about 50 days will be needed.

There are some approximate methods in order to overcome this difficulty. Herman *et al.*[6] have used free energy simulations(slow-change method) to estimate quantitatively the ratio of the binding constants of (S) and (R) isomers of a HIV-1 protease inhibitor, JG365. Kollman *et al.*[10] have presented the application of free energy perturbation theory/molecular dynamics to predict the consequence of replacing each of the seven peptide bonds in JG365 : ACE(acetyl)-Ser-Leu-Asn-HEA(hydroxyethylamine analog of Phe-Pro)-Ile-Val-NME(N-methyl) by ethylene of fluoroethylene isosteres. Both of two authors have used the following cyclic scheme.

$$\begin{array}{ccc} [1]\ \text{HIV-PR}+\text{JG365} & \Leftrightarrow & \text{HIV-PR}:\text{JG365}\ [2] \\ \Updownarrow & & \Updownarrow \\ [3]\ \text{HIV-PR}+\text{JG365}'' & \Leftrightarrow & \text{HIV-PR}:\text{JG365}''\ [4] \end{array}$$

The sum of the free energy changes for the entire cycle must be zero, and this condition can be rewritten as

$$\Delta\Delta G° = \Delta G°_{12} - \Delta G°_{34} = \Delta G°_{13} - \Delta G°_{24} \quad (4)$$

where the first difference is related to the experimentally measurable affinities (e.g. in terms of inhibition constants) and the second difference is computed. The first simulation reports the free energy difference due to conversion of the JG365 into JG365″ for free inhibitor peptide ($\Delta G°_{13}$) and the second one gives an estimate of the free energy difference between JG365 and JG365″ when the inhibitor peptide is bound to the HIV protease ($\Delta G°_{24}$). To make computation easier, Herman have used two approximations. (i) MD of only part of free inhibitor may give the free energy difference of whole inhibitor ($\Delta G°_{13}$). (ii) For the protease-inhibitor complex, the moving atoms were limited to those found within 10 Å spheres around enzyme active sites. These approximations make free energy calculation rather tractable. But the chemical perturbation (small structural difference) between JG365 and JG365″ is very small in comparison with whole structure. This method is still unusable for the comparison of inhibitors with entirely different structures (e.g. MVT101 *vs.* JG365). More approximation has been introduced in free energy perturbation method by Gunsteren *et al.*[8] They assumed that the initial slope in a graph of the free energy versus perturbation parameter points towards a final value of the free energy which lies near the exact one. The proposed approximation requires a factor of 1000 less computing effort than a full energy perturbation calculation. Even

with this approximation, however, the comparison of potencies of MVT101 and JG365 is still prohibitive.

In this work, the potential energies of geometry-optimized E, I and EI of eq. (1) were calculated by using mainly MM. It does not need seriously long CPU time. The definition of binding energy for this system is as follows.

$$\Delta E_P = E_P(\mathrm{E}) + E_P(\mathrm{I}) - E_P(\mathrm{EI}) \equiv E_{binding} \quad (5)$$

E_P : potential energy

Eq. (6)-(8) show the relationship between binding energy and free energy. As $E_{binding}$ is more positive, inhibitor (I) binds more tightly to protease (E).

$$-\mathrm{RT}\ln K = \Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = \Delta U^\circ + \Delta(\mathrm{PV})^\circ - T\Delta S^\circ \quad (6)$$

Since $\Delta(\mathrm{PV})^\circ = 0$ for this system, the equation (6) can be rewritten as follows:

$$-\mathrm{RT}\ln K = \Delta U^\circ - T\Delta S^\circ \quad (7)$$

$$= \Delta E_p^\circ + \Delta E_k^\circ - T\Delta S^\circ \quad (8)$$

E_K : kinetic energy

Computational Details

Addition of all hydrogen atoms to the raw PDB structures, analyses of the complexes, enzymes and inhibitors and manipulation of the structures were carried out using the molecular modeling package *InsightII version 2.1.0*.[13] Energy minimization and some dynamics were performed using *Discover version 2.8*[12,14] using CVFF force field[15] on IBM 530H and M20 workstations.

