Free Energy Calculations on Protein Stability: Thr-157 → Val-157 Mutation of T4 Lysozyme

Liem X. Dang, Kenneth M. Merz, Jr., and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143. *Received May 15, 1989*

Abstract: We present free energy perturbation calculations on the X-ray structure of native T4 lysozyme, the Thr-157 → Val-157 mutant of T4 lysozyme, and on a tetrapeptide model that represents the denatured state of the native and mutant enzymes. This is the first free energy perturbation calculation on protein stability in a case where a precise comparison with experiment can be made. The relative calculated free energy difference between native and denatured models for native and the Thr → Val mutant was 1.9 ± 1.1 kcal/mol. This value is in good agreement with the experimental value of 1.6 kcal/mol, supporting our approach. Contrary to expectation, the van der Waals rather than the electrostatic component of the computed free energy is the dominant determinant of this differential stability. The structures that emerge from the simulations after mutation differ in important ways from the X-ray structures. This emphasizes the difficulties in reproducing X-ray structure with incomplete protein representation as well as limited simulation times (40 ps here). Nonetheless, the hydrogen-bonding patterns in the simulated structures are physically reasonable.

T4 lysozyme has become the paradigm for the study of the dependence of protein stability on protein sequence and three-dimensional structure.[1] Both the X-ray structures and the thermodynamics of denaturation of the protein and many mutants are available.[2] Thus the question can be posed: Can theoretical molecular dynamics/free energy perturbation methods simulate the relative free energies of protein stability and the differences in X-ray structures of native and mutant enzymes? These free energy methods have been shown to be very useful in studies of solvation free energies,[3,4] effect on drug structure of protein–ligand binding,[5] and effect of site-specific mutations on enzyme ligand binding and catalysis.[6,7]

Two applications of these methods to protein stability have been presented,[3,8] but in neither case was comparison made to experimental free energies. One of the inherent difficulties in making such a comparison is the fact that no structural model is available for the denatured enzyme. Thus, a free energy calculation on protein stability must make an unproven assumption about the denatured protein structure. The nature of the calculation can be summarized by the following model:

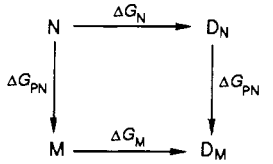

Table I

MD	$\Delta G_{forward}$	$\Delta G_{backward}$	ΔG_{av}	$\Delta\Delta G$
native (ΔG_{PN})			Thr → Val	
Thr → Val	5.97	−7.08	6.53 ± 0.53	
Val → Thr	−5.21	8.22	6.73 ± 1.50	
denatured (ΔG_{PD})				
Thr → Val	4.56	−4.84	4.70 ± 0.14	
$\Delta G_{PN} - \Delta G_{PD}$(calc)				1.91 ± 1.1
$\Delta G_N - \Delta G_M$(exp)				1.6

The relative free energy of denaturation of native (N) and mutant (M) protein in $\Delta\Delta G = \Delta G_N - \Delta G_M$. Theoretical studies can be carried out in which one mutates the native protein into the mutant (ΔG_{PN}) and the denatured native protein D_N, into the denatured mutant, D_M (ΔG_{PD}). In the case of the protein simulations, one can use the X-ray structure to carry out the calculation. For the denatured form, no structure is available. We decided to use a tripeptide model of sequence ANB, which is a part of the native sequence, and mutate this into sequence AMB in solution. In the specific application reported here, we used residues 156–158 (Gly-Thr-Trp) of T4 lysozyme with blocked N and C terminal groups, and mutated Thr → Val in solution to determine an approximate value of ΔG_{PD}. Then, the value of ΔG_{PN} was determined in two different simulations that mutated Thr-157 → Val-157 and Val-157 → Thr-157 in the proteins. In each case we started with the X-ray structure of the respective protein.

