Bound Water Molecules at the Interface between the HIV-1 Protease and a Potent Inhibitor, KNI-272, Determined by NMR

Yun-Xing Wang,[†] Darón I. Freedberg,[†] Paul T. Wingfield,[‡] Stephen J. Stahl,[‡] Joshua D. Kaufman,[‡] Yoshiaki Kiso,[§] T. Narayana Bhat,[⊥] John W. Erickson,[⊥] and Dennis A. Torchia*[,†]

Contribution from the Molecular Structural Biology Unit, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20892, Protein Expression Lab, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Bethesda, Maryland 20892, Structural Biochemistry Program, NCI-Frederick Cancer Research and Development Center, SAIC-Frederick, Frederick, Maryland 21702, and Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607, Japan

Received July 29, 1996[⊗]

Abstract: KNI-272 is a peptidomimetic transition state analog inhibitor, having very high specificity and binding affinity for the HIV-1 protease. In order to understand the interactions that enhance drug binding to the protease, we recorded 2D water/NOESY and water/ROESY spectra to identify water molecules that bind tightly to the protease/KNI-272 complex. Well-ordered water molecules are observed at the protease/inhibitor interface in the crystal structure of the complex that have short interproton distances to the Ile50/150, Ala28/128, and Asp29/129 amide protons. The cross peaks between these protein protons and water protons, observed in water/NOESY and water/ROESY spectra, provide strong evidence that these water molecules are present in the solution structure of the complex. Analysis of measured NOE and ROE cross relaxation rates indicates that, in solution, these water molecules have long residence times, at least 1 ns and possibly greater than 7 ns. The presence of long-lived hydration water molecules at the protein/inhibitor interface suggests that interactions involving these water molecules contribute to the potency of the inhibitor. Hence, consideration of the potential role of hydration water molecules in stabilizing protein/inhibitor structures could contribute to improved drug design and to a better understanding of the mechanisms of drug resistance.

Introduction

HIV-1 protease is a primary target of AIDS antiviral agents,[1] and crystal structures of the protease[2] have been used to design increasingly more potent inhibitors. The structure of a typical potent inhibitor complements that of the substrate binding site and maximizes hydrophobic and hydrogen-bonding interactions with it.[2] In addition, a tetrahedrally hydrogen-bonded water molecule, W301, bridges the flaps of the protease to the inhibitor in nearly every crystal structure of the inhibited HIV protein.[2a] This observation has been utilized to design a novel class of cyclic urea type inhibitors[3a] whose specificity derives from their ability to mimic and replace W301, as observed in the crystalline state[3a] and in solution.[3b] W301 and several other water molecules are observed at the protease/drug interface in the crystal structure of the protease complexed with KNI-272, Figure 1,[4] a highly potent inhibitor, currently in clinical trials.[5] This observation suggests that, in addition to W301, these water molecules may have an important role in the formation of the protease/KNI-272 complex. Herein we report NMR studies