Binding energy of MVT-101 to synthetic HIV-protease

Complex+4 Å hydration shell. The crystal structure of complex of MVT-101 and HIV-1 protease were obtained from Protein Data Bank (file, 4HVP). The coordinates of hydrogens were generated by the *Builder* module of *Insight II program*. The condition of pH=5.4 was used, because *in vitro* experiment for activity evaluation had been carried out at pH=5.4.[16,17] The pH has an effect on the protonation or deprotonation of amino acid residues which have their own native pK values.

Next step was to make hydration shell enclosing the complex of HIV-1 protease and MVT-101. This was done in order to mimic the real environment of *in vitro* experiment. This task can be accomplished in two ways , one is to make hydration shell of constant thickness from the surface of the complex, the other is to make hydration sphere centered at designated point and of assigned radius. The former is more adequate for such big system (total 3250 atoms) as enzyme complex because the system would require longer computation time in the energy minimization for the latter method. (In the latter method, the system would require 3741 water molecules in enclosing the whole complex, whereas in the former only 793 water molecules.) In this work, the hydration shell of 4 Å thickness was used.

And then energy minimization was performed with the CVFF (Consistent Valence Forcefield) of *Discover*. A cut-off of 16 Å was used for non-bonded interactions. (Even 8 Å is judged to be sufficient for this force field.[7,17]) And the option of IGRPCK=0 was used.[18] Steepest descents method was applied for the first 500 cycles of each minimization and then minimization was continued till energy gradient became less than 0.01 Kcal/molÅ with conjugate gradients method.[16] During whole minimization, a distance dependent dielectric function, $D(r)=4r$ was used. This 'high dielectric' model with $D(r)=4r$ had been favored by Whitlow and Teeter in their calculations on crambin[19] and by Novotny *et al.* in their work on antibody-antigen complexes.[20]

PRO GLN ILE THR LEU TRP GLN ARG PRO LEU VAL THR ILE
ARG ILE GLY GLY GLN LEU LYS GLU ALA LEU LEU ASP THR
GLY ALA ASP ASP THR VAL LEU GLU GLU MET ASN LEU PRO
GLY LYS TRP LYS PRO LYS MET ILE GLY GLY ILE GLY GLY
PHE ILE LYS VAL ARG GLN TYR ASP GLN ILE PRO VAL GLU
ILE ABA GLY HIS LYS ALA ILE GLY THR VAL LEU VAL GLY
PRO THR PRO VAL ASN ILE ILE GLY ARG ASN LEU LEU THR
GLN ILE GLY ABA THR LEU ASN PHE

Figure 4. Amino acid sequence of synthetic [Aba[67,95]] HIV-1 protease. The six residues that differ from natural HIV-1 protease, Figure 1, are underlined. Aba(L-α-Amino-n-butyric acid) was used in place of the two Cys residues, effectively substituting a methyl for the thiol group in the Cys side chain, to reduce synthetic difficulties associated with Cys deprotection and to ease product handling.[21]

Table 1. The number of water molecules within hydration shell. The thickness of shell is 4 Å for complex, 3 Å for protease, not constant for inhibitor

PDB	Complex	Protease	Inhibitor
hvp4	793	557	216
hvp5	811	538	273
hvp7	798	577	221
hvp9	833	538	295

Enzyme+3 Å hydration shell. HIV-1 protease which had been used in biological assays with MVT-101 was prepared by total chemical synthesis.[21,22] The target amino acid sequence is shown in Figure 4.

This synthetic enzyme has rather different amino acid sequence from that of natural expressed HIV-1 protease but still have the specific proteolytic activity characteristic of the HIV-1 protease[21] and have a turnover number comparable to that reported by Darke *et al.* for the enzyme derived from bacterial expression. On the other hand, synthetic protease has an advantage over natural expressed protease in that it can be prepared relatively more quickly for extensive biological assays of many inhibitors.

Crystallographic data for above synthetic protease were obtained from PDB file (3HVP). In this case, 3 Å (not 4 Å) hydration shell of 557 water molecules was adopted so that the extra water molecules (793-557=216) may be used in enclosing MVT-101. Table 1 exhibits the number of water molecules which are used in enclosing complex, protease and inhibitor in each case. Dielectric function, pH, cut-off, and minimization procedures are same as the previous calculation.