Methods

The native protein simulations involved the placement of a cap of 266 water molecules with a radius of 15 Å from the Thr-157 CB atom. The Val-157 simulation was set up similarly. The systems were minimized and then equilibrated for 6 ps with a nonbonded cutoff of 8 Å, at 300 K, with a nonbonded list update every 50 time steps. During these simulations, only those residues and water molecules within 15 Å of atom CB of residue 157 were allowed to move with the use of the "Belly" option within AMBER[10] 3.0 of Singh et al. The denatured protein simulations were modeled with the use of residues 156–158 of T4 lysozyme with a CH_3 group prior to residue 156 and a $NHCH_3$ group following residue 158. This tetrapeptide was place in a box consisting of 736 water molecules. The system was minimized and then equilibrated for 6 ps with periodic boundary conditions at a constant pressure of 1 atm and a constant temperature of 300 K. Finally, the free energy simulations were carried on these systems using the slow-growth procedure and the free energies for both the forward ($\lambda = 1 \rightarrow 0$) and reverse ($\lambda = 0 \rightarrow 1$) directions were obtained. In all calculations, a total of 40 ps was used in each direction with a time step of 1 fs. SHAKE[11] was used to con-

(1) Alber, T.; Matthews, B. W. *Methods Enzymol.* **1987**, *154*, 511.
(2) Alber, T.; Dao-pin, S.; Wilson, K.; Wozniak, J. A.; Cook, S. P.; Matthews, B. W. *Nature* **1987**, *330*, 41. Grutter, M. G.; Gray, T. M.; Weaver, L. H.; Alber, T.; Wilson, K.; Matthews, B. W. *J. Mol. Biol.* **1987**, *197*, 315.
(3) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564.
(4) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607.
(5) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574.
(6) Hwang, J.-K.; Warshel, A. *Biochemistry* **1987**, *26*, 2669.
(7) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551.
(8) Wong, C. F.; McCammon, J. A. In *Structure, Dynamics and Function of Biomolecules*; Ehrenberg, A., Rigler, R., Eds.; Springer-Verlag: Berlin, 1986.
(9) These calculations were carried out for 10 ps versus the 40 ps used for the Thr → Val simulations. In each calculation we started with an equilibrated structure, which was obtained as described in Table I. Free energies for only the forward direction ($\lambda = 1 \rightarrow 0$) were determined for the native protein and the tripeptide model in solution using the slow-growth method. The calculations on the electrostatic interaction of the NH group in the Thr-157 and Val-157 structures with the rest of protein were handled similarly.

(10) Singh, U. C.; Wiener, P. K.; Caldwell, J. W.; Kollman, P. A. AMBER(UCSF), Version 3.0, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, 1986.

0002-7863/89/1511-8505$01.50/0

Table II

no.	hydrogen bonds[a] donor	acceptor	what happens Thr	Val
1	157 at HOG1	159 at OD1	OK, not reformed[b]	
2	155 at HOG	157 at OG1	OK, not reformed[b]	
3	159 at HN	157 at OG1	OK, not reformed[b]	
4	4–6 H_2O	159 at OD1, OD2	stable[c]	stable[c]
5	162 at HNZ3	159 at OD2	new[d]	
6	163 at HND1	159 at OD2		new[e]
7	163 at HND2	159 at OD2		new[e]
8	155 at HOG	151 at O	stable[f]	equilibration breaks, reformed[f]
9	155 at HN	151 at O	stable[f]	stable[f]
10	157 at HN	155 at OG1	stable[f]	stable[f]

[a] Groups are considered hydrogen bonded if the distance between heteroatoms is less than 3 Å. [b] Stable during the equilibration of Thr structure (see Figure 1b), broken when Thr turns to Val (see Figure 1c), not reformed when Val mutated to Thr (see Figure 1d). [c] Four to six H_2O molecules make stereochemically acceptable hydrogen bonds with the Asp-159 CO_2^- throughout the simulations. [d] During equilibration, Lys-162 NH_3^+ moves to hydrogen bond with Asp-159 OD2. [e] During the simulation, the amide group of Gln-163 moves to hydrogen bond with Asp-159 OD2. [f] These hydrogen bonds stay preserved throughout the simulations; although the 155 HOG···151 O hydrogen bond breaks during equilibration and then re-forms for the rest of the simulations (see Figures 1 and 2).

strain all bond lengths to their equilibrium values. The perturbation group[4] was taken to be the whole amino acid residue 157, but only interactions between this residue and other residues were included in the evaluation of $\Delta\Delta G$. Thus, we are assuming that the intragroup energies are similar in the native and denatured proteins.

