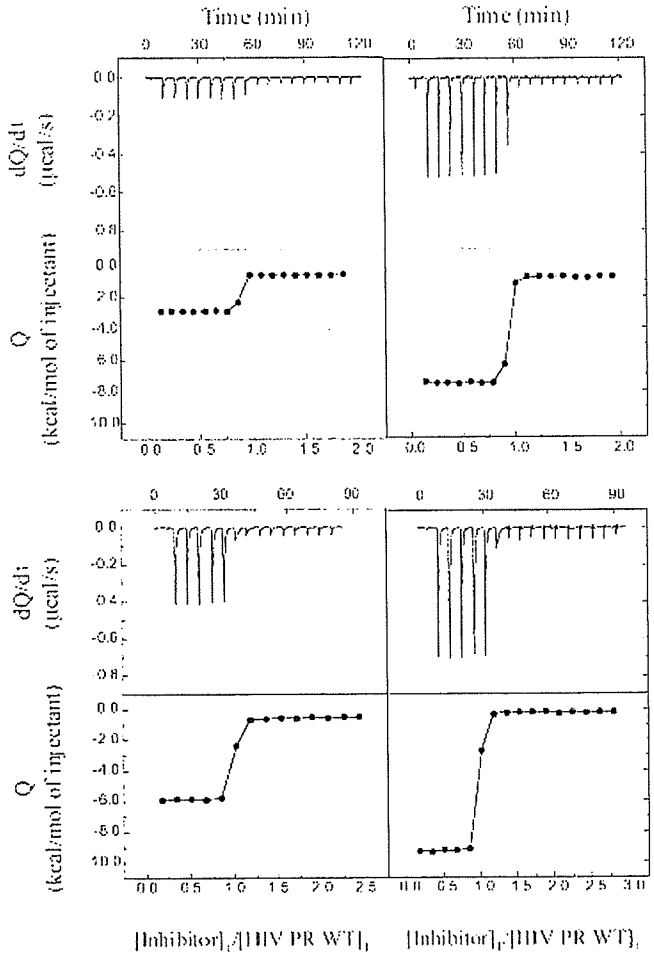


Figure 1. Typical calorimetric titrations of KNI-577 (top) and KNI-764 (bottom) into a solution of HIV-1 wild-type protease. The experiments were performed in sodium acetate 10 m*M*, pH 5.0 and DMSO 2%. The experiments on the left panels were performed at 20°C and on the right panels at 30°C. Additional experiments at 25°C were also used to determine the heat capacity change for each inhibitor against each protease used in these studies.

Crystallization

Crystals of HIV-1 protease WT were grown in hanging drops of equal amounts of reservoir solution (100 m*M* MOPS pH 7.0, 250 m*M* NaCl, 10 m*M* DTT, 3 m*M* NaN_3) and protein solution (6 mg/mL, 1 m*M* sodium acetate pH 5.0, 2 m*M* NaCl). Previously, the inhibitor (KNI-764 or KNI-577) was added to the protein solution from a concentrated (≈15 m*M*) stock solution in 100% DMSO to a final ratio of protease to inhibitor of 1:2. The drops were equilibrated against a volume of 1 mL of reservoir solution. DMSO was added to the reservoir to a final percentage equal to the one expected in the drops after equilibration (≈7%). Crystals grew in 1-2 days at 4°C.

Direct attempts to crystallize the V82F/I84V double mutant protease with the inhibitors KNI-764 and KNI-577 were not successful. In order to overcome this difficulty, crystallization was made with acetyl-pepstatin under the same conditions as reported above for the protease WT. Because acetyl-pepstatin is a less potent inhibitor than the KNI's, the ratio of fprotease to inhibitor was increased up to 1:10. Because the stock solution of acetyl-pepstatin (≈8 m*M*) was made in pure water, no DMSO was added to the reservoir. Crystals grew in 1-2 days at 4°C.

Once double mutant protease crystallized with acetyl-pepstatin, a ligand replacement technique was employed to obtain crystals of double mutant with KNI-764 and KNI-577. Crystals of protease with acetyl-pepstatin were washed and soaked in drops made with equal volumes of precipitant buffer and KNI-764 or KNI-577 solution (at 1 m*M* sodium acetate pH 5.0, 2 m*M* NaCl). Again, DMSO was added in the reservoir to the same percentage as the one expected in the drops after equilibration (≈7%).

Structure Determination

Four different data sets (wild-type and mutant HIV protease with either KNI-764 or KNI-577) were collected to 2.0 Å resolution on a RAxis IV Image Plate system using a rotating anode source equipped with Osmic mirrors (Molecular Structure Corporation, Texas). Diffraction data were processed with DENZO and SCALEPACK.[28] Autoindexing and consideration of systematically absent reflections revealed that the crystals belong to space group $P2_12_12$, unit-cell parameters a = 57.7 Å, b = 85.8 Å, c = 46.5 Å with two monomers in the asymmetric unit. Molecular replacement searches using the coordinates of the HIV protease from the crystal structure of the complex with inhibitor KNI-272 (PDB ID: 1HPX) as the search molecule yielded a clear solution.

An electron density map, calculated with phases based on the protein only, showed well-resolved density that allowed fitting all portions of the inhibitor as well as the side chains of mutated residues. Models consisting of protein coordinates plus those of the inhibitors were refined using the Crystallography and NMR System (CNS)[29] with a residual target that did not include non-crystallographic symmetry restraints. Rebuilding and correction of the models was guided by σA-corrected 2Fo-Fc electron density maps. R and R-free (calculated with randomly selected 10% of the reflections) were used to monitor refinement of the model.[30] The quality of the structures was assessed with the program PROCHECK.[31] The final refinement and model statistics are shown in Table I.

Coordinates

The coordinates have been deposited in the Protein Data Bank (accession codes 1MRW and 1MRX for the complex of KNI-577 with the wild type and drug resistant mutants, and 1MSM and 1MSN for the complex of KNI-764 with the wild type and drug resistant mutants, respectively).

RESULTS AND DISCUSSION

The Inhibitors

KNI-577 and KNI-764 are highly potent inhibitors of HIV-1 protease with K_i values in the subnanomolar range against wild-type enzyme.[10] At 25°C, KNI-577 and KNI-764 bind to wild-type protease with K_d values of 0.2 and 0.03 n*M* respectively. Structurally, the two inhibitors are identical except for the functional group at the P2′ position

TABLE I. Statistics for Data Collection and Refinement

Data sets	HIV Protease WT	HIV Protease 82/84MT	HIV Protease WT	HIV Protease 82/84MT
Inhibitor	KNI-764	KNI-764	KNI-577	KNI-577
Space group		P21212		
Cell constants		a = 58.1Å b = 86.0Å c = 46.5Å		
Resolution	20–2.0Å	20–2.0Å	20–2.0Å	20–2.0Å
Observed reflections	122330	121342	126994	81856
Unique reflections	15861	15875	16090	13461
Completeness (%)	97.7	98.5	98.3	82.9
(High resolution shell)	87.7	95.9	88.7	75.8
R sym[1]	0.067	0.078	0.103	0.072
Refinement				
R crystal/R free	0.209/0.236	0.222/0.249	0.218/0.246	0.208/0.250
Model composition[2]				
Amino acids (atoms)	198 (1516)	198 (1522)	198 (1516)	198 (1522)
Ligands (atoms)				
KNI-764 or KNI-577	1 (41)	1 (41)	1 (37)	1 (37)
Waters	147	123	112	128
Total atoms	1704	1686	1665	1687
Stereochemistry[3]				
r.m.s. Bond length (Å)	0.0063	0.0066	0.0065	0.0062
r.m.s. Angles (°)	1.22	1.28	1.31	1.23
<B-factor protein>	24.2	27.2	28.3	26.2
<B-factor water>	31.5	35.2	33.6	34.1
<B-factor ligand>	17.7	25.6	26.0	22.2

[1] $R_{sym} = (\Sigma|I_{hkl} - \langle I \rangle|)/(\Sigma I_{hkl})$, where I_{hkl} is the observed intensity and $\langle I \rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections.
[2] Model composition shows two monomers in asymmetric unit.
[3] Over 90% of main chain dihedrals fall within the 'most allowed regions' of the Ramachandran plot.

(see Fig. 2). In KNI-577, this position is occupied by a symmetric t-butylamide group, whereas in KNI-764 the position is occupied by a larger methylbenzylamide group. In KNI-577, there is one rotatable bond between the amide and the t-butyl group, and in KNI-764 there are two rotatable bonds between the amide and the phenyl ring. Both inhibitors bind to the protease in a process favored by enthalpic and entropic interactions.[10] At 25°C, the differences in affinity are due to the more favorable binding enthalpy of KNI-764 (2.9 kcal/mol) (see Fig. 1 and Table II), suggesting that this inhibitor establishes better interactions with wild-type protease than does KNI-577. KNI-577 on the other hand, exhibits a more favorable entropic contribution to the Gibbs energy of binding than does KNI-764 (1.8 kcal/mol), but this is not sufficient to overcome the more favorable binding enthalpy of KNI-764.

Analysis of the crystal structures of KNI-577 and KNI-764 bound to wild-type protease reveals that a group of conserved water molecules at the inhibitor–protein interface (Fig. 3) plays a crucial role in establishing a tight atomic packing density within the binding cavity. These water molecules have been shown to be immobilized within NMR timescales in the case of another allophenylnorstatine (Apns)-based inhibitor.[13] Also, because these water molecules establish important hydrogen bonds and van der Waals contacts between inhibitor and protein, they have been shown to contribute significantly to the observed binding enthalpy.[14]

Inhibitor Response to V82F/I84V Drug Resistant Mutation

Despite their structural similarities, the two inhibitors respond differently to the same drug resistant mutation V82F/I84V. Figures 3 and 4 show the structures of the inhibitors bound to wild-type and V82F/I84V protease mutant. Previously, Reiling et al[15] determined the structure of KNI-764 with a version of wild-type protease differing in six residues from the one reported here.

Although the structures of the four complexes reported here are highly similar, small localized differences are observed among them (Figs. 3-4). In all complexes, both monomers (A and B) of the HIV-1 protease interact with the inhibitor. In the case of the complexes with KNI-764, residue 82 (monomer A) interacts with the phenyl ring of the inhibitor, and residue 84 (monomer A) interacts with the same phenyl ring as well as with the methylbenzylamide moiety. In monomer B, residue 84 interacts with the dimethylthioproline group, but residue 82 does not interact with the inhibitor. As a result of these interactions, the substitutions present in the mutant (V82F/I84V) bring about small but significant changes in both the position and the conformation of the inhibitor. The largest changes involve movements of the aromatic rings of the inhibitor. The rotation in the plane of the ring of the methylbenzylamide by about 25° results in displacements of several atoms by up to 1.3 Å. Displacements of the same magnitude are observed for the 3-hydroxy-2-methylbenzoyl moiety, and

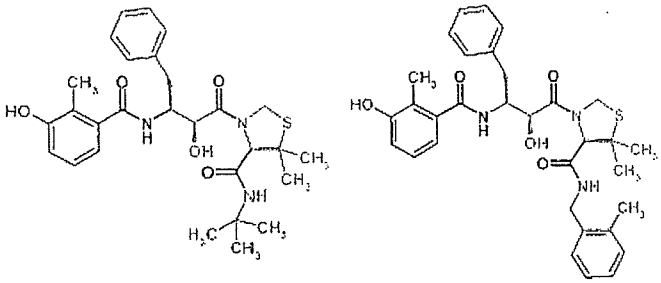


Figure 2. The chemical structure of KNI-577 (left) and KNI-764 (right). Both inhibitors share the same allophenyl-norstatine scaffold at the P1 position (red) and the same functional groups at the P2 (blue) and P1′ positions (green). The only difference is at the P2′ position (magenta).

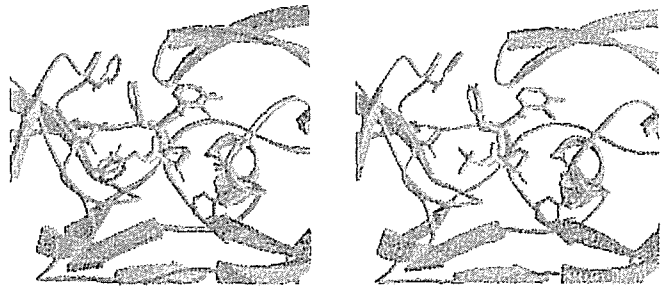


Figure 4. Superposition of the complexes of KNI-764 (left) and KNI-577 (right) bound to wild-type (cyan) and double mutant (purple) protease.

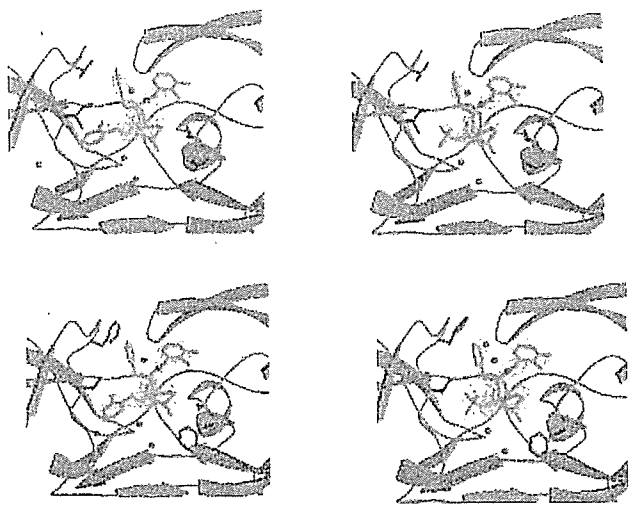


Figure 3. Complexes of KNI-764 and KNI-577 bound to wild-type (upper panels) and V82F/I84V double mutant (lower panels) HIV-1 protease. The electronic densities of KNI-764 (left panels) and KNI-577 (right panels) bound to the protease are shown. Water molecules considered in the structure-based thermodynamic analysis are shown in red.

smaller changes occur for the phenyl ring in Apns and the dimethylthioproline moieties.