Inhibitor. In the case of complex and enzyme, it was assumed that the optimized geometry of molecule in solution

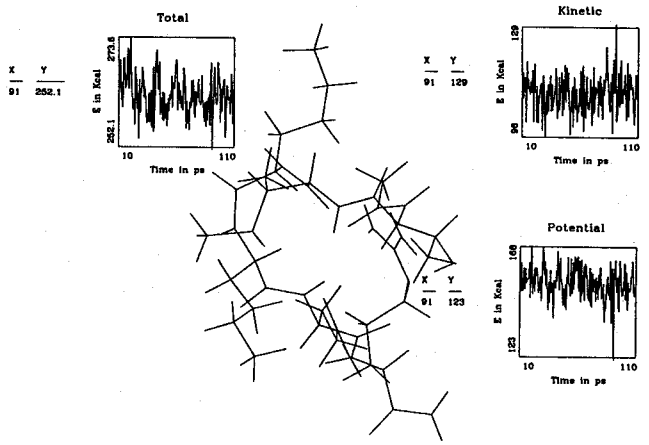

Figure 5. The minimum potential energy conformation of MVT-101. Three graphs demonstrate the variations of total, kinetic and potential energies for 110ps dynamics simulation at 303 K. From the graph of potential energy, minimum potential is achieved at 91ps. And the conformation at the very instant is shown above. X denotes time and Y energy.

would not be seriously different from the X-ray structure. The assumption may be rather reasonable since the energy minimization with hydration shell has been performed to mimic the environment in solution. But for inhibitor, we could not find any coordinate of MVT-101 in cambridge crystallographic data. The optimized geometry of inhibitor MVT-101 was obtained by molecular dynamics in acceptable computation time after all atoms except MVT-101 inhibitor were removed from the HIV protease-MVT-101 complex. During the molecular dynamics, the water environment in solution was mimicked implicitly by using dielectric function, $D(r)=4r$, without including any water molecule explicitly. If the MD had been performed with explicit hydration layer of water molecules, the unwelcome phenomenon of evaporation would have been inevitable.

At first, MTV-101 was minimized until the maximum energy gradient became less than 0.01 kcal/mol$\AA$ and then the temperature was raised to 303 K[5] in 10ps for equilibration and 100ps MD was followed at that temperature. The data collection was done at every 500fs. Figure 5 shows the variations of potential, kinetic and total energies for 110ps MD at 303 K.

From the graph of potential energy, the conformation at 91ps corresponds to the minimum of potential energy. The conformation of minimum potential energy was taken and hydrated with 216 water molecules and then its potential energy was minimized according to the same procedure established as before.

Binding energy of Acetyl-pepstatin to natural HIV-1 protease

Richards *et al.* had investigated the inhibition constants of acetyl-pepstatin against HIV-1 protease at several pH values.[23] So also in our work, the binding energies of acetyl-pepstatin to HIV-1 protease were calculated at several pH values; pH=4.7, 5.0 and 7.0. The complex of expressed HIV-1 protease and acetyl-pepstatin was hydrated as in Table 1 and then three sets of different pH's were prepared and each of those three was energy minimized. The HIV-1 protease which had been used in evaluating inhibition constants of acetyl-pepstatin is naturally expressed one and the X-ray structure of it is contained in PDB with the file name of 3PHV. The same procedure as the case of MVT-101 was followed in the MD and MM of acetyl-pepstatin.

Table 2. The comparison between experimental inhibition constants, K_i and calculated binding energies, $E_{binding}$ of several inhibitors

PDB	Enzyme	Inhibitor	pH	K_i	RT1nK_i	$E_{binding}$
hvp4	syn*	MVT-101	5.4	760	4.00	164.58
hvp5	nat#	acetyl-p	4.7	20	1.80	268.94
			5.0	35	2.14	268.94
			7.0	>1000	>4.16	267.92
hvp7	syn	JG365	5.5	0.24	−0.86	173.95
hvp9	nat	A74704	4.7	4.5	0.91	273.02

RT=0.602 kcal/mol (at 30 ℃). The unit of K_i is nM and energies in kcal/mol. *syn: synthetic protease. #nat: natural protease.