Results

The results of the free energy calculations are shown in Table I. As one can see, mutating Thr → Val in water involves a free energy change of 4.70 ± 0.14 kcal/mol, mutating Thr → Val in the protein involves a free energy change of 6.53 ± 0.53 kcal/mol, and mutating Val → Thr in the protein involves a free energy change of −6.73 ± 1.50 kcal/mol. These quoted standard deviations are based on averaging the free energies for forward and backward simulations and do not refer to standard deviations within a given simulation.

The calculated $\Delta\Delta G$ of 1.91 ± 1.1 kcal/mol is in good agreement with experiment. What is the cause of the greater stability of the Thr-157 protein, compared to Val-157? It is clear that the structure of the native protein that Thr-157 is involved in a network of hydrogen bonds, as noted in Table II and Figure 1a: The Thr HOG donates a proton to the carboxylate oxygen of Asp-159 and OG accepts hydrogen bonds from Asp-159 NH and Thr-155 HOG. Despite the length and poor angle of the 157 HOG···159 CO_2^- hydrogen bonds discussed by Alber et al.,[2] and given that a charged residue is involved, its existence is not unreasonable. Is this hydrogen-bonding network the key to the greater stability of the Thr-157 protein? When one considers the electrostatic and van der Waals components to ΔG_{PN} and ΔG_{PD}, one finds specifically $\overline{\Delta G_{PN}}$(elec) = 4.16 kcal/mol and $\overline{\Delta G_{PN}}$(vdw) = 2.46 kcal/mol, whereas $\overline{\Delta G_{PD}}$(elec) = 3.94 kcal/mol and $\overline{\Delta G_{PD}}$(vdw) = 0.76 kcal/mol. Thus, the larger differential component is the van der Waals one. To further examine this point, we also mutated the partial charges of the Thr-157 C_β, and O_γ and HO_γ from their normal values of 0.17, −0.55, 0.31 to −0.07, 0.0, 0.0, respectively, in the native Thr-157 structure and in the tripeptide Gly-Thr-Trp in solution.[9] The $\Delta\Delta G$ associated with these mutations is 3.8 kcal/mol in the protein and 3.9 kcal/mol in the tripeptide model. Thus, there appears to be no differential stabilization of Thr versus Val from the hydrogen-bonding properties of Thr, given that hydrogen bonding in our force field is represented mainly by electrostatic terms. Alber et al.[2] suggested that some of the differential stability is due to the presence of the Asp NH to Thr OG hydrogen bond in the native protein, which is lacking in the Val-157 mutant. Thus, we mutated the charges on the Asp-159 NH group from their normal values on N (−0.463) and H (0.252) to −0.211 and 0.0 respectively.[9] The $\Delta\Delta G$ for this charge mutation is 22.5 kcal/mol in the native protein and 21.9 kcal/mol in the Val mutant. Thus, the unsatisfied hydrogen bonding of the Asp-159 NH can explain 0.6 kcal/mol of the instability of the Val mutant.

What is nature of this van der Waals contribution to protein stability? It could come from more effective packing of the Thr O_γ than Val C_γ in the protein interior, leading to stronger dispersion attraction. It also could come from greater exchange repulsion from the larger CH_3 group in the valine mutant.

The structural results are worthy of note and are given in Figures 1 and 2. In each of these figures we present the following: the X-ray structure (Figures 1a and 2a), the structure (Figures 1b and 2b) after 6 ps of molecular dynamics equilibration, to which hydrogen bonding and the water cap has been added, the structure after forward mutation (Figures 1c and 2c), and the structure after reverse mutation (Figures 1d and 2d). It is clear from Table II that the crystallographically observed hydrogen bonds involving residues other than Thr-157 are preserved in both Thr-157 and Val-157 protein structures. But Figures 1 and 2 and Table II show how the Thr structures after the Thr → Val → Thr (see Figure 1d) mutation or the Val → Thr mutation (see Figure 2c) do not re-form the crystallographically observed hydrogen bonds for the Thr-157 residue. Nonetheless, Thr-157 does form hydrogen bonds with H_2O to replace the crystallographic hydrogen bonds of the Thr-157. For example, in the structure after Thr → Val → Thr mutation (see Figure 1c), three H_2O molecules are within reasonable distance and have a reasonable orientation to interact favorably with the Thr OH.