In the complexes with KNI-577 all of the differences between wild type and mutant are smaller than those observed with KNI-764 with one exception: the displacement of the phenyl ring in Apns is significantly larger than in the other complex. In the rest of the structure, including the 3-hydroxy-2-methylbenzoyl moiety and the dimethylthioproline, the changes are in general less than 0.3 Å. The *t*-butyl group, which in KNI-577 replaces the methylbenzylamide of KNI-764, hardly moves at all compared to the large movement experienced by the methylbenzylamide of KNI-764. The effect of these structural changes on the differences in thermodynamic properties of the four complexes is discussed below.

Thermodynamic Response to Drug Resistant Mutation

KNI-764 and KNI-577 lose significant enthalpic interactions (2.3 and 2.6 kcal/mol, respectively) when facing the drug resistant mutant (Table II). However, KNI-764 is able to partially compensate for the loss in binding enthalpy by a gain in binding entropy. When the effects of the enthalpy loss and the entropy gain are combined, it becomes evident why KNI-764 loses affinity by only a factor of 25 while KNI-577 loses affinity by a factor of 260. In absolute terms, KNI-764 maintains a nanomolar affinity against the drug-resistant mutant while the affinity of KNI-577 becomes close to 100 n*M*. In fact, KNI-764 inhibits the V82F/I84V drug-resistant mutant with an affinity similar to that of first-generation protease inhibitors against wild-type protease.[9–11] The affinity of KNI-577, on the other hand, weakens below the level required for effective inhibition.

Isothermal titration calorimetric experiments performed at different temperatures (Fig. 5) indicate that the binding of both inhibitors is associated with a negative change in heat capacity. The magnitude of the heat capacity change is different for each inhibitor, and their response to the drug-resistant mutant is also different. KNI-764 exhibits a ΔC_p of -360 ± 20 cal K^{-1} mol^{-1} against the wild type and -390 ± 15 cal K^{-1} mol^{-1} against the mutant. KNI-577, on the other hand, exhibits a ΔC_p of -420 ± 20 cal K^{-1} mol^{-1} against the wild type and -380 ± 20 cal K^{-1} mol^{-1} against the mutant, i.e. the ΔC_p value for KNI-764 decreases while the ΔC_p value for KNI-577 increases.

The solvent-accessible surface areas of KNI-577 and KNI-764 are 760 and 820 Å^2 respectively, of which 598 and 657 Å^2 correspond to non-polar surface. Upon binding to the wild-type enzyme, KNI-577 buries 573 Å^2 of its non-polar surface and 163 Å^2 of its polar surface. KNI-764, on the other hand, buries 607 Å^2 of its non-polar surface and 162 Å^2 of its polar surface. Against the V82F/I84V mutant KNI-577 buries 558 Å^2 of its non-polar surface, while KNI-764 buries 619 Å^2. The amount of buried polar surface remains unchanged for both inhibitors. The actual increase in the non-polar surface area buried by KNI-764 against the mutant protein reveals the adaptability of this inhibitor by its capacity to bury itself deeper into the mutant binding cavity. At the thermodynamic level, this structural feature provides an important contribution to the compensating entropy change. The opposite effect is seen with KNI-577, which buries a smaller non-polar surface area against the mutant protease.

TABLE II. Binding Thermodynamics of Inhibitors to Wild Type and V82F/I84V Drug-Resistant Mutant at 25°C

	ΔG (cal/mol)	ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	ΔC_p [cal k^{-1} mol^{-1}]
KNI-764 → WT	−14300 ± 100	−7600 ± 300	−6700 ± 300	−360 ± 20
KNI-764 → V82F/I84V	−12400 ± 100	−5300 ± 200	−7100 ± 200	−390 ± 15
KNI-577 → WT	−13200 ± 100	−4700 ± 200	−8500 ± 200	−420 ± 20
KNI-577 → V82F/I84V	−9900 ± 100	−2100 ± 200	−7800 ± 200	−380 ± 20

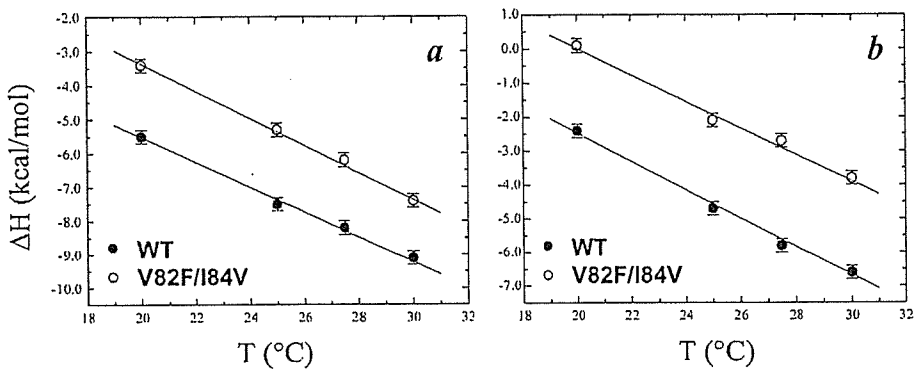


Figure 5. Determination of the change in heat capacity (ΔC_p) upon binding for a) KNI-764 and b) KNI-577 to wild-type and V82F/I84V mutant proteases. ΔC_p is equal to the slope of the enthalpy versus temperature plot.

The loss in binding enthalpy against the mutant experienced by both inhibitors can be explained in terms of the loss of van der Waals contacts when the inhibitors, selected against the wild-type enzyme, face an altered binding cavity. The gain in entropy experienced by KNI-764 could originate from several factors. An increase in solvation entropy would be consistent with the more negative heat capacity change observed with the mutant.[16] Another potential source of compensation is an increase in conformational entropy. One can envision a situation in which the loss of interactions is accompanied by an increase in the conformational degree of freedom of ligand and protein, especially if the loss occurs in non-conformationally constrained regions of the inhibitor. In this respect we observe an increase in the crystallographic B-factors for the KNI-764/mutant complex, while the opposite is true for KNI-577.

Structure-Based Thermodynamic Analysis

The crystallographic structures and the thermodynamic data discussed here indicate that the partial compensation of the binding enthalpy loss with an actual gain in binding entropy provides the thermodynamic mechanism to escape the deleterious effects of the V82F/I84V drug resistant mutation, and that this behavior originates from the better adaptability of KNI-764 to the binding site distortions introduced by the mutations. Recently, a refined structural parameterization of the enthalpy change associated with the binding of small ligands was presented.[14] Four set of parameters were considered to be important for a quantitative account of the binding enthalpy: (1) the interactions between ligand and protein, reflected in changes in solvent accessible surface areas for ligand and protein; (2) the conformational change associated with binding; (3) the presence of water molecules at the protein–ligand interface (a cutoff of 6 Å for completely buried water molecules was reported); and (4) any effects due to protonation/deprotonation associated with ligand binding.

In order to assess whether the thermodynamic response can be explained quantitatively in terms of structure, a parameterization of the binding energy developed previously was utilized. If the parameters reported by Luque and Freire[14] [$\Delta H(25) = \Delta H_{conf} - 7.35\Delta ASA_{np} + 31.06 \times \Delta ASA_{pol}$] are applied to the four structures presented here, the results shown in Figure 6 are obtained. In all cases, the reported conformational enthalpy change (ΔH_{conf}) of 5900 cal/mol for the protease molecule upon binding was used.[14] Protonation effects for these two inhibitors do not contribute significantly to the binding enthalpy, as demonstrated by microcalorimetric titrations performed with buffers having different heats of ionization (data not shown).[17] According to the published report,[14] water molecules at protein–inhibitor interface contribute significantly to the binding enthalpy. In the KNI-764 complex with the wild type, six water molecules are buried within 6 Å of the inhibitor (2, 5, 10, 27, 99, 148) while in the complex with the drug resistant mutant only 5 water molecules (11, 22, 47, 91, 111) are buried. In the complex of KNI-577 with the wild–type protease, five water molecules are buried within 6 Å (1, 5, 6, 49, 84), and the same number are found with the resistant mutant (6, 7 57, 87, 123). The water molecules that satisfy the published criteria are shown in Figure 3. As shown in Figure 6, the parameterization predicts the binding enthalpies within 1.4 kcal/mol of the experimental value. For KNI-577 against the mutant protease, the error is somewhat larger and amounts to ≈2 kcal/mol. In all cases, the loss of binding enthalpy against the

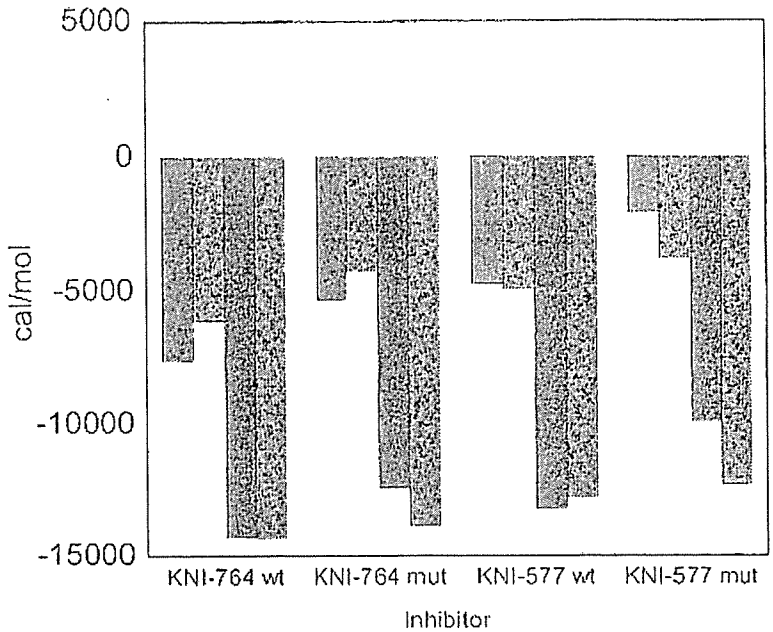


Figure 6. Experimental and calculated binding enthalpies (open and hatched bars) and binding Gibbs energies (solid and crossed bars). The standard deviation between experimental and calculated enthalpies is 1.4 kcal/mol. The standard deviation between experimental and calculated Gibbs energies is 1.3 kcal/mol.

TABLE III. Structure-Based Thermodynamic Analysis of Inhibitor Complexes with Wild Type and Drug-Resistant Mutant at 25°C

	KNI-764 wild type	KNI-764 mutant	KNI-577 wild type	KNI-577 mutant
ΔASA_{pol}[1] Å^2	−588 (−322)	−535 (−333)	−527 (−337)	−499 (−325)
ΔASA_{np}[1] Å^2	−848 (−889)	−884 (−932)	−756 (−814)	−789 (−829)
ΔH_{int} kcal/mol	−12.0	−10.1	−10.8	−9.7
$-T\Delta S_{solv}$ kcal/mol	−23.3	−25.0	−20.7	−22.1
$-T\Delta S_{conf}$ kcal/mol	18.4	18.5	16.1	16.8
$-T\Delta S_{rt}$ kcal/mol	2.4	2.4	2.4	2.4
ΔG_{el} kcal/mol	0.2	0.4	0.2	0.3
ΔG_{calc} kcal/mol	−14.3	−13.8	−12.8	−12.3

[1]The changes in solvent accessible surface area include the contributions due to the water molecules at the protein/inhibitor interface. The number in parenthesis represent the changes in solvent accessible surface areas calculated without taking into consideration the water molecules at the protein/inhibitor interface. ΔH_{int} is the intrinsic binding enthalpy assuming that the protease is in the same conformation in the free and bound states. The experimental enthalpy is calculated by adding the enthalpy corresponding to the change in conformation (5.9 kcal/mol) to this value. ΔS_{solv} and ΔS_{conf} are the calculated changes in solvation and conformational entropy, respectively. ΔS_{rt} correspond to the loss of translational degrees of freedom and is approximated by the cratic entropy[16]. ΔG_{el} is the coulombic contribution to the Gibbs energy[16].

drug resistant mutant is correctly predicted. Each water molecule optimizes the atomic packing density between inhibitor and protein and contributes on the order of −1.2 kcal/mol to the binding enthalpy. Against the drug resistant mutation, KNI-764 loses one of the buried water molecules but is able to partially compensate for this loss by accommodating itself within the binding pocket. It ends up suffering a smaller enthalpy loss than KNI-577, which maintains the same number of buried water molecules with the mutant.