Table 3. The effect of HIV-1 protease on a naturally occurring inhibitor, acetyl-pepstatin[23]

pH	Salt	K_i (nM)
4.7	20 mM acetic acid, 20 mM Mes 40 mM Tris, ionic strength μ=0.82 M	20
5.0	0.1 M sodium acetate, 4 mM EDTA, 5 mM mercaptoethanol, 1000 mM NaCl μ=1.06 M	35
7.0	0.1 M Mes, 4 mM EDTA 5 mM mercaptoethanol 1000 mM NaCl	>1000

temp=30 ℃

Binding energies of JG365 and A74704

Also for JG365 and A74704, all procedure were similar to the previous cases of MVT-101 and acetyl-pepstatin, and the number of water molecules in the hydration shell is contained in Table 1.

Results and Discussion

The results of calculation are given in Table 2. The extent of binding of HIV-1 protease and acetyl-pepstatin depends on pH in solution as indicated in Table 3.

The pH dependence of inhibition constant and binding energy. This pH dependence was simulated in the *Builder* module of *InsightII.* The extent of protonation of enzyme, inhibitor and their complex at different pH is explained in Table 2. The variation of pH value causes the variation in total number of atoms of system and consequently leads to the variation of energy after minimization. Table 2 shows that the pH effect on inhibitor activity is exhibited also in our calculation as in experiment. For example, both experiment and calculation say that Acetyl-pepstatin has larger activity at pH 4.7 (or 5.0) than at pH 7.0. However, this is true only qualitatively not quantitatively because the binding energy of acetyl-pepstatin at pH=4.7 is larger than that

at pH=7.0 by 1.02 kcal/mol while the difference at experiment ($RT\ln K_i$) is 2.36 kcal/mol. On the other hand, the difference of $RT\ln K_i$ at pH=4.7 and pH=5.0 is 0.34 kcal/mol but is zero kcal/mol in our work.[24] This may be contributed from the different salt environments of real experiments.

According to the result in Table 2, acetyl-pepstatin is found to have same binding energy at pH=4.7 and 5.0. In the minimization of complex with acetyl-pepstatin, the calculated energies are exactly same, 1512.00 kcal/mol at both pH. This is because any residue in complex does not behave in the different ways at both pH. There is no amino acid which has pK value in the range of 4.7-5.0.

· But this result is not compatible with the experiment in Table 3. Why not? And what will happen if the free energies is calculated at both pH? Even if the free energies of system at both pH are calculated, they should be exactly same values because system does not have any difference at both pH; same atoms, same coordinates. Therefore they should have same inhibition constant from eq. (3). In experiment, there exist different numbers and kinds of salts around system at pH 4.7 and 5.0, although system itself has exactly same number of proton(or hydrogen) at both pH. The different salt environment of system should be mimicked in calculation for more exact result but only protonation of system at both pH is considered mainly because of computational difficulty at the present stage.

The inhibitor dependence of inhibition constant and binding energy. The same pH and salt environment are prerequisite to get reliable analysis of $-RT\ln K_i$ *vs.* $E_{binding}$. In Table 1, the inhibition constants of MVT-101 and JG365 were evaluated at almost same pH; 5.4, 5.5. And acetyl-pepstatin and A74704 were assayed at exactly same pH; 4.7. As mentioned before, synthetic protease was used in biological assays of MVT-101 and JG365, while bacterial expressed protease was used for acetyl-pepstatin and A74704. Therefore the comparison within only same class is desirable. A74704(K_i=4.5 nM) is known as superior inhibitor to acetyl-pepstatin (K_i=20 nM) from biological assays and this was reflected in our work. The $E_{binding}$ of A74704 is 273.02 kcal/mol and for acetyl-pepstatin 268.94 kcal/mol. Also JG365 (K_i=0.24 nM) is stronger inhibitor than MVT-101 (K_i=760 nM) and our results (173.95, 164.58 kcal/mol each) agree with the experiment qualitatively.

From Eqs. (4)-(7) and $K=K_i/Ki'$,[24] the relation (9) is satisfied for each inhibitor.