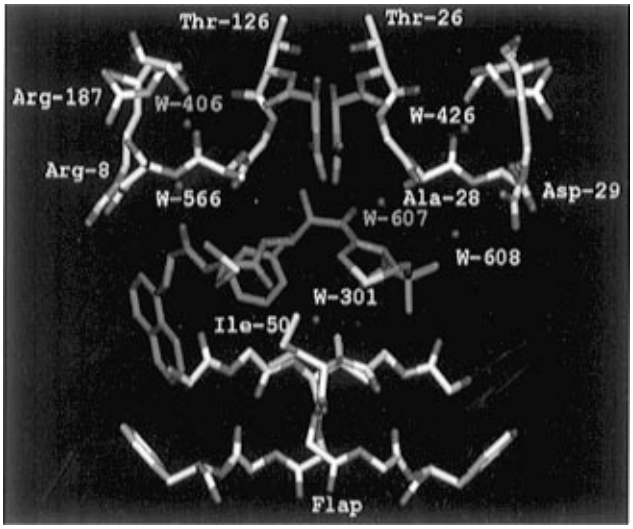

Figure 1. View of the active site of the protease/KNI-272 crystal structure[4] showing the locations of the six ordered water molecules. W426 and W406 occupy nearly equivalent positions relative to each monomer. The remaining four water molecules occupy unique sites. Atoms colored by type: oxygen, red; nitrogen, is blue; carbon, (protein) white, (KNI-272) orange. Cross relaxation rates of water molecules labeled in white (derived from W301/I150/I150, W608/D29, W426/A28, and W566/D129 NOEs and ROEs) are listed in Table 1. As discussed in the footnotes to Table 1, WNOESY/WROESY cross peak signal intensities for W406 and W607 (pink labels) could not be determined because of signal overlap.

* To whom correspondence should be addressed. Phone: (301) 496-5750. Fax: (301) 402-1512. E-mail: torchia@yoda.nidr.nih.gov.
[†] National Institutes of Health.
[‡] National Institute of Arthritis and Musculoskeletal and Skin Diseases.
[§] Kyoto Pharmaceutical University.
[⊥] SAIC-Frederick.
[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1996.

(1) Richman, D. D. *Nature (London)* **1995**, *374*, 494−495.
(2) (a) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543−585. (b) Erickson, J. W. *Nature, Struct. Biol.* **1995**, *2*, 523−529.
(3) (a) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C.-H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. *Science (Washington, D.C.)* **1994**, *263*, 380−384. (b) Grzesiek, S.; Bax, A.; Nicholson, L. K.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Eyermann, C. J.; Torchia, D. A.; Hodge, C. N.; Lam, P. Y. S.; Jadhav, P. K.; Chang, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 1581−1582.

S0002-7863(96)02612-1 CCC: $12.00

whose goals are to determine if these water molecules are present in solution and to estimate their residence times.

Results and Discussion

NOEs and ROEs between water protons and protein protons provide a means to detect water molecules bound to proteins.[6−9] Such interactions can be detected in water/NOESY (WNOESY) and water/ROESY (WROESY) spectra.[8] These spectra also yield cross relaxation rates, σ_{noe} and σ_{roe}, parameters related to the water residence time, τ_r.[8] Models of water dynamics show[6] that when $\tau_r \ll 1/\omega$ (ω is the ^{1}H Larmor frequency) σ_{noe} equals σ_{roe} and is positive. As τ_r increases, σ_{roe} monotonically increases; in contrast, as τ_r increases, σ_{noe} changes sign, attaining a slow limit value of $-\sigma_{roe}/2$, when $\omega^2\tau_r^2 \gg 1$, i.e., $\tau_r > 2$ ns at 500 MHz.

Numerous cross peaks are observed in the amide WNOESY and WROESY spectra of the protease/KNI-272 complex, Figure 2. The spectra were recorded at 45 °C using a ^{2}H(85%)/^{15}N-(95%) labeled protein sample to improve the sensitivity of the I50/I150 signals, which are severely broadened at 34 °C, by conformational exchange of the protease flaps.[10] Deuteration improves sensitivity and resolution in WNOESY and WROESY experiments by eliminating most ^{1}HN−^{1}HC dipolar interactions, and also simplifies the interpretation of the data by reducing spin diffusion.

The crystal structure reveals six well-ordered water molecules (*B*-factors < 26) at the KNI-272 binding site,[4] Figure 1. Five of these waters have relatively short internuclear distances to the amide protons of A28/128, D29/129, and I50/150, Table 1. In addition, each of these amide protons is at least 4.2 Å from the closest exchangeable proton, Table S1, and has an H/D exchange lifetime in excess of 15 min. Hence, their cross peaks are assigned to NOE/ROE interactions with water molecules.

No other signals observed in Figure 2 can be exclusively assigned to a direct dipolar interaction between a water proton and an amide protein proton. The positive signals (drawn with dashed contour lines in Figure 2) observed in the WROESY spectrum (e.g., G16/116, S37/137) arise from amide protons in rapid exchange with bulk water. This conclusion is consistent with H/D exchange experiments which show that all amides having positive WROESY cross peaks have exchange lifetimes of less than ∼5 min, Table S1.