The global changes in solvent accessible surface area are shown in Table III. These changes include the sum protein contributions, as well as those made by the inhibitor, taking into consideration the water molecules at the protein–inhibitor interface. For comparison, the global changes calculated without considering the water molecules are also shown. It is clear that the inclusion of the water molecules makes a significant contribution, especially to the polar surface that is buried upon binding. It is also seen that against the mutant the enthalpy loss originates from a diminished number of polar contacts reflected in a loss in buried polar surface area coupled to a gain in buried nonpolar surface area. The larger error observed for KNI-577 against the mutant protease could be due to an overestimation of the number of water molecules buried at the interface. As a result of this error, KNI-577 is predicted to lose ≈1 kcal/mol less than KNI-764 against the mutant. In the

TABLE IV. Structure-Based Dissection of Heat Capacity Changes

	$\Delta C_{p,\mathrm{exp}}$ (cal k^{-1} mol^{-1})	$\Delta C_{p,\mathrm{solv}}$ (cal k^{-1} mol^{-1})	$\Delta\Delta C_p$ (cal k^{-1} mol^{-1})	$\sigma(\Delta H)$ kcal/mol
KNI-764 → WT	−360 ± 20	−220	140	4.9
KNI-764 → V82F/I84V	−390 ± 15	−251	139	4.9
KNI-577 → WT	−420 ± 20	−196	224	6.3
KNI-577 → V82F/I84V	−380 ± 20	−218	162	5.3

$\Delta\Delta C_p = \Delta C_{p,\mathrm{solv}} - \Delta C_{p,\mathrm{exp}}$
$\sigma(\Delta H) = \sqrt{RT^2\,\Delta\Delta C_p}$

structure-based calculations presented here, we have applied rigid cutoff criteria derived from a statistical analysis of different protein–ligand complexes[14] (completely buried waters within 6 Å of inhibitor), according to which KNI-764 loses one water molecule against the mutant but KNI-577 does not. If KNI-577 also loses one water molecule, the predicted enthalpy would be ≈−2.6 kcal/mol, which is much closer to the experimental value of −2.1 kcal/mol. It seems that a critical factor required to calculate enthalpy changes from structure is an accurate determination of the number of buried water molecules in a complex. A statistical-based criterion like the one used here may not work in every situation since some water molecules may exchange with the solvent even if they are close to the inhibitor and appear to be completely buried from the solvent in the crystal structure.

The binding of KNI-764 and KNI-577 is favored by entropic contributions to the Gibbs energy of −6.7 and −8.5 kcal/mol respectively. The overall entropy contribution is composed of two major terms, a favorable term associated with the desolvation of the inhibitors and some regions of the protein and an unfavorable term associated with the loss of conformational and rotational/translational degrees of freedom. In our calculations, the binding energy was approximated by assuming that the Gibbs energy associated with the closing of the flaps was very small. This assumption is supported by NMR data[18] indicating that the flaps exist in rapid dynamic equilibrium and that the fully closed conformation is also present in solution. The solvation entropy change was calculated from the changes in solvent accessible surface as described before[16] including a correction term for the new enthalpy parameterization, and is shown in Table III. Conformational entropy changes were calculated as described before[16,19,20,21] and are also shown in Table III. According to this analysis, KNI-764 was correctly predicted to gain solvation entropy with the drug resistant mutant. Again, the error obtained for KNI-577 against the drug-resistant mutation is larger than the one obtained for KNI-764. The overall change in conformational entropy was calculated to be larger for KNI-764 than KNI-577. This is an expected result since KNI-764 has an additional rotatable bond and a larger interaction surface with the protease. For each inhibitor, the conformational entropy changes were predicted to be similar for wild type and mutant proteases within the error of the calculations.

Overall, the structure-based calculations predict the binding thermodynamics of the inhibitor to wild-type protease with higher accuracy than to the mutant (Fig. 6 and Table III). Against the wild type, the error in ΔG is 0.2 kcal/mol, whereas against the mutant it is 1.9 kcal/mol and always an overestimation. In our calculations, we have assumed that the conformational energetics of wild type and mutant protein are the same. In fact, this may not be the case. Previously, Todd et al[9] determined that the native conformation of the V82F/I84V mutant of the HIV-1 protease is more stable than the wild type protease and if this difference is due to a higher stability of the open conformation it will lower the affinity of inhibitors by about a factor of ten or 1.4 kcal/mol. If this is indeed the case, the error in the structure-based calculation of ΔG will go down to 0.5 kcal/mol and will be close to the one observed for the wild-type protease.

The heat capacity values calculated from the changes in accessible surface areas listed in Table IV underestimate the experimental values, suggesting that other effects in addition to changes in solvation contribute to the observed values. Since linked protonation reactions are negligible, a reasonable alternative is the presence of a significant contribution associated with the change in conformation of the protease between the free and bound states. In the unbound state, the protease exists in a highly dynamic conformational equilibrium that defines a conformational ensemble characterized by significant fluctuations, particularly in the flap region;[18] in the bound state those conformational fluctuations are significantly dampened.[18] Since structural fluctuations are coupled to enthalpy fluctuations, these changes would be reflected in the heat capacity of the protease because heat capacity is proportional to the average square fluctuation of the enthalpy:[22,23]

$$C_p = \frac{<\Delta H^2> - <\Delta H>^2}{RT^2} \qquad (1)$$

If the differences between the experimental heat capacities and those calculated from the changes in solvation are assumed to be due to diminished structural fluctuations in the bound state, a reduction of ≈5.4 kcal/mol in the width of the enthalpy distribution of the protease would be expected upon going from the unbound to the bound state. The typical width of the enthalpy distribution for a globular protein of similar molecular weight is on the order of 30 kcal/mol.[23]

While water molecules buried at the protein–ligand interface contribute significantly to the binding enthalpy, they appear not to contribute significantly to the Gibbs energy of binding. The enthalpy gain is compensated by

the entropic loss associated with trapping the water molecules within the protein structure. If the structure-based thermodynamic analysis is performed without considering the water molecules using a procedure described previously,[16] the enthalpy is underestimated by 6-7 kcal/mol; however, the Gibbs energy is calculated reasonably accurately (1.9 kcal/mol standard deviation). If the water molecules are included in the analysis and their entropy loss is considered an adjustable parameter, only a marginal improvement in the calculated Gibbs energy is obtained (1.3 kcal/mol), suggesting that the enthalpic gain is significantly compensated by the entropy loss of immobilizing the water molecules at the protein–ligand interface. Similar observations regarding the role of interface water have been made recently for antigen–antibody complexes.[24]

CONCLUSIONS

The structural and thermodynamic information presented here describes a mechanism by which an inhibitor can minimize the effects of mutations conferring drug resistance. At the thermodynamic level, the detrimental effects of the mutation on the binding affinity are minimized by the phenomenon of enthalpy/entropy compensation. At the structural level, the thermodynamic response originates from the better adaptability of KNI-764 compared to KNI-577. The better adaptability of KNI-764 is due to the presence of an asymmetric functionality and an additional rotatable bond at the same position where KNI-577 has only one rotatable bond and a symmetric functionality. While inhibitor flexibility is important for adaptability, it must be noted that it should be located at positions facing variable regions of the target and that the asymmetry of the functional group provides a significantly larger number of alternate conformations that are able to establish strong interactions with the target molecule.

REFERENCES

1. Kaplan AH, Michael SF, Wehbie RS, Knigge MF, Paul DA, Everitt L, Kempf DJ, Norbeck DW, Erickson JW. Selection of multiple human immunodeficiency virus type 1 variants that encode viral proteases with decreased sensitivity to an inhibitor of the viral protease. Proc Natl Acad Sci 1994;91:5597–5601.
2. Ho DD, Toyoshima T, Mo H, Kempf DJ, Norbeck D, Chen C, Wideburg NE, Burt SK, Erickson JW, Singh MK. Characterization of human immunodeficiency virus type 1 variants with increased resistant to a C2-symmetric protease inhibitor. J Virol 1994;68:2016–2020.
3. Condra JH, Schleif WA, Blahy OM, Gabryelski LJ, Graham DJ, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M, Titus D, Yang T, Teppler H, Squires KE, Deutsch PJ, Emini EA. *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature 1995;374:569–571.
4. Roberts NA. Drug-resistance patterns of saquinavir and other HIV proteinase inhibitors. AIDS 1995;9:S27–S32.
5. Tisdale M. HIV protease inhibitors - resistance issues. Int Antiviral News 1996;4:95–107.
6. Hong L, Treharne A, Hartsuck JA, Foundling S, Tang J. Crystal structures of complexes of a peptidic inhibitor with wild-type and two mutant HIV-1 proteases. Biochemistry 1996;35:10627–10633.
7. Ala PJ, Huston EE, Klabe RM, McCabe DD, Duke JL, Rizzo CJ, Korant BD, DeLoskey RD, Lam PYS, Hodge CN, Chang CH. Molecular basis of HIV-1 protease drug resistance: structural analysis of mutant proteases complexed with cyclic urea inhibitors. Biochemistry 1997;36:1573–1580.
8. Jadhav PK, Ala P, Woerner FJ, Chang CH, Garber SS. Cyclic urea amides: HIV-1 protease inhibitors with low nanomolar potency agains both wild type and protease inhibitor resistant mutants of HIV. J Med Chem 1997;40:181–191.
9. Todd MJ, Luque I, Velazquez-Campoy A, Freire E. The thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 2000;39:11876–11883.
10. Velazquez-Campoy A, Freire E. Incorporating target heterogeneity in drug design. J Cell Biochem 2001;S37:82–88.
11. Velazquez-Campoy A, Kiso Y, Freire E. The binding energetics of first and second generation HIV-1 protease inhibitors: implications for drug design. Arch Biochem Biophys 2001;390:169–175.
12. Ohtaka H, Velazquez-Campoy A, Xie D, Freire E. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of amprenavir and TMC-126 to wild type and drug-resistant mutants of the HIV-1 protease. Protein Science 2002;11:1908–1916.
13. Wang YX, Freedberg DI, Wingfield PT, Stahl SJ, Kaufman JD, Kiso Y, Bhat TN, Erickson JW, Torchia DA. Bound water molecules at the interface between the HIV-1 protease and a potent inhibitor, KNI-272, determined by NMR. J Am Chem Soc 1996;118:12287–12290.
14. Luque I, Freire E. Structural parameterization of the binding enthalpy of small ligands. Proteins 2002;49:181–190.
15. Reiling KK, Endres NF, Dauber DS, Craik CS, Stroud RM. Anisotropic dynamics of the JE-2147-HIV protease complex: drug resistance and thermodynamic binding mode examined at 1.09 Å structure. Biochemistry 2002;41:4582–4594.
16. Luque I, Freire E. A system for the structure-based prediction of binding affinities and molecular design of peptide ligands. Methods Enzymol 1998;295:100–127.
17. Velazquez-Campoy A, Luque I, Todd MJ, Milutinovich M, Kiso Y, Freire E. Thermodynamic dissection of the binding energetics of KNI-272, a powerful HIV-1 protease inhibitor. Protein Sci 2000;9:1801–1809.
18. Freedberg DI, Ishima R, Jacob J, Wang YX, Kustanovich I, Louis JM, Torchia DA. Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamic calculations. Protein Sci 2002;11:221–232.
19. Lee KH, Xie D, Freire E, Amzel LM. Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. Proteins Struct Func and Genetics 1994;20:68–84.
20. DAquino JA, Gómez J, Hilser VJ, Lee KH, Amzel LM, Freire E. The magnitude of the backbone conformational entropy change in protein folding. Proteins 1996;25:143–156.
21. DAquino JA, Freire E, Amzel LM. Binding of small molecules to macromolecular targets: evaluation of conformational entropy changes. Proteins 2000;4:93–107.
22. Hill TL. Introduction to statistical thermodynamics. Addison-Wesley: Reading, MA, 1960.
23. Gomez J, Hilser JV, Xie D, Freire E. The heat capacity of proteins. Proteins: Struc, Func and Genetics 1995;22:404–412.
24. Yokota A, Tsumoto K, Shiroishi M, Kondo H, Kumahai I. The role of hydrogen bonding via interfacial water molecules in antigen/antibody complexation. J Biol Chem 2003;278:5410–5418.
25. Todd M, Semo N, Freire E. The structural stability of the HIV-1 protease. J Mol Biol 1998;283:475–488.
26. Velazquez-Campoy A, Vega S, Freire E. Amplification of the effects of drug-resistance mutations by background polymorphisms in HIV-1 protease from African subtypes. Biochemistry 2002;41:8613–8619.
27. Sigurskjold BW. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. Anal Biochem 2000;277:260–266.
28. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;277:307–326.
29. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J, Kuszewski J, Nilges M, Pannu NS, Read R, Rice L, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystall D 1998;54:905–921.
30. Brunger AT. Free R value: cross-validation in crystallography. Methods Enzymol 1997;277:366–396.
31. Laskowski R. PROCHECK: a program to check the sterochemical quality of protein structures. J Appl Cryst 1993;26:283–291.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters
Tetrahedron Letters 45 (2004) 5965–5968

Design and synthesis of a novel water-soluble Aβ1-42 isopeptide: an efficient strategy for the preparation of Alzheimer's disease-related peptide, Aβ1-42, via *O–N* intramolecular acyl migration reaction

Youhei Sohma, Masato Sasaki, Yoshio Hayashi,* Tooru Kimura and Yoshiaki Kiso*

Department of Medicinal Chemistry, Center for Frontier Research in Medicinal Science, Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607-8412, Japan

Received 19 May 2004; revised 9 June 2004; accepted 11 June 2004

Abstract—A novel water-soluble isopeptide of Alzheimer's disease-related peptide Aβ1-42, '26-*O*-acyl isoAβ1-42', which could efficiently convert to intact Aβ1-42 under physiological conditions via *O–N* intramolecular acyl migration, was synthesized providing a new system useful for investigation of biological function of Aβ1-42.