Discussion

The results presented here suggest that, in some cases, one can use the thermodynamic cycle/perturbation method to analyze the effect of protein stability of site-specific mutation. However, a number of caveats should be emphasized. First, in contrast to the application to some cases of protein–ligand interactions, there is considerably more uncertainty in what structure to use to evaluate ΔG_{PN}, where one transforms the "denatured protein" into its site-specific mutant. We have made a simple attempt at this by using a tetrapeptide model for this region of the denatured protein. Obviously, in a 40-ps simulation one cannot span all the relevant conformations of the very floppy tetrapeptide. Furthermore, even if one could, it is not clear how relevant this sample would be to the actual denatured protein. Thus, all one can say at this point is that our result is reasonable and gives numbers that are in good agreement with experiment. One should note that our procedure may well give an "upper bound" for the amount of solvent exposure in the denatured protein, so that the calculations using this procedure are likely to overestimate $\Delta\Delta G$. Our calculations do overestimate it, but the error bars in the calculation make this overestimation not definitive. Other cases must be studied to see if the procedure consistently overestimates $\Delta\Delta G$.

A second uncertainty in the calculations is the structural hysteresis in the calculation of the native protein mutation ΔG_{PN}. There is a larger uncertainty in this calculated number, and it is an average of four different mutations. It is clear that the

(11) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 326.