Amide protons from residues such as T4/104 and T12/112 are each within 4.5 Å of a labile (OH) protein proton, but not a water proton, in the crystal structure, Table S1. Hence, their cross peaks are assigned to an indirect mechanism, consisting of an NOE between the amide proton and the labile proton, which rapidly exchanges with solvent. Direct and indirect dipolar mechanisms can contribute to cross peak intensities of amide protons that are within 4.5 Å of both a labile proton and a water proton. The mechanism(s) assigned to each cross peak is listed in Table S1.

Cross relaxation rates, Tables 1 and S1, can be used to estimate the τ_r values of bound water molecules. The values

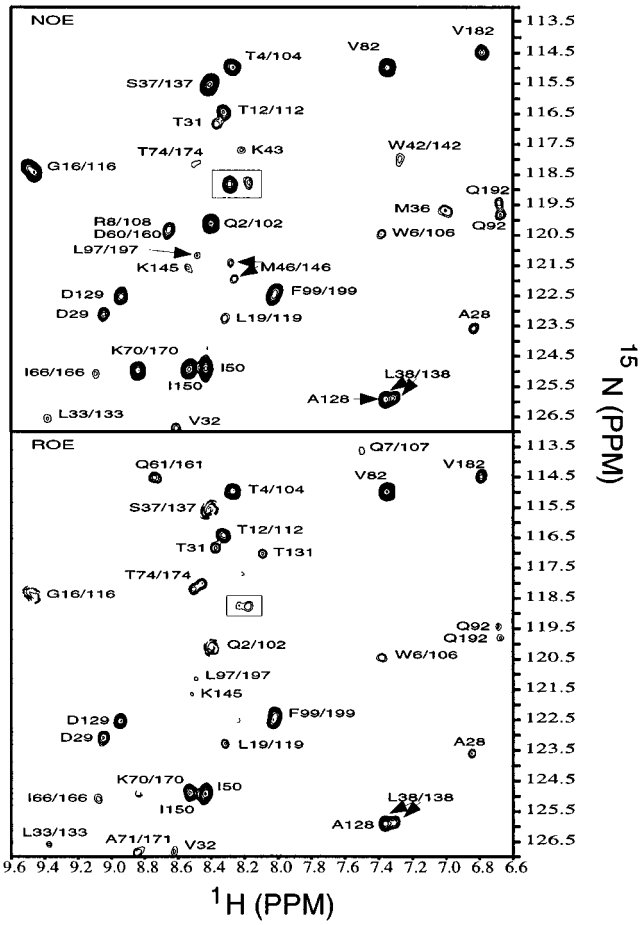

Figure 2. Portions of the WNOESY (top) and WROESY (bottom) difference spectra recorded with respective mixing times of 90 and 50 ms for the KNI-272 complex at 45 °C, 50 mM NaAc buffer, pH 4.55. Other spectral parameters are as in Figure S1. Signal assignments are taken from Wang *et al.*[15] Cross peaks within boxes cannot be unambiguously assigned because of signal overlap. Note that since cross peaks having *negative* intensities correspond to direct NOE and ROE interactions with water molecules,[8] these are drawn with *unbroken* contour lines. *Positive* cross peaks in the WROESY spectrum (*drawn with dashed contours*) are due to amide hydrogen exchange with solvent water.

of R_σ are all less than −0.4, indicating that $\tau_r > 1$ ns. This conclusion is confirmed by a simple calculation of the cross relaxation rates, Table 1, which assumes that (a) τ_r is greater than the overall correlation time, τ_o (7.4 ns at 45 °C), and (b) the interproton distances in solution are equal to those calculated using the hydrogen atom coordinates derived from the crystal structure by X-PLOR.[11] The agreement with experiment is reasonable, Table 1, when one considers that a 25% error in an interproton distance results in a 4-fold error in the calculated relaxation rate. The best agreement between the experimental and calculated rates is obtained for the tetrahedrally coordinated W301, which is expected to have the most precisely defined hydrogen atom positions.[4]