Amyloid β peptides (Aβ) are the main proteinaceous component of the amyloid plaques found in the brains of Alzheimer's disease (AD) patients.[1] Neuritic plaques, pathognomonic features of AD, contain abundant fibrils formed from Aβ, which have been found to be neurotoxic in vivo and in vitro.[2] The predominant forms of Aβ consist of mainly 40- and 42-residue peptides (designated Aβ1-40 and Aβ1-42, respectively), which are proteolytically produced from amyloid precursor protein (APP) by enzymatic reactions.[3] Since Aβ1-42 is suggested to play a more critical role in amyloid formation and the pathogenesis of AD than Aβ1-40, many studies using synthetic Aβ1-42 have been carried out to clarify the involvement of Aβ1-42 in AD.[4]

However, Aβ1-42 is a highly hydrophobic peptide and forms aggregates in various media. This aggregation is attributed to its intermolecular hydrophobic interaction and hydrogen bond formation among peptide chains, leading to the formation of extended β-sheet structures.[5] Hence, the highly agglutinative potency of Aβ1-42 results in the synthetic difficulty of this peptide,[5] so-called 'difficult sequence'-containing peptide.[6] Due to the low solubility and broad elution of Aβ1-42 under acidic or neutral conditions, the conventional HPLC purification of synthesized Aβ1-42 in the aqueous TFA–acetonitrile system is too laborious to remove impurities accumulated during the solid-phase peptide synthesis (SPPS). Furthermore, the biological experiments using Aβ1-42 are problematic because of the large extent of aggregation in a standard storage solution such as dimethylsulfoxide (DMSO).[7] Therefore, the 'in situ' system that can prepare intact Aβ1-42 in solubilized form under physiological conditions would be a powerful tool in understanding the inherent pathological function of Aβ1-42. To create such system, (1) a novel propeptide possessing a high solubility during HPLC purification and long-term storage as a solution and (2) a capability of intact Aβ1-42 production under physiological conditions are required.

O–N intramolecular acyl migration reaction is a well-known reaction observed in Ser/Thr-containing peptides.[8] In our previous study regarding the synthesis of difficult sequence-containing peptides, we disclosed a novel and efficient method based on this migration reaction of synthesized '*O*-acyl isopeptides'. This method remarkably improved the synthetic yields of difficult sequence-containing pentapeptides.[9] The result also indicated that the branched ester structure in *O*-acyl isopeptides could suppress the unfavorable nature seen in the difficult sequence-containing peptides. Namely, the insertion of the ester bond into peptide chain can

Keywords: Water-soluble Aβ1-42 isopeptide; Alzheimer's disease; Aβ1-42; *O–N* intramolecular acyl migration reaction; Difficult sequence; 26-*O*-acyl isoAβ1-42.
* Corresponding authors. Tel.: +81-75-595-4635; fax: +81-75-591-9900; e-mail: kiso@mb.kyoto-phu.ac.jp

0040-4039/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2004.06.059

probably disrupt the secondary structure formed by the inherent peptide chain, leading to the improvement of coupling and deblocking efficacy during SPPS. In addition, *O*-acyl isopeptides with the newly formed amino group can possess reasonable H_2O- and MeOH-solubility required in HPLC purification by the formation of salt. Furthermore, from recent research of water-soluble prodrugs[10] and *O*-acyl isopeptides[9] based on *O–N* intramolecular acyl migration, it is established that the purified *O*-acyl isoform can completely be converted to the original *N*-acyl form in a short time with no side reaction at pH 7.4.

Based on this background, we conceived the idea that the *O–N* intramolecular acyl migration of *O*-acyl isopeptides could be applied to the synthesis of Aβ1-42 via a novel water-soluble isopeptide of Aβ1-42, that is, '26-*O*-acyl isoAβ1-42 (26-AIAβ42, **2**)'. This idea would overcome the problems in the synthesis and storage of Aβ1-42 (Fig. 1). Although there are two Ser residues in Aβ1-42 at the positions 8 and 26 with the capability of *O–N* intramolecular acyl migration, we selected the Ser^{26} for the *O*-acylation, since the adjacent Gly^{25} does not epimerize during ester bond formation (Fig. 1).

As depicted in Scheme 1, Fmoc-Ala-O-chlorotrityl resin (0.465 mmol/g, **3**) was employed and Fmoc-protected amino acids were sequentially coupled using the DIPCDI–HOBt method (2 h) after removal of each Fmoc group with 20% piperidine–DMF (20 min) to give peptide resin **4**. After Boc-Ser-OH was introduced to **4**, the obtained **5** was coupled with Fmoc-Gly-OH at the β-hydroxy group of Ser using the DIPCDI–DMAP method in CH_2Cl_2 to obtain ester **6**. 26-AIAβ1-42-resin (**7**) was obtained through the coupling of additional amino acid residues by the conventional manner. Finally, 26-AIAβ1-42 (**2**) was obtained as a major product by the treatment of TFA–*m*-cresol–thioanisole–H_2O (92.5:2.5:2.5:2.5) for 90 min followed by NH_3I–dimethylsulfide for 60 min in TFA–H_2O (2:1).

In HPLC analysis of crude products (Fig. 2A), Aβ1–25 (DAEFRHDSGYEVHHQKLVFFAEDVG) was not observed as a by-product, although a very low rate (1.6%, HPLC yield) of Aβ26-42 (SNKGAIIGLMVGGVVIA) was detected. This indicates that (1) the esterification of the β-hydroxy group of Ser was successfully completed on the solid support and (2) the formed ester bond between Gly and Ser was stable in both piperidine and

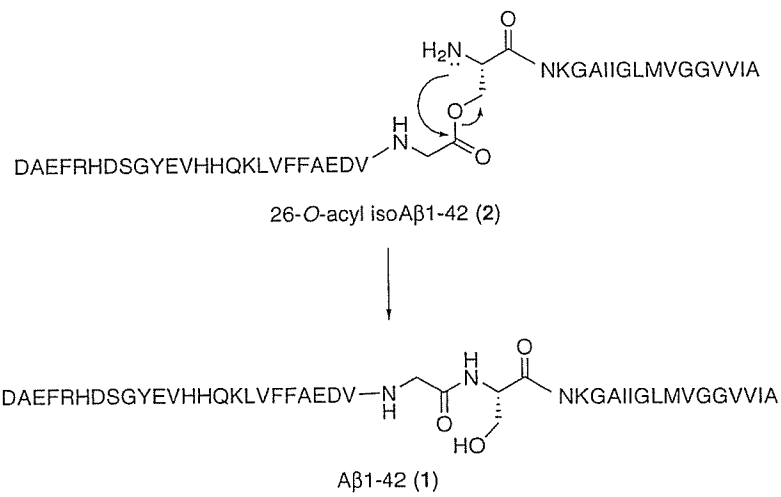
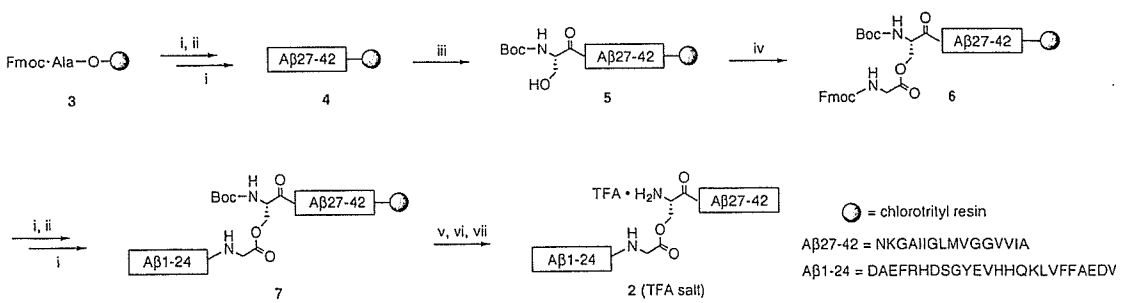


Figure 1. The production of Aβ1-42 (**1**) via the *O–N* intramolecular acyl migration reaction of 26-*O*-acyl isoAβ1-42 (**2**).



Scheme 1. Reagents and conditions: (i) 20% piperidine/DMF, 20 min; (ii) Fmoc-AA-OH (2.5 equiv), DIPCDI (1,3-diisopropylcarbodiimide 2.5 equiv), HOBt (2.5 equiv), DMF, 2 h; (iii) Boc-Ser-OH (2.5 equiv), DIPCDI (2.5 equiv), HOBt (2.5 equiv), DMF, 2 h; (iv) Fmoc-Gly-OH (3.0 equiv), DIPCDI (3.0 equiv), DMAP (0.2 equiv), CH_2Cl_2, 16 h×2; (v) TFA–*m*-cresol–thioanisole–H_2O (92.5:2.5:2.5:2.5), 90 min; (vi) NH_3 (20 equiv), dimethylsulfide (20 equiv), TFA–H_2O (2:1), 60 min, 0 °C; (vii) preparative HPLC (a linear gradient of CH_3CN in 0.1% aqueous TFA)

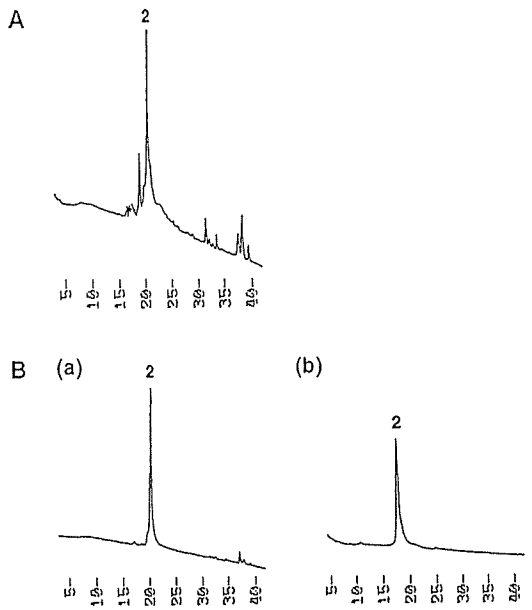


Figure 2. HPLC profiles of (A) crude and (B) purified 26-AIAβ1-42 (2). Analytical HPLC was performed using a C18 reverse phase column (4.6×150 mm; YMC Pack ODS AM302) with binary solvent system: a linear gradient of CH_3CN (0–100% CH_3CN, 40 min for A and B-a, 25–55% CH_3CN, 60 min for B-b) in 0.1% aqueous TFA at a flow rate of 0.9 mL min^{-1} (temperature: 40 °C), detected at 230 nm.

TFA. The crude *O*-acyl isopeptide **2** was dissolved in hexafluoroisopropanol, applied to preparative HPLC, and eluted using 0.1% aqueous TFA–CH_3CN. Since **2** was eluted as a narrow single peak, we could easily purified by preparative scale HPLC to give pure **2** (Fig. 2B) as TFA salt with the total isolated yield of 33.6%, calculated from the original loading of chlorotrityl resin.[11] This yield was higher than that obtained in the synthesis of **1** by a standard Fmoc-based SPPS (7.2%). Since **1** was eluted as a broad peak in preparative scale HPLC purification, it was laborious to isolate **1** from impurities as reported.[5] In addition, in the synthesis of **2**, no conversion to **1** was observed.

The water-solubility of **2** (TFA salt) was 15 mg mL^{-1}, which was 100-fold higher than that of Aβ1-42 (**1**, 0.14 mg mL^{-1}). Interestingly, the fact that a slight modification of peptide chain by the insertion of one ester bond could drastically increase the solubility of the insoluble original peptide with 42-residues suggests that *O*-acyl isopeptide totally break the secondary structures responsible for the insolubility of the original peptide. As we demonstrated that *O*-acyl isopeptide could suppress the unfavorable nature of difficult sequence-containing pentapeptides in the previous study,[9] the present result in the synthesis of **2** indicates that this method is a powerful strategy for increasing the solubility even in large peptides.

As shown in Figure 3, purified **2** was completely converted to Aβ1-42 (**1**) at room temperature in phosphate buffered saline (PBS, pH 7.4) with no side reaction. This migration was rapid with a half-life of 2.6 min, while TFA salt of **2** was stable at 4 °C for at least 30 days as either a solid state or a DMSO solution. Moreover, a slower migration was observed at pH 4.9 (PBS) with a

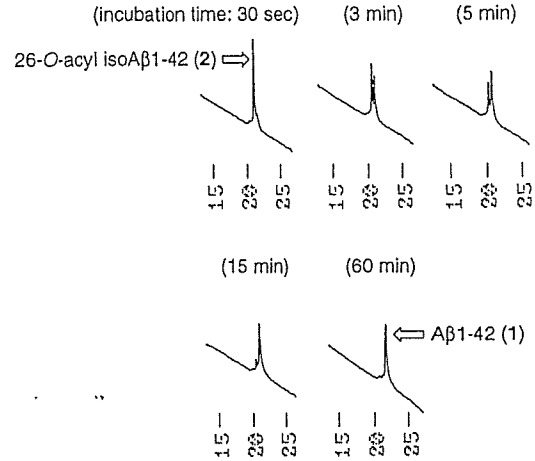


Figure 3. HPLC profiles of the conversion of 26-AIAβ1-42 (**2**) to Aβ1-42 (**1**) via *O–N* intramolecular acyl migration in PBS (pH 7.4, 25 °C). Analytical HPLC was performed using a C18 reverse phase column (4.6×150 mm; YMC Pack ODS AM302) with binary solvent system: a linear gradient of CH_3CN (0–100% CH_3CN for 40 min) in 0.1% aqueous TFA at a flow rate of 0.9 mL min^{-1} (temperature: 40 °C), detected at 230 nm.

half-life of 3 h, and no migration at pH 3.5 (acetate buffer) after incubation for 3 h. This rapid migration under physiological conditions enables to produce intact monomer Aβ1-42 in situ to investigate inherent biological function of Aβ1-42 in AD. The conversion of **2** (TFA salt) in water for 48 h at room temperature followed by lyophilization yielded Aβ1-42 (**1**) quantitatively as TFA salt with the purity of >95%.[12]

In conclusion, we synthesized a novel water-soluble isopeptide of Aβ1-42, (26-AIAβ42, **2**), which is a 26-*O*-acyl isopeptide of Aβ1-42. This *O*-acyl isopeptide has a higher water-solubility than that of Aβ1-42 (**1**), and can migrate to intact Aβ1-42 (**1**) under various conditions while it is stable under storage conditions. This suggests that the synthesis of Aβ1-42 via 26-AIAβ42 could overcome the solubility problem and give a novel tool for the biological evaluation system, in which 26-AIAβ42 can be stored in a solubilized state before the use and rapidly produce the intact Aβ1-42 in situ during biological experiments.