The effectiveness of binding energy. For an inhibitor,

$$-RT\ln K_i = -RT\ln K_i' + \Delta E_P^\circ + E_K^\circ - T\Delta S^\circ \quad (9)$$

Between inhibitors,

$$\delta(-RT\ln K_i) = \delta\Delta E_P^\circ + \delta\Delta E_K^\circ - T\delta\Delta S^\circ. \quad (10)$$

where $\delta\Delta E_P^\circ = \delta E_{binding}$

The $\delta(-RT\ln K_i)$ between MVT-101 and JG365 is 4.86 kcal/mol and $E_{binding}$ is 9.37 kcal/mol. So the contribution of kinetic energy and entropic term to the free energy is inferred to be −4.51 kcal/mol. In the case of acetyl-pepstatin and A74704, $\delta(-RT\ln K_i)$ is 0.89 kcal/mol and $\delta E_{binding}$ is 4.08 kcal/mol and $\delta\Delta E_K^\circ - T\delta\Delta S^\circ = -3.19$ kcal/mol. From this result, it is inferred that the contribution of kinetic energy and entropic term to free energy is not generally similar but may not be major effect. The magnitude of $\delta(-RT\ln K_i)$ seems to be proportional to $\delta E_{binding}$ rather than to $\delta\Delta E_K^\circ - T\delta\Delta S^\circ$. But it does not mean that our result is applicable to quantitative prediction of inhibition constant. Instead it can be said that qualitative prediction in trends of inhibition strength will be probable.

Conclusion

In this work, we defined and calculated "binding energy" and examined whether there is a parallelism between calculated "binding energy" and experimental inhibition constant. From Results and Discussion, the defined binding energy seems to reflect activity (K_i) trend of inhibitors qualitatively. This method has some advantages; (i) it does not require long computation time. (ii) it is applicable to qualitative comparison of activity of inhibitors which have even dissimilar structures.

In this work, only four inhibitors were calculated because those are all data set provided by PDB. After a further accumulation of PDB data, the more extensive study on parallelism between $-RT\ln K_i$ and $E_{binding}$ should be performed in order to confirm whether "binding energy" is a good descriptor for regression analysis of this system. If the entropy effects between different complexes are quite different, the regression between $-RT\ln K_i$ and $E_{binding}$ will have small correlation coefficient, r. In the case of large correlation coefficient, however, the inhibition constants of new inhibitor candidates will be predictable approximately and give efficient informations for inhibitor synthesis. Of course, it must be another important subject to presume the structure of new inhibitor candidate-HIV protease complex appropriately.

Acknowledgment. This work has been supported by the Korea Science and Engineering Foundation, the Ministry of Education, the Korea Research Foundation, and the S. N. U. Daewoo Research Fund.

References

1. Clavel, F.; Guyader, M.; Guetard, D.; Salle, M.; Montagnier, L.; Alizon, M. *Nature* **1986**, *324*, 691.
2. Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M. *Nature (London)* **1986**, *326*, 662.
3. W. Greenlee. *J. Med. Res. Rev.* **1990**, *10*, 173.
4. Won, Y. D.; Yu, S. W. *Chemworld* **1993**, vol. 33 no. 2 pp 99-104.
5. Brooks, B. R.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
6. Alexander Tropsha; Jan Hermans, *Protein Engineering* **1992**, vol. 5, no. 1, pp 29-33.
7. van Gunsteren, W. F.; Berendsen, H. J. C., Groningen Molecular Simulation(GROMOS). Biomos. Groningen. 1987.
8. Gerber, P. R.; Mark, A. E.; van Gunsteren, W. F. *Journal of Computer-Aided Molecular Design* **1993**, *7*, pp 305-323.
9. Weiner, P. K.; Kollman, P. A. *J. Comp. Chem.* **1981**, *2*, 287.
10. Piotr Cieplak; Kollman, P. A. *Journal of Computer-Aided*