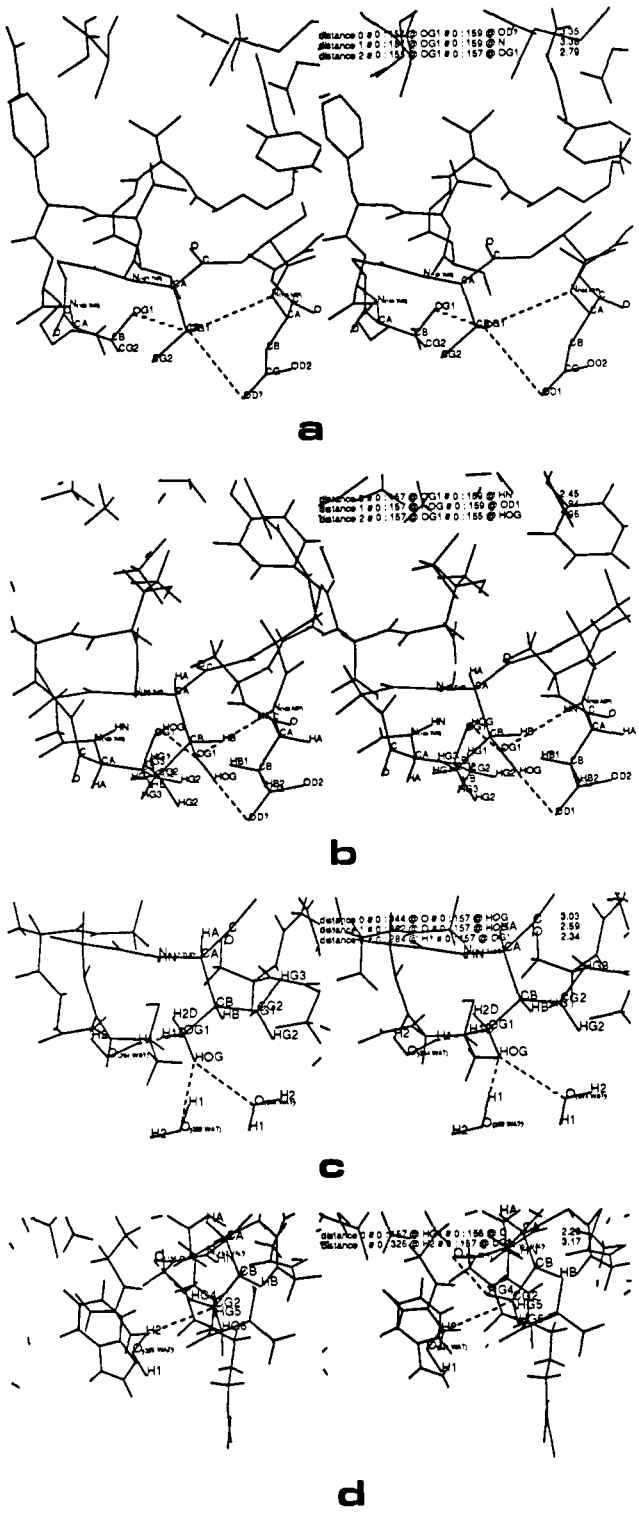

Figure 1. Structures of Thr-157 T4 lysozyme: (a) X-ray structure. (b) The equilibrated structure obtained from molecular dynamics simulations; the H1D and H2D are dummy protons and HOG is the "real" proton. (c) The final Thr-157 structure after mutation, i.e., Thr-157 → Val-157 → Thr-157. (d) The Thr-157 mutant obtained from free energy simulation of Val-157 → Thr-157; the HG5 and HG6 are the dummy protons and HG4 is the "real" proton.

hydrogen bond structure of the system with Thr → Val → Thr is quite different from the original Thr X-ray structure. Again, in 40 ps, one expects to limit sampling of the configuration of the system. Nonetheless, Thr-157 is located on the outside of the protein and it is not unreasonable to speculate that in solution, as opposed to the crystal, the hydrogen-bonding structure might be somewhat different and could have some aspects in common with the structure after Thr → Val mutations. Again, further analysis of this is required to address these issues, but to get a definitive answer may require orders of magnitude more simulation than carried out here.