Acknowledgements

This research was supported in part by the Frontier Research Program and grants from the Ministry of Education, Science and Culture of Japan. Y.S. is grateful for Research Fellowships of JSPS for Young Scientists. We thank Dr. Z. Ziora and Dr. M. Skwarczynski for useful discussions, and Mr. A. Taniguchi, Ms. F. Fukao, Ms. Y. Fukusako, Ms. M. Kimura, and Mr. Y. Chiyomori for technical assistance. We are grateful to Mr. T. Hamada for measurement of mass spectra.

References and notes

1. Selkoe, D. J. *Nature* 1999, *399*, A23–31.
2. Geula, C.; Wu, C. K.; Saroff, D.; Lorenzo, A.; Yuan, M. L.; Yankner, B. A. *Nat. Med.* 1998, *4*, 827–831.

3. Sinha, S.; Lieberburg, I. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11049–11053.
4. For examples: (a) Younkin, S. G. *Ann. Neurol.* **1995**, *37*, 287–288; (b) Dahlgren, K. N.; Manelli, A. M.; Stine, W. B., Jr.; Baker, L. K.; Krafft, G. A.; LaDu, M. J. *J. Biol. Chem.* **2002**, *277*, 32046–32053; (c) Bitan, G.; Tarus, B.; Vollers, S. S.; Lashuel, H. A.; Condron, M. M.; Straub, J. E.; Teplow, D. B. *J. Am. Chem. Soc.* **2003**, *125*, 15359–15365; (d) Hou, L.; Shao, H.; Zhang, Y.; Li, H.; Menon, N. K.; Neuhaus, E. B.; Brewer, J. M.; Byeon, I.-J. L.; Ray, D. G.; Vitek, M. P.; Iwashita, T.; Makula, R. A.; Przybyla, A. B.; Zagorski, M. G. *J. Am. Chem. Soc.* **2004**, *126*, 1992–2005.
5. For examples: (a) Burdick, D.; Soreghan, B.; Kwon, M.; Kosmoski, J.; Knauer, M.; Henschen, A.; Yates, J.; Cotman, C.; Glabe, C. *J. Biol. Chem.* **1992**, *267*, 546–554; (b) Quibell, M.; Turnell, W. G.; Johnson, T. *J. Chem. Soc., Perkin Trans. 1* **1995**, *6*, 2019–2024; (c) Shen, C.-L.; Murphy, R. M. *Biophys. J.* **1995**, *69*, 640–651; (d) Fukuda, H.; Shimizu, T.; Nakajima, M.; Mori, H.; Shirasawa, T. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 953–956; (e) Milton, S. C. F.; Milton, R. C. D.; Kates, S. A.; Glabe, C. *Lett. Peptide Sci.* **1999**, *6*, 151–156; (f) Murakami, K.; Irie, K.; Morimoto, A.; Ohigashi, H.; Shindo, M.; Nagano, M.; Shimizu, T.; Shirasawa, T. *Biochem. Biophys. Res. Commun.* **2002**, *294*, 5–10.
6. For a review, see: Sheppard, R. *J. Peptide Sci.* **2003**, *9*, 545–552.
7. Stine, W. B., Jr.; Dahlgren, K. N.; Krafft, G. A.; Ladu, M. J. *J. Biol. Chem.* **2003**, *278*, 11612–11622.
8. (a) Stewart, J. M. In *The Peptides*; Gross, E., Meienhofer, J., Eds.; Academic: NY, 1981; Vol. 3, pp 169–201; (b) Tamamura, H.; Kato, T.; Otaka, A.; Fujii, N. *Org Biomol. Chem.* **2003**, *1*, 2468–2473; (c) Mouls, L.; Subra G.; Enjalbal, C.; Martinez, J.; Aubagnac, J.-L. *Terahedro Lett.* **2004**, *45*, 1173–1178.
9. Sohma, Y.; Sasaki, M.; Hayashi, Y.; Kimura, T.; Kiso, Y *Chem. Commun.* **2004**, 124–125.
10. (a) Kimura, T.; Ohtake, J.; Nakata, S.; Enomoto, H. Moriwaki, H.; Akaji, K.; Kiso, Y. Peptide Chemistr 1994. Ohno, M., Ed.; Protein Rearch Foundation: Osaka 1995; pp 157–160; (b) Oliyai, R.; Stella, V. J. *Bioorg. Med Chem. Lett.* **1995**, *5*, 2735–2740; (c) Kiso Y.; Kimura, T.; Ohtake, J.; Nakata, S.; Enomoto, H. Moriwaki, H.; Nakatani, M.; Akaji, K. *Peptides: Chem istry, Structure and Biology*; Mayflower Scientific: Eng land, 1996; pp 157–159; (d) Hamada, Y.; Ohtake J.; Sohma, Y.; Kimura, T.; Hayashi, Y.; Kiso, Y. *Bioorg Med. Chem.* **2002**, *10*, 4155–4167; (e) Hayashi, Y. Skwarczynski, M.; Hamada, Y.; Sohma, Y.; Kimura T.; Kiso, Y. *J. Med. Chem.* **2003**, *46*, 3782–3784 (f) Skwarczynski, M.; Sohma, Y.; Kimura, M.; Hayashi Y.; Kimura, T.; Kiso, Y. *Bioorg. Med. Chem. Lett.* **2003** *13*, 4441–4444.
11. Spectral data for **2**: MALDI-MS (TOF): M_{calc}: 4514.04 $M + H_{found}$: 4515.26; HPLC analysis at 230 nm: purity wa >96%.
12. Spectral data for **1**: MALDI-MS (TOF): M_{calc}: 4514.04 $M + H_{found}$: 4515.48; HPLC analysis at 230 nm: purity wa >95%; the retention time on HPLC (0–100% CH_3CN fo 40 min, 230 nm) of synthesized **1** was identical to that c commercially available Aβ1-42 (Peptide Institute, Inc Osaka, Japan).

Acta Crystallographica Section B
Structural
Science
ISSN 0108-7681

Rigid backbone moiety of KNI-272, a highly selective HIV protease inhibitor: methanol, acetone and dimethylsulfoxide solvated forms of 3-[3-benzyl-2-hydroxy-9-(isoquinolin-5-yloxy)-6-methylsulfanylmethyl-5,8-dioxo-4,7-diazanonanoyl]-*N*-*tert*-butyl-1,3-thiazolidine-4-carboxamide

Mitsunobu Doi,[a]* Tooru Kimura,[b] Toshimasa Ishida[a] and Yoshiaki Kiso[b]

[a]Osaka University of Pharmaceutical Sciences, Nasahara, Takatsuki, Osaka 569-104, Japan, and [b]Department of Medicinal Chemistry, Kyoto Pharmaceutical University, Yamashinaku, Kyoto 607-8412, Japan

Correspondence e-mail: doit@gly.oups.ac.jp

Received 16 March 2004
Accepted 29 May 2004

When crystals of kynostatin (KNI)-272, a highly selective HIV protease inhibitor containing allophenylnorstatine [(2*S*,3*S*)-3-amino-2-hydroxy-4-phenylbutyric acid], were grown in three different solvent systems (methanol, acetone and dimethylsulfoxide solutions), the local conformations around the hydroxymethylcarbonyl (HMC) moiety, which mimics the structure of the transition state, were similar in all three forms. The peptide backbones were slightly bent, but their structures differed from typical sheets, turns or helixes. Although the isoquinoline ring at the *N*-terminal showed conformational variations, a remarkable similarity was observed in the *C*-terminal region, including the HMC moiety. Moreover, the conformational characteristics of the uncomplexed forms resembled those of the inhibitor within the KNI-272–HIV protease complex. This suggests that the structure of the *C*-terminal region of KNI-272 is rigid or very stable.

1. Introduction

KNI-272 (I) is a tripeptide mimic whose design was based on the substrate transition state concept of amide hydrolysis by aspartic protease (Fig. 1; Kiso, 1996; Kiso *et al.*, 1999). This peptide potently inhibits HIV protease activity (Mimoto *et al.*, 1992) and displays good pharmacokinetics and an excellent therapeutic index (Kageyama *et al.*, 1993). It is the unusual amino acid allophenylnorstatine [Apns = (2*S*,3*S*)-3-amino-2-hydroxy-4-phenylbutyric acid] which has the key structure that mimics the transition state (Mimoto, Imai, Tanaka, Hattori, Takahashi *et al.*, 1991; Mimoto, Imai, Tanaka, Hattori, Kisanuki *et al.*, 1991), although the other unusual amino acids of KNI-272, 5-isoquinolyloxy acetic acid (iQoa), methylthioalanine (Mta), thiazolidine-4-carboxylic acid (Thz) and *tert*-butylamine (^{t}Bu), are also highly optimized to enhance the inhibitory activity and selectivity for HIV protease (Sheha *et al.*, 2000). The P1—P1′ site of aspartic protease (Schechter, 1968) is formed at the hydroxymethylcarbonyl (HMC)–amide bond, between Apns[3] and Thz[4], and both *cis* and *trans* conformers can be presumed for this bond. However, NMR and molecular modeling studies indicate that the *trans* conformer of the HMC–amide bond should predominate (Kato *et al.*, 1994; Ohno *et al.*, 1996), which is consistent with the crystal structure (Doi *et al.*, 2001), the structure of the inhibitor within the KNI-272–HIV protease complex (Baldwin *et al.*, 1995) and the finding that a wide moiety which includes

DOI: 10.1107/S0108768104013047

an HMC group will tend to converge towards a certain conformation (Ohno *et al.*, 1996). Although no strong factor limiting the peptide conformation has been found within the chemical structure of KNI-272, the above results suggest the existence of a stable conformation for the moiety with the HMC group. Since the properties of solvents affect the peptide structures (Karle *et al.*, 1990; Awasthi *et al.*, 2001), the inherent conformational variations of peptides seem to be observed in the crystal structures grown from different solvent systems. To test that hypothesis, the comparison among such crystal structures could be a probe. Therefore, we solved three solvated crystal forms of KNI-272 and compared the stable conformations with previously solved structures.

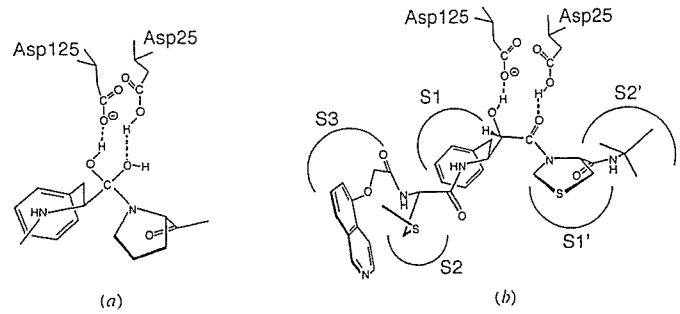
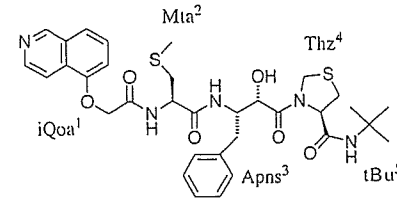


Figure 1
The substrate transition state concept (Kiso, 1996; Kiso *et al.*, 1999). (*a*) Substrate transition state of amide hydrolysis by HIV protease. A water molecule is added to the carbonyl carbon of a substrate activated by catalytic Asp residues of HIV protease. An intermediate takes a tetrahedral coordination for a brief time and then the C—N bond is broken. (*b*) Interaction between KNI-272 and HIV protease. Hydrogen bonds with catalytic Asp residues resemble those of the substrate transition state, but in this case the C—N bond is never broken.



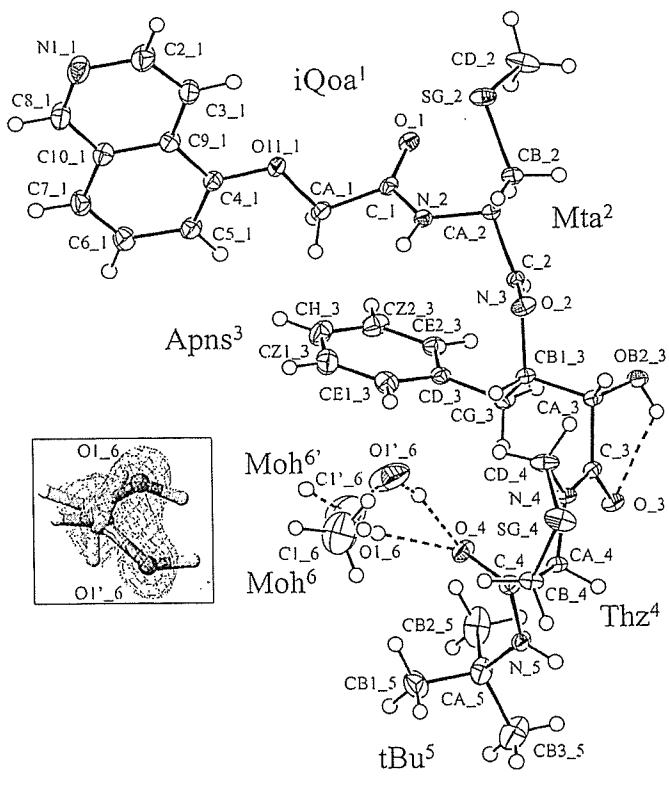


Figure 2
Structure of the ME form with displacement ellipsoids drawn at the 50% probability level. Dotted lines represent hydrogen bonds to a disordered methanol molecule. The inset shows the electron density around the solvents at the 0.8 and 2.4 σ level.