Although ordered water molecules are found outside the KNI-272 binding site in the X-ray structure, in general, these waters do not have long residence times. For example, water molecules 485/408/439/418 have *B*-factors in the range of 21−23 Å^2. These water molecules are partially enclosed by cuplike indentations on the protein surface and are hydrogen-bonded to the amide protons of T31/131/T74/174, respectively. In addition one proton of each water molecule is within 2.5 Å of

(4) Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Minoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581−590.
(5) Chokekijchai, S.; Shirasaka, T.; Weinstein, J. N.; Mitsuya, H. *Antiviral Res.* **1995**, *28*, 25−38.
(6) Otting, G.; Liepinsh, E.; Wüthrich, K. *Science (Washington, D.C.)* **1991**, *254*, 974−980.
(7) Otting, G.; Liepinsh, E. *Acc. Chem. Res.* **1995**, *28*, 171−177.
(8) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 627−638.
(9) Clore, G. M.; Bax, A.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 5671−5676.
(10) Nicholson, L. K.; Yamazaki, T.; Torchia, D. A.; Grzesiek, S.; Bax, A.; Kaufman, J. D.; Stahl, S. J.; Wingfield, P. T.; Lam, P. Y. S.; Jadhav, P. K.; Hodge, C. N.; Domaille, P. J.; Chang, C.-H. *Nature, Struct. Biol.* **1995**, *2*, 274−280.
(11) Brünger, A. T. *X-PLOR Manual Version 3.1: A System for X-RAY Crystallography and NMR*; Yale University: New Haven, CT, 1992; p 382.

Table 1. Distances between protons of Ordered Water Molecules and Amide Protons at the Protease/KNI-272 Binding Site,[4] Together with a Comparison of the Experimental and Calculated Cross Relaxation Rates[a]

residue	water[c]	r	σ_{noe}	σ_{roe}	R_σ[d]
A28[b]	426	3.7, 4.1	−0.64 (−0.26)	1.08 (0.53)	−0.59 (−0.5)
D29	608	2.5, 2.9	−0.41 (−2.34)	0.99 (4.71)	−0.41 (−0.5)
D129	566	3.3, 3.6	−1.59 (−0.49)	3.05 (0.99)	−0.52 (−0.5)
I50	301	2.6, 2.6	−3.47 (−2.62)	7.54 (5.28)	−0.46 (−0.5)
I150	301	2.4, 2.4	−4.00 (−3.98)	6.97 (8.00)	−0.57 (−0.5)

[a] In s^{-1}, calculated values in parentheses, derived using the slow motion limit equations, $\sigma_{roe} = d\tau_o/\sum r^6 = -2\sigma_{noe}$, where d is the dipolar coupling constant, 3.610×10^{-9} $s^{-2}/Å^{-6}$, τ_o is the overall correlation time, 7.4 ns, and r is the interproton distance between the amide proton and each water proton, in Å, listed in column 3. It is assumed for simplicity that the order parameter of the dipolar interaction is unity. If internal motion occurs, the order parameter is less than 1, leading to a reduction in the calculated cross relaxation rates.[13] [b] The A128 NH is 3.6 Å from a water 406 proton, but because signals of A128 and either L38 or L138 overlap, cross relaxation rates cannot be measured. However, the relatively strong NOE and ROE cross peaks seen at the A128 signal position, Figure 2, are consistent with a dipolar interaction between the A128NH and W406. [c] A sixth water molecule, W607, observed at the active site of the complex, Figure 1, is not within 4.5 Å of an amide proton, but is ~2.3 Å from the δ2-methyl protons of L123. Cross peaks are observed at the L123 δ-chemical shift in WNOESY and WROESY difference spectra, recorded using a $^{13}C/^{15}N$-labeled protease sample, not shown; however, the δ-methyl signals of L19/119/63/163 overlap those of L123, and preclude measurement of L123/W607 cross peak intensities. [d] $R_\sigma = \sigma_{noe}/\sigma_{roe}$ with experimental uncertainty ±0.05 for I50/150 and ±0.1 otherwise.

the Thr amide proton of its hydrogen-bonded partner. Nevertheless, ROE cross peaks are weak and NOE cross peaks are either weak or absent for these four amide protons, Figure 2. In addition, the values of R_σ are in the range of 0.2−0.35 or less, Table S1, indicating that either the residence times of these water molecules are short, < 1 ns, or the water/protein interproton vectors undergo large amplitude rapid reorientation.