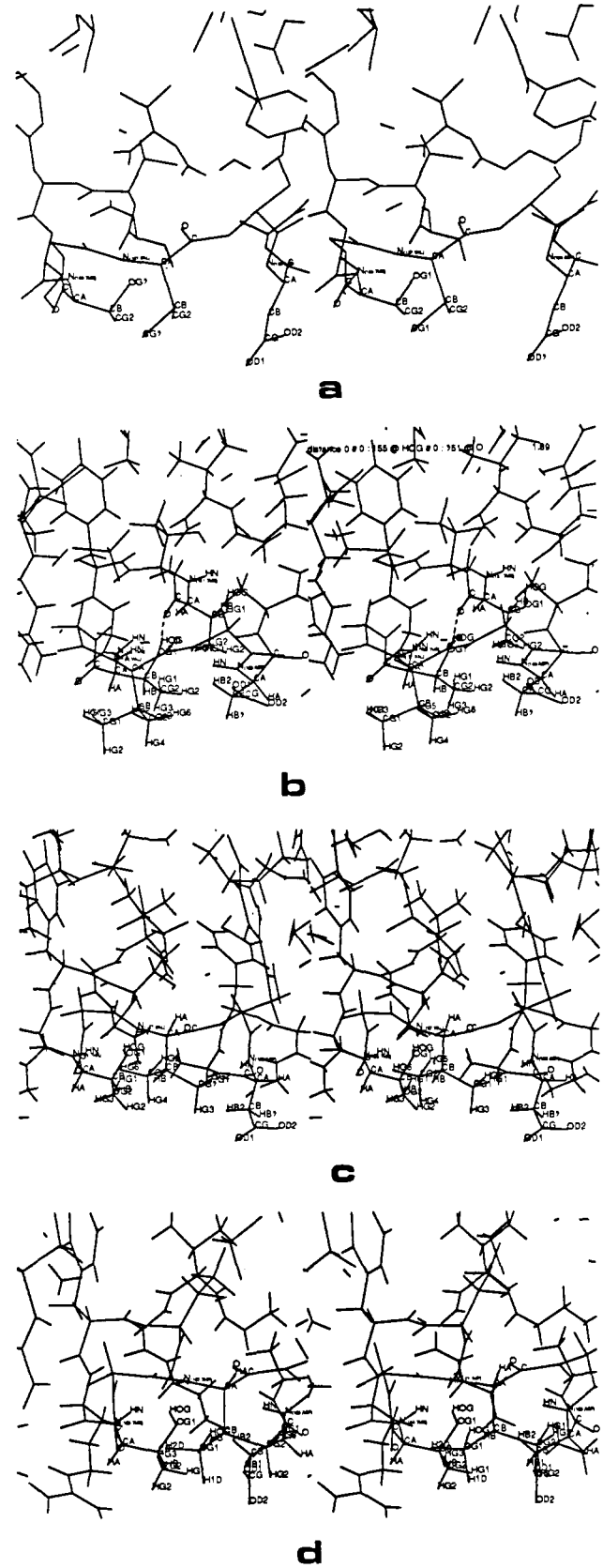

Figure 2. Structures of mutant Val-157 T4 lysozyme: (a) X-ray structure. (b) The equilibrated structure obtained from molecular dynamics simulations. (c) The final Val-157 structure after mutation, i.e., Val-157 → Thr-157 → Val-157. (d) The Val-157 mutant obtained from free energy simulation of Thr-157 → Val-157.

A third caveat concerns molecular mechanical force field and simulation protocol, involving the use of a simple molecular mechanical model and keeping the part of the protein further than 15 Å from the mutation frozen. This "belly approach" appears at least as reasonable and has seemed to work effectively in many different systems. Again, the molecular mechanical parameters are clearly far from perfect, but the use of the same parameters in ΔG_{PN} and ΔG_{PD} may allow for significant cancellation of errors.

One of the most significant and interesting results found in this study has been the fact that the differential stability $\Delta\Delta G$ is mainly a van der Waals rather than electrostic effect. How definite is this conclusion? By using the "slow-growth" procedure, we can unambiguously compute both the electrostatic and van der Waals contribution to the free energy changes. Thus, our calculated result is on firm ground. And it is further supported by our model mutation, in which we zero the partial charges on the O–H group in native and denatured states. Again, this result is consistent with actual mutation calculation in that the electrostatic contribution to the stability of both native and denatured models is nearly identical. It is clear that only the OH and not a CH_3 group can "fit" into the native structure at position 157 and gain more dispersion attraction and avoid exchange repulsion in this position.

One of the most important results of this study is the further demonstration of the power and utility of free energy component analysis.[12] No only can free energy calculations give free energies that can be related to those determined experimentally, but model calculations, such as the zeroing of the charges on the O–H group noted above and mutation of the Asp-159 N–H charges to zero in both Thr and mutant Val native structure, show the power of model calculations using this method to give new insight into protein stability. One of the important functions of theoretical calculation on molecules is to give useful mechanistic insights, and it is clear that the free energy approach fullfills this function in many cases.

(12) Daggett, V.; Brown, F.; Kollman, P. *J. Am. Chem. Soc.*, in press.

Conclusion

In summary, we have carried out free energy simulations on the Thr-157 → Val-157 mutation in T4 lysozyme. The calculations are successful in reproducing the experimental $\Delta\Delta G$ of protein stability, and this success suggests, at least in this case, a tripeptide model is adequate to represent the denatured protein. It seems counterintuitive that a hydrophilic residue like Thr will provide more stability to the native protein than to the more solvent-exposed denatured protein. Our calculations suggest that, even with the excellent hydrogen-bonding network of Thr-157 in the native protein, the electrostatic hydrogen bonding is no better than in the denatured native protein. Nonetheless, van der Waals energies contribute significantly to the differential stabilization of the Thr-157 than the Val-157 proteins because of greater dispersion attraction or less exchange repulsion of the smaller OH group compared to a CH_3 group. The calculations are less successful in their ability to reproduce all the hydrogen-bonding details of the crystal structures, which is not surprising given the limited time and simple environmental representation used in these simulations and the possibility that surface groups are indeed more mobile in solution than reflected in the crystal.

Acknowledgment. Most of the calculations were carried out at the Cornell National Supercomputer Center facility and at the San Diego Supercomputer Center through supercomputer support provided to P.A.K. by the NSF (Grant DMB-87-14775 and CHE-85-10066). Research support from the NIH (Grant GM-29072) to P.A.K. is acknowledged, as is a grant from Burroughs Wellcome Co. to P.A.K.

Registry No. Lysozyme, 9001-63-2.