2. Experimental

KNI-272, which was synthesized as described previously (Mimoto *et al.*, 1992), was dissolved in approximately 60–80% aqueous methanol with warming and the minimum amount of methanol needed to prevent precipitation. The solution was then sealed in a vial at room temperature and crystals of the aqueous methanol form (ME form) were grown for 3–4 d. Crystals of the acetone solvated form (AC form) were grown from KNI-272 acetone solution containing 5–10% hexane. Crystals of the dimethylsulfoxide solvated form (DM form) were grown from KNI-272 dissolved in dimethylsulfoxide, to which was added a small amount of water (5–10%). In this case, the solution was sealed in a vial overnight. All crystals were briefly passed through 100% glycerol on a nylon loop (Hampton Inc.) and then flash-frozen under a nitrogen stream at 90 K. Data collection was carried out on a Bruker SMART APEX CCD area detector using Mo $K\alpha$ radiation. Intensities were integrated using *SMART* and *SAINT* (Bruker, 1998), and the empirical absorption corrections were applied using *SADABS* (Sheldrick, 1996). The structures were solved using *SHELXS97* (Sheldrick, 1997) and refined with *SHELXL97* (Sheldrick, 1997). H atoms were placed at calculated positions (C—H 0.95–1.00 and N—H 0.88 Å) using isotropic displacement parameters [U_{iso} = 1.5 U_{eq} (C) for methyl H atoms and 1.2 U_{eq} (parent atom) for all other atoms] and were included in the structure-factor calculations. The H atom in the hydroxyl group of Apns^3 was found in a differential Fourier map and was fixed during the refinements. In the ME form, the hydroxyl H atoms of the disordered methanol molecules were also found in a differential Fourier map and they too were fixed during the refinements. The absolute configurations were determined from the material amino acids and were consistent with the suggested Flack parameters (Flack, 1983). The crystal and experimental data are summarized in Table 1.[1]

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: OG5001). Services for accessing these data are described at the back of the journal.

Table 1
Crystal and experimental summaries of the ME and AC forms and referenced structures of the MPD form.

MPD represents KNI-272 crystals grown from aqueous 2-methylpentane-2,4-diol solution (Doi *et al.*, 2001).

	ME	AC	DM
Formula	$C_{33}H_{41}N_5O_6S_2$	$C_{33}H_{41}N_5O_6S_2$	$C_{33}H_{41}N_5O_6S_2$
Solvation	CH_4O	$2C_3H_6O$	C_2H_6SO
M_r	699.87	783.98	745.96
System	Monoclinic	Monoclinic	Monoclinic
Space group	$P2_1$	$P2_1$	$C2$
a (Å)	10.5077 (9)	10.6395 (7)	29.390 (6)
b (Å)	13.288 (1)	13.2130 (9)	12.882 (3)
c (Å)	13.515 (1)	14.692 (1)	10.609 (2)
β (°)	101.924 (1)	98.931 (1)	103.373 (3)
V (Å^3)	1846.3 (3)	2040.4 (2)	3907.6 (13)
Z	2	2	4
T (K)	90.0 (2)	90.0 (2)	90.0 (2)
D_x (g cm^{-3})	1.259	1.276	1.268
F(000)	744	836	1584
Wavelength (Å)	0.7107	0.7107	0.7107
μ (mm^{-1})	0.196	0.187	0.241
No. of reflections (obs)	16 068	23 703	12 220
R_{int}	0.0191	0.0161	0.0344
No. of reflections (used)	8028	8940	7717
θ_{max} (°)	27.1	27.1	26.7
Flack parameter	−0.01 (4)	0.01 (3)	0.09 (6)
R	0.0333	0.0286	0.0443
wR	0.0835	0.0729	0.1136
Goodness-of-fit	1.026	1.034	0.992
$(\Delta/\sigma)_{max}$	0.011	0.007	0.009
Fraction of θ_{max}	0.998	0.997	0.988
$\Delta\rho_{max}$ (e Å^{-3})	0.338	0.285	0.835
$\Delta\rho_{min}$ (e Å^{-3})	−0.301	−0.218	−0.475

3. Results and discussion

The structures of the ME, AC and DM forms are shown in Figs. 2, 3 and 4, respectively. In the ME form, electron densities persist in the solvent region (Fig. 2, inset) and are interpreted as being a methanol molecule disordered to two sites (Moh[6] and Moh[6']). Both the disordered methanol molecules are located at the positions close to the thiazolidine ring (Thz[4]) and hydrogen-bonded to the carbonyl oxygen of Thz[4]: O1_6···O_4 2.842 (4) and O1'_6···O_4 2.706 (4) Å. The AC form contains two acetone molecules (Ace[6] and Ace[7]). Ace[6] interacts with the amide nitrogen of Mta[2] [N_2···O1_6 3.019 (2) Å], but no hydrogen bond is formed with Ace[7] (Fig. 3). Ace[6] is sandwiched between the Thz[4] and the isoquinoline (iQoa[1]) rings, and the hydrophobic interactions among these residues seem to contribute to the peptide conformation of the AC form. A dimethylsulfoxide molecule (Dms[6]) is present within the DM form (Fig. 4) and is hydrogen-bonded to the amide nitrogen of Mta[2]: N_2···O1_6 2.832 (3) Å. This interaction pattern is similar to that of Ace[6] in the AC form. The relative disposition of Dms[6] for the peptide is also similar to that of Ace[6], but the isoquinoline ring is not folded to Dms[6]. Apparently KNI-272 is 'hygroscopic' for the solvents used, so that its crystallization is accompanied by solvation and the properties of solvents affect the conformations of KNI-272.

Within the substrate–HIV protease complex, at least one water molecule is used for the hydrolysis of the amide bond and should therefore be located in close proximity to the substrate (Kiso, 1996). NMR analysis has shown that water molecules bind between KNI-272 and HIV protease, stabilizing the complex structure (Wang, Freedberg, Wingfield *et al.*, 1996). It thus appears that the hygroscopicity of KNI-272 supports the structure mimicking the substrate transition state within the complex.

According to the substrate transition state concept (Kiso, 1996; Kiso *et al.*, 1999), the HMC group of the Apns[3] residue should assume a conformation that mimics the substrate of HIV protease (Fig. 1), which is a homodimeric protein with a symmetric catalytic center (Pearl & Taylor, 1987). It is known, moreover, that when the asymmetric KNI-272 is bound, catalytic Asp residues assume different ionization states that contribute to a structure that mimics the transition state (Wang, Freedberg, Yamazaki *et al.*, 1996). This means that the structure around the HMC group is the most important when considering the interaction between KNI-272 and the catalytic center of the protease. The bond distances around the HMC moiety of the Apns[3] residue clearly distinguish the C*A*_3—O*B*2_3 hydroxyl bond from the C_3—O_3 carbonyl bond (Table 2), and no trend toward tautomerism between the hydroxyl and carbonyl groups is seen in any of the crystal forms. Such clear localization of the hydroxyl and carbonyl groups would provide support for the structure mimicking the transition state within the complex. Bond rotation is chemically permitted for C*A*_3—C_3, but the conformations of the hydroxyl and carbonyl groups are synperiplanar in the three forms: O*B*2_3—C*A*_3—C_3—O_3 −20.9 (2), −24.1 (2) and −24.6 (4)° for the ME, AC and MD forms, respectively. A hydrogen bond is formed between O*B*2_3 and O_3, making a five-membered ring at the HMC group, which contributes to the stability of the synperiplanar conformations.

Similar geometry was previously found around the HMC group of KNI-272 crystals grown from aqueous 2-methylpentane-2,4-diol solution (MPD form; Doi *et al.*, 2001). It is noteworthy that geometries found at the HMC structure are similar despite the different physical properties of the respective solvent systems (*e.g.* polarity and dielectric constant). Within the structure of the KNI-272–HIV protease complex (the 1HPX form; Baldwin *et al.*, 1995), the HMC group interacts with two Asp residues and there are many contacts between KNI-272 and the protein. Nevertheless, the geometry of HMC in the 1HPX form is surprisingly similar to those of the uncomplexed structures, which implies that the local conformation of the HMC moiety in single crystals reflects the active conformation of the inhibitor–enzyme complex.

The peptide backbones of all the forms are slightly bent but cannot be classified as a typical sheet, turn or helix (Table 3). For comparison, the ME, AC, DM and MPD forms of KNI-272 were fitted to the 1HPX form using the four Cα atoms of residues 2–5 (Fig. 5). Conformational variation is observed in the isoquinoline ring of the iQoa[1] residue, which would reflect the molecule–molecule contacts within the crystal and solvent environments. Similar variation is observed for the side chains of Mta[2] and Apns[3], but their spatial distributions are significantly smaller than that of iQoa[1]. This suggests that the

Table 2
Selected geometry around the HMC moiety of the Apns residue.

MPD represents KNI-272 crystals grown from aqueous 2-methylpentane-2,4-diol solution (Doi *et al.*, 2001). 1HPX represents the structure of the inhibitor within the HIV protease–KNI-272 complex (Baldwin *et al.*, 1995).

	ME	AC	DM	MPD	1HPX
CB1_3–CA_3	1.548 (2)	1.548 (2)	1.545 (4)	1.529	1.53
CA_3–OB2_3	1.409 (2)	1.407 (2)	1.409 (3)	1.412	1.38
OB2_3–H	0.855	0.874	0.848	0.757	0.97
CA_3–C_3	1.533 (2)	1.525 (2)	1.528 (4)	1.534	1.52
C_3–O_3	1.236 (2)	1.228 (2)	1.233 (3)	1.240	1.23
N_3–*CB1_3*–*CA_3*–*OB2_3*	−65.2 (2)	−67.2 (1)	−72.1 (3)	−65.1	−75
CB1_3–CA_3–C_3–O_3	101.9 (2)	98.5 (1)	97.5 (3)	107.9	89
OB2_3–CA_3–C_3–O_3	−20.9 (2)	−24.1 (2)	−24.6 (4)	−14.2	−33
OB2_3···O_3	2.693 (2)	2.704 (1)	2.693 (3)	2.669	2.68
OB2_3–H···O_3	2.187	2.359	2.200	2.269	2.22
∠*OB2_3*–H···O_3	117.7	103.8	117.0	114.2	108

conformation of iQoa[1] is flexible and reflects the specific environment of the KNI-272 molecule, and that this flexibility is favorable for stabilizing the inhibitor molecule on the enzyme. In the 1HPX form, the isoquinoline ring is located on the protein surface, bending into the cavity of the active center, which would delay the release of the inhibitor from the protease. By contrast, the backbones of the Mta[2]–Apns[3]–Thz[4]–'Bu[5] moieties are very similar to one another (Fig. 5), with torsion angles ($\varphi3'$, $\psi3$, $\varphi4$, $\psi4$ and $\varphi5$) that are indicative of the similarity of the backbone conformations (Table 3). In the MPD form, the hydrating waters (disordered to three sites) are located beside the phenyl ring of Apns[3], forcing the side chain to swing, thereby causing slight shifts in the CB1_3 position related to the $\varphi3$ angle. In addition, the Thz[4] ring of the 1HPX form shows a 'puckering' that differs slightly from the other forms (*e.g.* the $\chi4$ angle). Water molecules 301 and 607 are located beside Thz[4] within the complex and seem to affect the 'puckering' of the Thz[4] ring. Analysis of molecular dynamics using NOE restraints has shown there to be three predominant families of KNI-272 structures (conformations *A*, *B* and *C*) with conformational variations at iQoa[1] (Ohno *et al.*, 1996). Of those, conformation *A* was similar to the 1HPX form. In particular, the backbone of Apns[3]–Thz[4]–'Bu[5] of conformation *A* showed remarkable similarity to that of the 1HPX form. These similarities suggest the existence of a stable conformation of the backbone of the Mta[2]–Apns[3]–Thz[4]–'Bu[5] moiety. In other words, this moiety is conformationally restricted and reflects the active form in the complex.

The findings summarized above suggest that KNI-272 is pre-organized to the active form so that, except for the iQoa[1]

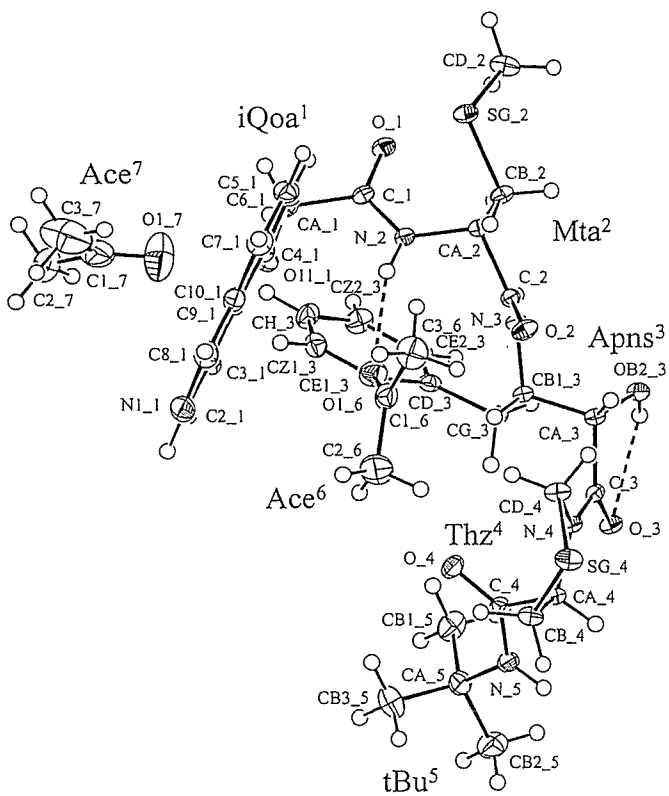


Figure 3
Structure of the AC form with displacement ellipsoids drawn at the 50% probability level. Dotted lines represent hydrogen bonds.