Six other ordered (B < 30 Å^2) water molecules observed in the crystal structure (water molecules 430, 432, 440, 477, 489, and 509) have at least one proton within 3 Å of an amide proton, but their cross peaks are either very weak or not observed. The fact that these water molecules are fully solvent accessible and located on the surface of the protein may explain their short residence times and/or rapid reorientation. Other factors being equal, it is reasonable to suppose that a solvent-exposed bound water molecule will have a shorter residence time than an interior water. Evidence has been presented that positionally disordered buried water in interleukin-1β has as a residence time of 1 ns or greater.[12] The observation that these six surface water molecules in the protease/KNI-272 complex have relatively small B-factors is probably due to the fact that, except for W509, they occupy sites that are close to the surfaces of two protein molecules in the crystal; hence, they may be ordered by crystal packing forces.

Our finding that four to six water molecules at the KNI-272 binding site have large τ_r values differs from our observation that only one to three such water molecules are found at the inhibitor binding site in the complex formed by the protease and the cyclic urea type inhibitor DMP323.[13] DMP323 is also a specific, potent inhibitor of the protease, which uses its urea oxygen, and possibly some of its four hydroxyl groups, as water mimics to interact with the protease.[3,13] In contrast, in the complex formed with KNI-272, a transition state analog,[14] the X-ray[4] and NMR data indicate that water molecules play a more significant role in the structure of the complex.

Taken together, these results illustrate the distinct ways that inhibitors are able to interact with the protease. In addition, they indicate that interactions involving water molecules, in addition to W301, should be considered in inhibitor design, and that the impact of protease mutations upon such interactions may be a significant mechanism of drug resistance.

Experimental Section

Production of $^2H/^{15}N$-Labeled HIV-1 Protease. HIV-1 protease (strain HXB2) was produced in *E. coli* as described by Cheng *et al.*[16] using the expression vector pET11a and host bacterial strain BL21 (DE3).[17] Cysteine to alanine mutations at positions 67 and 95 were introduced using the polymerase chain reaction[18] in order to avoid intermolecular disulfide formation. Plasmids were transformed into BL (DE3) *E. coli* cells and plated on LB ampicillin plates. Single colonies were picked and plated (one colony per plate) on LB plates made using 100% D_2O and which contained 100 μg/mL carbenicillin. Colonies were plated so as to produce a lawn of cells. After overnight incubation at 37 °C, the cells from about half the area of the plate were transferred directly into a 2-L Braun Model MD fermentor containing 0.5 L of minimal media made with 100% D_2O containing $^{15}NH_4Cl$ (5 g/L). All media additives were deuterated, and protein purification proceeded as described.[19] Assessment of protein labeling, made by electrospray mass spectrometry, indicated ca. 85% ^2H labeling (of nonexchangeable hydrogen) and 95% ^{15}N labeling of the protein.

Signal Assignments and Nomenclature. The HIV protease is a homodimer made up of two identical polypeptide chains, each containing 99 amino acids. We number the amino acid residues in each monomer 1−99 and 101−199, respectively. The 2-fold symmetry of the dimer is broken when the protease binds an asymmetric inhibitor such as KNI-272. In an asymmetric complex, chemical shifts of residues i and $i + 100$ are not degenerate, and monomer specific signal assignments of the protease/KNI-272 complex have been reported.[15] The amide assignments relevant to this work are listed in Table S1.