Molecular Design **1993**, *7*, pp 291-304.
11. Biosym Technologies Inc., *Discover*, User guide, ver 2.8 March **1992**, pp 2-115.
12. Biosym Technologies Inc., *Discover*, User guide, ver 2.8 March **1992**, pp 2-97.
13. Biosym Technologies Inc., *InsightII*, User guide, ver 2.1.0 March 1992.
14. Biosym Technologies Inc., *Discover*, User guide, ver 2.8 March **1992**, pp 2-4.
15. Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Woliff, J.; Genest, M.; Hagler, A. T. *Proteins* **1988**, *4*, pp 31-47.
16. Clare E. Sansom; Jin Wu; Irene T. Weber *Protein Engineering* **1992**, vol. 5, no. 7, pp 659-667.
17. Maria Miller *et al*, *Nature* **1989**, *337*, pp 576-579.
18. Some residues are so big that they are greater than half the cut-off distance. In this case, *Discover* stops that process. So larger cut-off distance should be adopted. However, computing time will increase in third power as cutoff distance increases linearly. Besides, even cutoff 20 Å made a trouble during the computation, though it had worked well at the start of process. As size of amino acid rather varies during minimization, cutoff value should be sufficient during whole process. The option of IGRPCK=0 makes it possible to continue the process with ignoring the instant expand of residue, if the process did not have any trouble at the start of MM.
19. Whitlow, M.; Teeter, M. M. *J. Amer. Chem. Soc.* **1986**, *108*, pp 7163-7172.
20. Novotny, J., Bruccolen, R .E.; Saul, F. A. *Biochemistry* **1989**, *28*, pp 4735-4749.
21. Alexander Wlodawer *et al*, *Science* **1989**, *245*, pp 616-621.
22. Schneider, J.; Kent, S. B. H. *Cell* **1988**, *54*, 363.
23. Anthony D. Richards *et al*, *FEBS letters* **1989**, *247*, pp 113-117.
24. Of course this needs to be explained more. The equilibrium constant, K in eq. (3) is unitless but inhibition constant K_i has unit of nM. Therefore $RT\ln(K_i/K_i')$, K_i' some calibration constant, has closer relation to $E_{binding}$. So the comparison between $\delta(RT\ln K_i)$ and $E_{binding}$ is more desirable to that between $RT\ln K_i$ and $E_{binding}$.

Synthesis and Configuration Analysis of Diastereomers of 5′-O-(2′-Deoxycytidyl)-3′-O-Thymidyl Phosphorothioate

Byung Jo Moon*, Hyun Ju Jeong, Sang Kook Kim, and Nam Hee Kim

Department of Biochemistry, College of Natural Sciences, Kyungpook National University, Taegu 702-701, Korea
Received September 7, 1995

A procedure is described for the synthesis of the title compound *via* phosphotriester intermediates. The preparation of R_p and S_p diastereomeric dinucleotide of d[Cp(S)T] was performed by the condensation of the protected deoxycytidine, the protected thymidine, 2,5-dichlorophenylphosphorodichloridothioate and 1-hydroxybenzotriazole in THF. Their designation of configuration at phosphorus as R_p and S_p follows from anaylsis of ^{31}P NMR spectroscopy and reverse-phase HPLC and the stereospecificity in the hydrolysis catalyzed by Nuclease S1 and snake venom phosphodiesterase. Diastereomerically pure R_p and S_p d[Cp(S)T] were utilized to synthesize oligonucleotides containing the XhoI recognition sequence with a phosphorothioate group at the cleavage site.

Introduction

Diastereomeric phosphorothioate analogues of nucleotide are important tools for elucidation of the stereochemistry of action of different classes of enzymes. For instance, the stereochemical course of action of RNase A.[1,2] To obtain a more complete insight into enzyme-substrate interactions of exo- and endonucleases in general, one needs the two diastereomeric phosphorothioate analogues of an appropriate dinucleoside monophosphate in their optically pure form.

To study the stereochemical course of reaction catalyzed by XhoI restriction endonuclease, we need to synthesize a oligodeoxynucleotide which contains the XhoI recognition sequence with a phosphorothioate group at the cleavage site. Deoxyoligonucleotides which contain phosphorothioate linkage at cleavage site of restriction endonucleases can be used to elucidate the stereochemical course of reaction catalyzed by restriction endonucleases. For instance, the sterochemical courses of action of EcoRI[3,4] and EocRV[5] have been established using oligonucleotide containing the appropriate recognition sequence with phosphorothioate internucleotidic linkage of known absolute configuration. Thus we have synthesized opically pure diastereomers of 5′-O-(2′-deoxycytidyl)-3′-O-thymidyl phosphorothioate (d[Cp(s)T]), which can be used as a dimeric building block to make oligonucleotides containing the appropriate recognition sequence of XhoI restriction endonuclease with phosphorothiate internucleotidic linkage of cleavage site.

*To whom correspondence should be addressed.