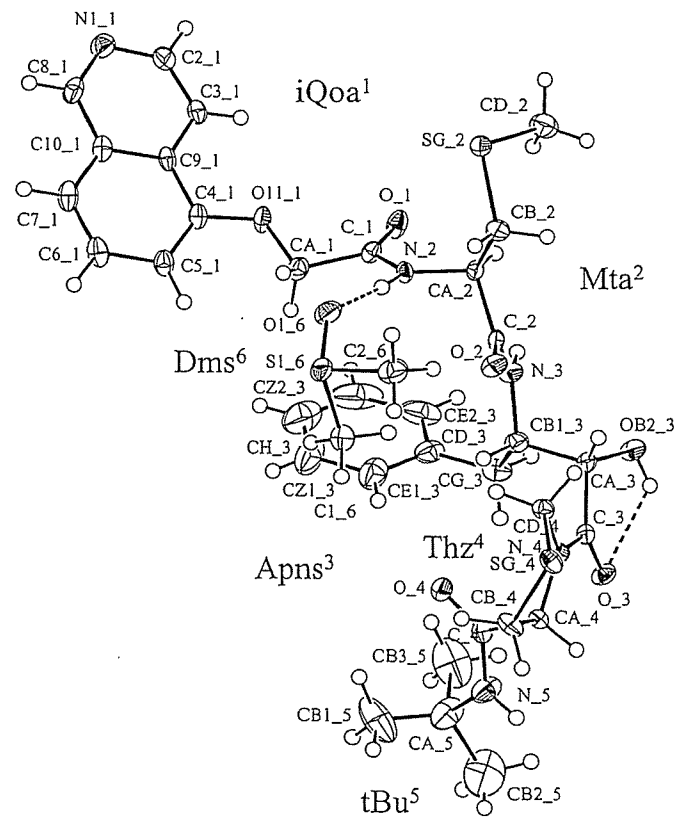


Figure 4
Structure of the DM form with displacement ellipsoids drawn at the 50% probability level. Dotted lines represent hydrogen bonds.

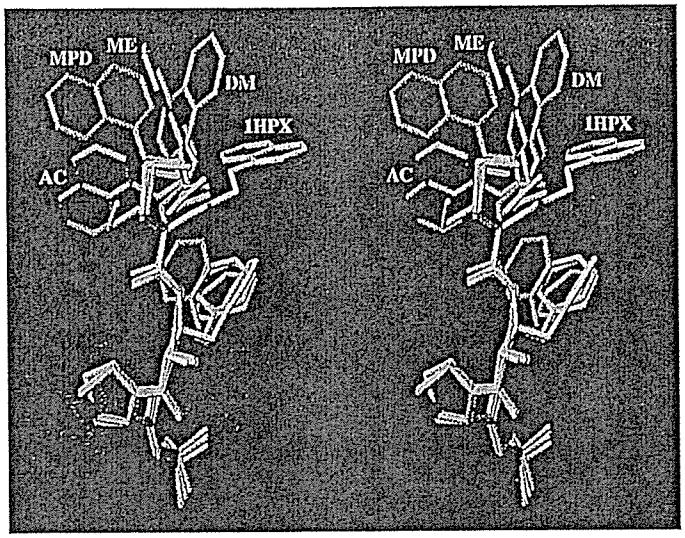


Figure 5
Stereoview of the superimposition for four KNI-272 molecules. Molecular fitting was carried out for the four Cα atoms of residues 2–5 with an r.m.s.d. of 0.03–0.08 Å. H atoms are omitted for clarity. The colors cyan, yellow, green, orange and gray represent the ME, AC, DM, MPD and 1HPX forms, respectively. The side chain of $Apns^3$ was disordered to two sites in the MPD form, but only one part is drawn for clarity. This figure is produced using *TRUBO-FRODO* (Roussel *et al.*, 2002).

residue, effective and selective binding to HIV protease is achieved without any large conformational changes. However, we still do not know what the key factor is. From studies of HIV protease inhibitors (Mimoto, Kato *et al.*, 1999; Mimoto, Hattori *et al.*, 2000), it seems that a certain combination of unusual amino acids, Apns, Thz and ^{t}Bu, mediates the actual restrictions for five rotatable bonds in the regions ($\varphi 3$, $\varphi 3'$, $\psi 3$, $\varphi 4$ and $\psi 4$) by steric hindrance. The basic skeleton of the Mta^2–$Apns^3$–Thz^4–$^t Bu^5$ moiety would therefore be very important for an HIV protease inhibitor having an HMC group. It has also been adapted to the recently developed water-soluble HIV protease inhibitor prodrugs, which also show potent activity (Sohma *et al.*, 2003).

References

Awasthi, S. K., Shankaramma, S. C., Raghothama, S. & Balaram, P. (2001). *Biopolymers*, **15**, 465–476.

Baldwin, E. T., Bhat, T. N., Gulnik, S., Liu, B., Topol, I. A., Kiso, Y., Mimoto, T., Mitsuya, H. & Erickson, J. W. (1995). *Structure*, **3**, 581–590.

Bruker (1998). *SAINTPLUS*, Version 5, and *SMART*, Version 5. Bruker AXS Inc., Madison, Wisconsin, USA.

Doi, M., Ishida, T., Katsuya, Y., Sasaki, M., Taniguchi, T., Hasegawa, H., Mimoto, T. & Kiso, Y. (2001). *Acta Cryst.* C**57**, 1333–1335.

Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.

Kageyama, S., Mimoto, T., Murakawa, Y., Nomizu, M., Ford, Jr, H., Shirasaka, T., Gulnik, S., Erickson, J., Takada, K., Hayashi, H., Broder, S., Kiso, Y. & Mitsuya, H. (1993). *Antimicrob. Agents Chemother.* **37**, 810–817.

Karle, I. L., Flippen-Anderson, J. L., Uma, K., Balaram, H. & Balaram, P. (1990). *Biopolymers*, **29**, 1433–1442.

Kato, R., Takahashi, O., Kiso, Y., Moriguchi, I. & Hirono, S. (1994). *Chem. Pharm. Bull.* **42**, 176–178.

Kiso, Y. (1996). *Biopolymers*, **40**, 235–244.

Kiso, Y., Matsumoto, H., Mizumoto, S., Kimura, T., Fujiwara, Y. & Akaji, K. (1999). *Biopolymers*, **51**, 59–68.

Mimoto, T., Hattori, N., Takaku, H., Kisanuki, S., Fukazawa, T., Terashima, K., Kato, R., Nojima, S., Misawa, S., Ueno, T., Imai, J., Enomoto, H., Tanaka, S., Sakikawa, H., Shintani, M., Hayashi, H. & Kiso, Y. (2000). *Chem. Pharm. Bull.* **48**, 1310–1326.

Mimoto, T., Imai, J., Kisanuki, S., Enomoto, H., Hattori, N., Akaji, K. & Kiso, Y. (1992). *Chem. Pharm. Bull.* **40**, 2251–2553.

Mimoto, T., Imai, J., Tanaka, S., Hattori, N., Kisanuki, S., Akaji, K. & Kiso, Y. (1991). *Chem. Pharm. Bull.* **39**, 3088–3090.

Mimoto, T., Imai, J., Tanaka, S., Hattori, N., Takahashi, O., Kisanuki, S., Nagano, Y., Shintani, M., Hayashi, H., Sakikawa, H., Akaji, K. & Kiso, Y. (1991). *Chem. Pharm. Bull.* **39**, 2465–2467.

Mimoto, T., Kato, R., Takaku, H., Nojima, S., Terashima, K., Misawa, S., Fukazawa, T., Ueno, T., Sato, H., Shintani, M., Kiso, Y. & Hayashi, H. (1999). *J. Med. Chem.* **42**, 1789–1802.

Ohno, Y., Kiso, Y. & Kobayashi, Y. (1996). *Bioorg. Med. Chem.* **4**, 1565–1572.

Pearl, L. H. & Taylor, W. R. (1987). *Nature*, **329**, 351–354.

Schechter, I. B. A. (1968). *Biochem. Biophys. Res. Commun.* **32**, 898–902.

Sheha, M. M., Mahfouz, N. M., Hassan, H. Y., Youssef, A. F., Mimoto, T. & Kiso, Y. (2000). *Eur. J. Med. Chem.* **35**, 887–894.

Sheldrick, G. M. (1996). *SADABS*. University of Göttingen, Germany.

Sheldrick, G. M. (1997). *SHELXS97* and *SHELXL97*. University of Göttingen, Germany.

Sohma, Y., Hayashi, Y., Ito, T., Matsumoto, H., Kimura, T. & Kiso, Y. (2003). *J. Med. Chem.* **46**, 4124–4135.

Roussel, A., Legaigneur, P., Inisan, A. G. & Cambillau, C. (2002). TURBO-*FRODO*, Version Linux.1. Universite Aix-Marseille II, Marseille, France.

Wang, Y. -X., Freedberg, D. I., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y. & Torchia, D. A. (1996). *Biochemistry*, **35**, 9945–9950.

Wang, Y.-X., Freedberg, D. I., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y., Bhat, T. N., Erickson, J. W. & Torchia, D. A. (1996). *J. Am. Chem. Soc.* **118**, 12287–12290.

Youhei Sohma
Yoshio Hayashi
Mariusz Skwarczynski
Yoshio Hamada
Masato Sasaki
Tooru Kimura
Yoshiaki Kiso
Department of Medicinal Chemistry, Center for Frontier Research in Medicinal Science, Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607-8412, Japan

O–N Intramolecular Acyl Migration Reaction in the Development of Prodrugs and the Synthesis of Difficult Sequence-Containing Bioactive Peptides*

Received 18 June 2004; accepted 8 July 2004
Published online 13 September 2004 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20136

Abstract: *N–O intramolecular acyl migration in Ser- or Thr-containing peptides is a well-known side reaction in peptide chemistry. It results in the mutual conversion of ester and amide bonds. Our medicinal chemistry study focused on the fact that the O-acyl product can be readily converted to the original N-acyl form under neutral or slightly basic conditions in an aqueous buffer and the liberated ionized amino group enhances the water solubility of O-acyl products. Because of this, we have developed a novel class of "O–N intramolecular acyl migration"-type water-soluble prodrugs of HIV-1 protease inhibitors. These prodrugs released the parent drugs via a simple chemical mechanism with no side reaction. In this study, we applied this strategy to important cancer chemotherapeutic agents, paclitaxel and its derivatives, to develop water-soluble taxoid prodrugs, and found that these prodrugs, 2'-O-isoform of taxoids, showed promising results with higher water solubility and proper kinetics in their parent drug formation by a simple pH-dependent chemical mechanism with O–N intramolecular acyl migration. These results suggest that this strategy would be useful in toxicology and medical economics.*

After the successful application of O–N intramolecular acyl migration in medicinal chemistry, this concept was recently used in peptide chemistry for the synthesis of "difficult sequence-containing

Correspondence to: Yoshiaki Kiso; email: kiso@mb. kyoto-phu.ac.jp

*This paper is dedicated to the memory of Murray Goodman. We will always remember his contributions to peptide science and to peptide societies all over the world, his enthusiasm for life and science encouraging all of us, and his warm, joyous, and humane personality.