Communications to the Editor

Highly Stereoselective Synthesis of Ganglioside GD_3[1]

Yukishige Ito,[†] Masaaki Numata,[‡] Mamoru Sugimoto,[‡] and Tomoya Ogawa*,[†]

RIKEN (The Institute of Physical and Chemical Research) Wako-shi, Saitama 351-01, Japan
Central Research Laboratory, MECT Co., Kitano Tokorozawa-shi, Saitama 359, Japan

Received May 22, 1989

Gangliosides, namely, sialic acid containing glycolipids, have attracted much attention because of the numerous biological roles they play in cellular recognition, differentiation, oncogenesis, and so on.[2] In spite of the structural diversity and prevalence in cell surfaces, synthetic studies toward such molecules have been limited to relatively simple ones.[3] This has been mainly due to the low efficiency encountered in the introduction of a sialic acid residue, especially when a secondary alcohol was used as a glycosyl acceptor.[4,5] Recently, we succeeded in obtaining a general solution for this critical problem, which affords α-glycosides of *N*-acetylneuraminic acid (NeuAc), the most representative in a sialic acid family, under nearly complete stereochemical control.[6] Our approach features an efficient use of the stereocontrolling phenylseleno or phenylthio auxiliary residing at the C-3 position of a NeuAc donor such as **9**. With this potent device in hand, our attention has been centered around the synthetic approach toward complex gangliosides with multiple NeuAc residues. Now, we report here the implementation of our strategy to the highly selective synthesis of ganglioside GD_3 (**1**), which has been isolated from various sources such as mammalian retina,[7a] bovine kidney,[7b]

* Author to whom correspondence should be addressed.
[†] RIKEN.
[‡] MECT Co.

(1) Synthetic Studies on Cell-Surface Glycans. **65**. Part 64: Kobayashi, M.; Yamazaki, F.; Ito, Y.; Ogawa, T. *Tetrahedron Lett.* In press.

(2) *Structure and Function of Gangliosides*; Svennerholm, L., Mandel, P., Dreyfus, H., Urbun, P.-F., Eds.; Plenum Publishing Co.: New York, 1980. *Sialic Acids, Chemistry, Metabolism and Function*; Schauer, R., Ed.; Springer-Verlag: Wien, 1982. *Ganglioside Structure, Function and Biomedical Potential*; Ledeen, R. W., Yu, R. K., Rapport, M. M., Suzuki, K., Eds.; Plenum Publishing Co.: New York, 1984.

(3) (a) Sugimoto, M.; Ogawa, T. *Glycoconjugate J.* **1985**, *2*, 5–9. (b) Sugimoto, M.; Numata, M.; Koike, K.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1986**, *156*, c1–c5. (c) Numata, M.; Sugimoto, M.; Koike, K.; Ogawa, T. *Carbohydr. Res.* **1987**, *163*, 209–225. (d) Numata, M.; Sugimoto, M.; Shibayama, S.; Ogawa, T. *Carbohydr. Res.* **1988**, *174*, 73–85.

(4) For leading references, see: (a) Ogawa, T.; Sugimoto, M. *Carbohydr. Res.* **1985**, *135*, c5–c9. (b) Paulsen, H.; von Deessen, U. *Carbohydr. Res.* **1986**, *146*, 147–153. (c) Okamoto, K.; Kondo, T.; Goto, T. *Tetrahedron* **1987**, *43*, 5909–5918, 5919–5928. (d) Kanie, O.; Kiso, M.; Hasegawa, A. *J. Carbohydr. Chem.* **1988**, *7*, 501–506.

(5) For enzymatic approaches, see: Simon, E. S.; Bednarski, M. D.; Whitesides, G. M. *J. Am. Chem. Soc.* **1988**, *110*, 7159–7163 and references cited therein.

(6) (a) Ito, Y.; Ogawa, T. *Tetrahedron Lett.* **1987**, *28*, 6221–6224. (b) Ito, Y.; Ogawa, T. *Tetrahedron Lett.* **1988**, *29*, 3987–3990. (c) A similar approach was recently reported by Goto et al.: Kondo, T.; Abe, H.; Goto, T. *Chem. Lett.* **1988**, 1657–1660.

0002-7863/89/1511-8508$01.50/0