NMR Spectra. WNOESY and WROESY spectra were recorded as previously described[8] on a Bruker AMX 500 MHz spectrometer. The sample of ca. 1.5 mM HIV-1 protease/KNI-272 complex was maintained at 45 °C in 50 mM NaAc buffer, pH 4.55, in 95% H_2O/5% D_2O. The proton ($F2$) and ^{15}N ($F1$) carrier frequencies were 4.58 and 120.0 ppm, respectively, with spectral widths of 8064.5 Hz ($F2$) and 2000 Hz ($F1$). A total of 128 hypercomplex points were collected in $F1$ with an acquisition time of 65 ms. Data were zero filled to 1024 complex points in $F2$ and 512 complex points in $F1$, multiplied by a Lorentz-to-Gauss window function with an inverse Lorentz width of 7.0 Hz and a Gaussian width of 20 Hz in both $F2$ and $F1$, and then Fourier transformed. Cross peaks arising from residual protein protons or inhibitor protons that resonate near water were identified in a control experiment[8] that selectively attenuates water-derived cross peaks. Relaxation rates in the laboratory ($\rho_1 + k_n$) and rotating ($\rho_2 + k_r$) were determined using the two-point approximation as described previously,[8] and are listed in Table S1.

Cross Relaxation Rates. The ratios of the measured signal intensities of cross peaks observed in sum and difference WNOESY

(12) Ernst, J. A.; Clubb, R. T.; Zhou, H.-X.; Gronenborn, A. M.; Clore, G. M. *Science* **1995**, *267*, 1813−1817.

(13) Wang, Y.-X.; Freedberg, D. I.; Grzesiek, S.; Torchia, D. A.; Wingfield, P. T.; Kaufman, J. D.; Stahl, S. J.; Chang, C.-H.; Hodge, C. N. *Biochemistry* **1996**, *35*, 12694−12704.

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

(15) Wang, Y.-X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996**, *35*, 9945−9950. Corrected Tables S2 and S3 of the Supporting Information of the *Biochemistry* paper have been submitted as a correction in *Biochemistry*. They are also included in the Supporting Information for the present paper.

(16) Cheng, Y-S. E.; McGowan, M. H.; Kettner, C. A.; Schloss, J. V.; Erickson-Viitanen, S.; Yin, F. H. *Gene* **1990**, *87*, 243−248.

(17) Higuchi, R.; Krummel, B.; Saiki, R. K. *Nucleic Acid Res.* **1988**, *16*, 7351−7367.

(18) Studier, F. W.; Rosenberg, A. I.; Dunn, J. J.; Dubendorf, J. W. *Methods Enzymol.* **1990**, *185*, 60−89.

(19) Yamazaki, T.; Hinck, A. P.; Wang, Y-X.; Nicholson, L. K.; Torchia, D. A.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Chang, C-H.; Domaille, P. J.; Lam, P. Y. S. *Protein Sci.* **1996**, *5*, 495−506.

and WROESY spectra were used to determine the normalized signal intensities,[8] ζ_{noe} and ζ_{roe}, listed in Table S1. These quantities together with the measured relaxation rates ($\rho_1 + k_n$) and ($\rho_2 + k_r$) were used, as described previously,[8] to derive the values of σ_{noe} and σ_{roe} (Table S1).

Acknowledgment. We thank Stephen Grzesiek and Ad Bax for the WNOESY and WROESY pulse sequences and Ann Frances Miller and Michelle Markus for critically reading the paper. This work was supported by the intramural AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

Supporting Information Available: Table S1, listing signal assignments, relaxation rates, interproton distances, H/D exchange lifetime cross peak mechanism classifications, calculation of the effect of amide spin diffusion on the cross relaxation rate determination, Figure S1, WNOESY and WROESY sum spectra, and corrected Tables S2 and S3 from ref 15, listing the ^{1}H assignments of KNI-272 and the ^{1}H, ^{13}C, and ^{15}N assignments of the HIV-1 protease in the protease/KNI-272 complex, respectively (12 pages). See any current masthead page for ordering and Internet access instructions.

JA962612I