Contract grant sponsor: Frontier Research Program and Ministry of Education, Science and Culture of Japan

Biopolymers (Peptide Science), Vol. 76, 344–356 (2004)

peptides." The strategy was based on hydrophilic O-acyl isopeptide synthesis followed by the O–N intramolecular acyl migration reaction, leading to the desired peptide. In a model study with small, difficult sequence-containing peptides, synthesized "O-acyl isopeptides" not only improved the solubility in various media and efficiently performed the high performance liquid chromatography purification, but also altered the nature of the difficult sequence during SPPS, resulting in the efficient synthesis of O-acyl isopeptides with no complications. The subsequent O–N intramolecular acyl migration of purified O-acyl isopeptides afforded the desired peptides as precipitates with high yield and purity. Further study of the synthesis of a larger difficult sequence-containing peptide, Alzheimer's disease-related peptide (Aβ1–42), surprisingly showed that only one insertion of the O-acyl group drastically improved the unfavorable nature of the difficult sequence in Aβ1–42, and achieved efficient synthesis of 26-O-acyl isoAβ1–42 and subsequent complete conversion to Aβ1–42 via the O–N intramolecular acyl migration reaction of 26-O-acyl isoAβ1–42. This suggests that our new method based on O–N intramolecular acyl migration is an important method for the synthesis of difficult sequence-containing bioactive peptides. © 2004 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 76: 344–356, 2004

Keywords: *O–N intramolecular acyl migration reaction; water-soluble prodrug; paclitaxel; difficult sequence-containing peptide; O-acyl isopeptide; Aβ1–42; 26-O-acyl isoAβ1–42*

INTRODUCTION

N–O intramolecular acyl migration is well known as a side reaction of Ser- or Thr-containing peptides.[1–5] β-Hydroxyl groups are acylated by the N–O shift under strong acidic conditions through a five-membered ring intermediate (Figure 1A). However, the resulting *O*-acyl isopeptides can be readily converted to original *N*-acyl compounds under neutral or slightly basic conditions in aqueous buffer. Therefore, after deprotection of protected Ser/Thr-containing peptides with strong acid, this O–N intramolecular acyl migration reaction is conventionally manipulated to obtain the desired peptides. On the other hand, the solubility of *O*-acyl isopeptides in aqueous media generally increases with the newly produced and ionized amino group, although these isopeptides have had limited attention in peptide synthesis so far. In addition, it was

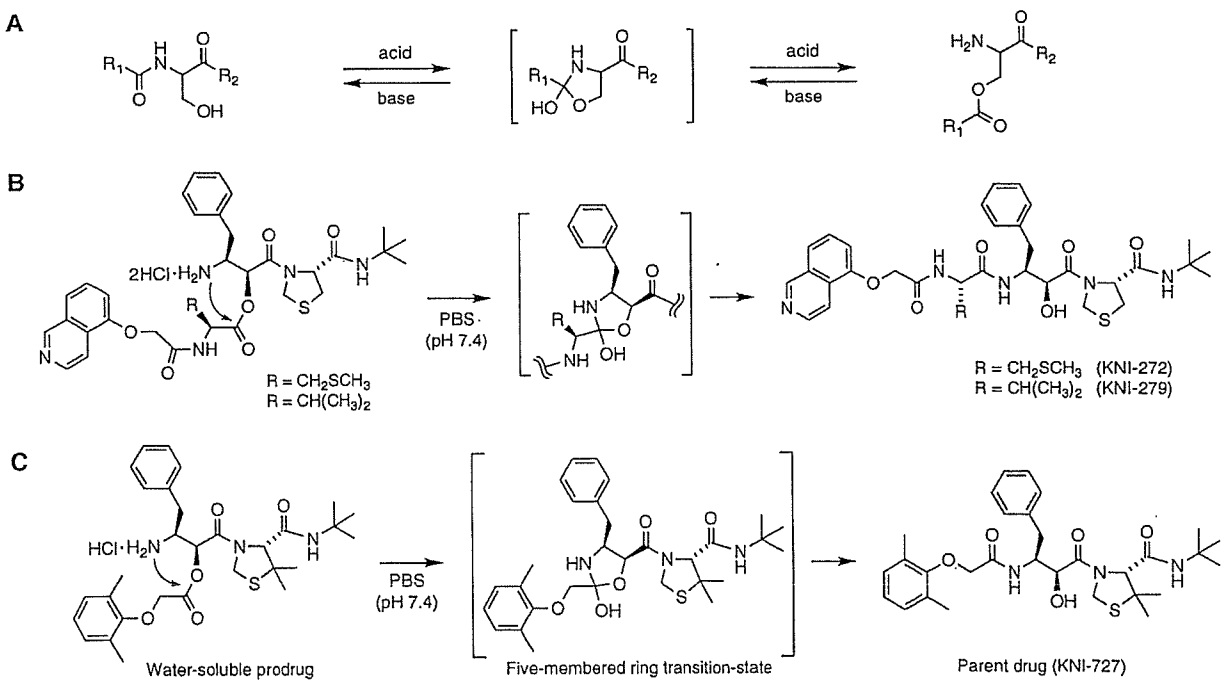


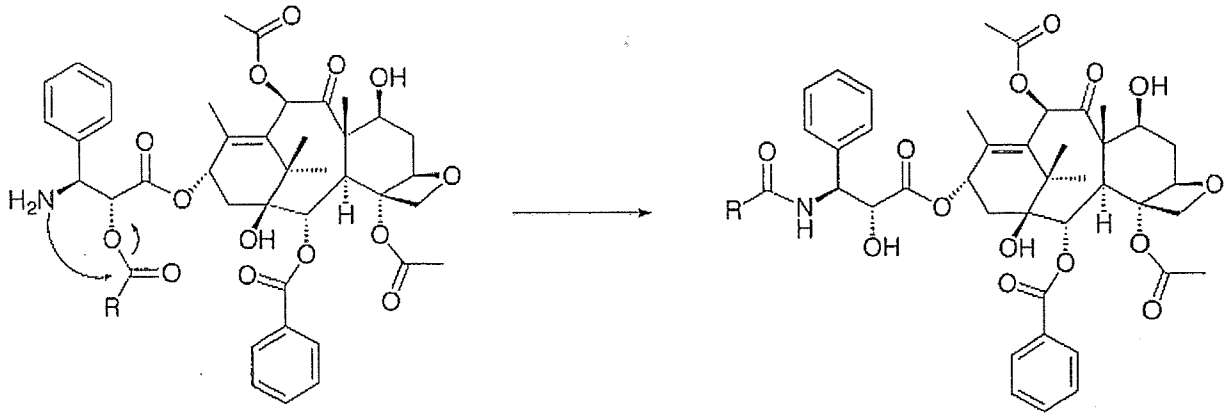
FIGURE 1 (A) Mechanism of N–O and O–N intramolecular acyl migration reaction; (B) O–N intramolecular acyl migration-type water-soluble prodrugs of KNI-272 and -279; (C) water-soluble prodrug of KNI-727.

predicted that the intramolecular acyl migration reaction would also occur in a compound with *N*-acylated α-hydroxy-β-amino acids, which has a similar 2-aminoethanol structure to β-hydroxy-α-amino acids. Utilizing these features, we have developed a novel class of "O–N intramolecular acyl migration"-type water-soluble prodrugs of HIV-1 protease inhibitors such as KNI-272, -279, and -727.[6–11] These inhibitors contain an α-hydroxy-β-amino acid residue, allophenylnorstatine [Apns, (2*S*, 3*S*)-3- amino-2-hydroxy-4-phenylbutanoic acid], which has a hydroxymethylcarbonyl (HMC) isostere derived from a natural scissile amino acid sequence "Phe–Pro."[12–20] As depicted in Figures 1B and C, these "O–N intramolecular acyl migration"-type water-soluble prodrugs of HIV-1 protease inhibitors, which are *O*-acyl isoforms of the parent drugs, increase water solubility and are easily converted to the corresponding *N*-acyl forms, i.e., parent drugs via O–N intramolecular acyl migration. This migration can be precisely controlled by pH and is completed in a short time with no side reactions under physiological conditions (pH 7.4). Hurley et al. also reported a study on the O–N-acyl migration of renin inhibitors.[21] From these findings, a new prodrug concept based on O–N intramolecular acyl migration was established, and the sparingly water-soluble drugs with an *N*-acyl-2-aminoethanol structure are generally potential targets for this water-soluble prodrug strategy.

Paclitaxel (**1**), one of the most important chemotherapeutic agents with promising antitumor activity,[22] has poor water solubility (0.00025 mg/mL). This is one of its major drawbacks, since the need to coinject with a detergent, Cremophor EL®, is thought to cause hypersensitivity reactions.[23] In this study, to solve this problem, we applied the O–N intramolecular acyl migration-type water-soluble prodrug strategy to paclitaxel, focusing on its *N*-benzoylphenylisoserine residue, an α-hydroxy-β-amino acid derivatives. This prodrug, designated "isotaxel," which is a 2′-*O*-isoform of paclitaxel, was synthesized and its potential as a prodrug was evaluated. Furthermore, we synthesized the water-soluble prodrug of other taxoid to understand the generalization of this strategy.

In the development of automated solid-phase peptide synthesis (SPPS), most peptides with ~50 amino acids are routinely prepared with no difficulty. However, the synthesis of specific sequences, so-called difficult sequence-containing peptides, are still problematic and the peptides often have low yield and purity in SPPS[24–29] (for a review, see Ref. 30). It is known that the difficult sequences are generally hydrophobic and promote aggregation in solvents during synthesis and purification. This aggregation is attributed to intermolecular hydrophobic interaction and a hydrogen-bond network among resin-bound peptide chains, resulting in the formation of extended secondary structures such as β-sheets.[24,25] To solve this problem, Mutter et al. developed building blocks, so-called pseudo-prolines, which are dipeptide derivatives, including Ser/Thr-derived oxazolidines or Cys-derived thiazolidine.[26–28] Sheppard and Johnson et al. also reported a building block, 2-hydroxy-4-methoxybenzyl (Hmb), a protecting group for the backbone amide nitrogen[29] (for a review, see Ref. 30). These special building blocks were designed to disrupt the secondary structure formed by interchain hydrogen bonding. However, in these approaches, prior modification of Fmoc-amino acids by two to six steps of solution-phase synthesis is required, and strong acids are required to remove the building blocks. Therefore, the development of novel methods using conventional amino acid derivatives are of great significance in the synthesis of difficult sequence-containing peptides. Since it is known that the difficult sequences are generally hydrophobic and promote aggregation in solvents during synthesis and purification, we applied the concept of O–N intramolecular acyl migration, namely, our idea was based on the synthesis of the hydrophilic *O*-acyl isopeptide followed by the O–N intramolecular acyl migration reaction leading to the desired peptide.

In this article, we review our application of the O–N intramolecular acyl migration reaction to the development of water-soluble prodrugs of taxoids and the development of a novel synthetic method for difficult sequence-containing peptides. In particular, in the latter case, we first evaluated the viability of our synthetic strategy through the synthesis of difficult sequence-containing small model peptides, and then applied it to the synthesis of a natural difficult sequence-containing large peptide, Alzheimer's disease-related peptide Aβ1–42. We successfully synthesized these "*O*-acyl isopeptides," which has not only improved solubility in various media and achieved efficient high performance liquid chromatography (HPLC) purification, but also changed the unfavorable nature of the difficult sequence during SPPS, resulting in the efficient synthesis of *O*-acyl isopeptides. The subsequent O–N intramolecular acyl migration of purified *O*-acyl isopeptides afforded the desired peptides in high yield and purity. It is interesting to note that only one *O*-acyl group insertion drastically improved the nature of the difficult sequence in Aβ1–42 with 42 amino acid residues. Therefore, this new method will contribute to the synthesis of larger difficult sequence-containing peptides.



2 (isotaxel); R = Ph
6; R = *i*-Pr

1 (paclitaxel); R = Ph
5 (canadensol); R = *i*-Pr

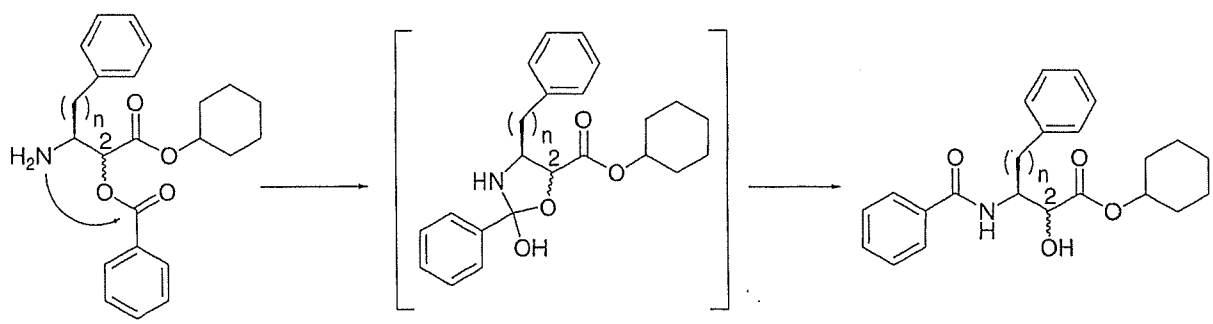
FIGURE 2 The O–N acyl migration of isotaxel **2** to paclitaxel **1** and **6** to canadensol **5**.

DESIGN AND SYNTHESIS OF ISOTAXEL: A NOVEL WATER-SOLUBLE PACLITAXEL PRODRUG WITH NO AUXILIARY AND NO BY-PRODUCT

To improve the poor water solubility of paclitaxel **1** (0.00025 mg/mL), many water-soluble prodrugs with modified hydroxyl groups at the C-2′ and/or C-7 positions with solubilizing moieties have been reported, and some are currently undergoing clinical evaluation.[31–33] However, none of these applications are presently in clinical use and the released auxiliary moieties may have some unfavorable effects in vivo.[34] These factors suggest that novel approaches for water-soluble prodrugs of **1** are needed.

In the O–N intramolecular acyl migration-type water-soluble prodrugs of HIV-1 protease inhibitors with Apns, the observed *O*-acyl migration of prodrugs such as KNI-272 (Figure 1B) under physiological conditions was so rapid ($t_{1/2} < 1$ min)[9] that it might prove difficult for paclitaxel, which has a similar α-hydroxy-β-amino acid residue, i.e., (*2R, 3S*)-phenylisoserine, to undergo a systemic distribution after injection. However, we also observed that *O*-benzoyl group migration in a prodrug of Bz–Apns–Thz–NHBu^t (KNI-565; Thz: thiazolidine-4-carboxylic acid) was relatively slow,[10] possibly due to the less electrophilic nature of the carbonyl carbon and the steric effect of the phenyl ring. Hence, we decided to perform a model study to create a new type of water-soluble paclitaxel prodrug (Figure 2).

To examine the effect of the benzoyl group and stereochemistry of α-hydroxy-β-amino acids on the kinetics of O–N acyl migration, we designed three



3a, n=1, 2-*S*
3b, n=1, 2-*R*
3c, n=0, 2-*R*

4a, n=1, 2-*S*
4b, n=1, 2-*R*
4c, n=0, 2-*R*

FIGURE 3 The O–N acyl migration in model compounds, designed to examine the effect of the benzoyl group and stereochemistry of α-hydroxy-β-amino acids on the kinetics.