*C h a p t e r  I I*

# STEREOSELECTIVE INCORPORATION OF UNSATURATED ISOLEUCINE ANALOGUES INTO PROTEINS *IN VIVO**

## II.1 Abstract

The unsaturated amino acids 2-amino-3-methyl-4-pentenoic acid (E-Ile) and 2-amino-3-methyl-4-pentynoic acid (Y-Ile) were prepared, and E-Ile was successfully separated into its SS, RR and SR, RS diastereomeric pairs. The translational activities of the SS-E-Ile, SR-E-Ile, and Y-Ile analogues were assessed using an *Escherichia coli* (*E. coli*) strain auxotrophic for isoleucine (Ile**)**. SS-E-Ile was incorporated into the test protein murine dihydrofolate reductase (mDHFR) in place of isoleucine at a rate of substitution of up to 72%, while SR-E-Ile showed no conclusive evidence of translational activity. At least one stereoisomer of Y-Ile also supported protein production, but the stereochemical purity of the amino acid samples was not sufficient to investigate stereochemical discrimination. *In vitro* ATP-$PP_i$ exchange assays indicate that SS-E-Ile is activated by the isoleucyl-tRNA synthetase (IleRS) at a rate comparable to isoleucine; SR-E-Ile is activated approximately 100 times more slowly.

**Sections of this chapter are excerpted from a manuscript accepted to ChemBioChem by Marissa L. Mock, Thierry Michon, Jan C. M. van Hest, and David Tirrell, 2005.*

## II.2 Introduction

Genetic engineering provides a tool with which one can prepare complex macromolecules possessing both precisely controlled architectures and specific catalytic or biological activity. Recent work has shown the advantages of using the biosynthetic machinery to produce new materials (for a review see reference [1]). The use of monomers other than the twenty canonical amino acids enables the introduction of new functionality into proteins, creating the potential for novel physical and chemical properties. Analogues of many of the canonical amino acids have been incorporated into proteins in *E. coli* using the wild-type biosynthetic machinery, e.g. [2, 3], while modifications of that machinery have permitted the incorporation of a still broader set of non-canonical amino acids [4-14]. Increasing the number of amino acid monomers that can be incorporated into proteins, and thereby the range of physical properties and chemistries available, requires detailed understanding of the biosynthetic apparatus.

Protein synthesis involves transcription of the information contained in DNA into mRNA and translation of the mRNA into polypeptide chains. The aminoacyl-tRNA synthetases (aaRSs) are essential to the fidelity of this process. Each aaRS selectively catalyzes the activation of the carboxylate group of the appropriate amino acid by reaction with adenosine triphosphate (ATP) to produce the aminoacyl adenylate, which reacts with terminal hydroxyl group of a cognate tRNA to produce aminoacyl-tRNA. The selectivity of the aaRSs is an important consideration in any attempt to incorporate nonnatural amino acids into proteins *in vivo*. Modifications of the aaRSs, through enlarging the active site [14-16] or decreasing editing activity [13, 17], have been shown to permit incorporation of analogues that are not usually incorporated into proteins. The

rational modification of aaRSs to allow use of a wider range of nonnatural amino acids requires an understanding of the mechanism(s) of selectivity of each individual aaRS.

The isoleucyl-tRNA synthetase (IleRS) has been well studied, in part because it must perform a significant feat of selective recognition as it discriminates its cognate amino acid isoleucine (Ile) from the natural amino acid valine (Val), which differs in chemical structure by only one methylene group. Pauling calculated that the additional binding energy contributed by the extra methylene group should at most result in a discrimination of 1/20 [18], while the erroneous substitution of Val for Ile actually occurs at a rate of about 1 in 3000 [19]. In fact, IleRS does misactivate Val (approximately 140 times more slowly than Ile [20, 21]) and later hydrolyzes the misactivated amino acid in an editing site located ~34 Å from the synthetic site of the enzyme [22-24]. Isoleucine contains two chiral centers, one at the alpha carbon and another at the beta carbon. The stereoisomer of 2-amino-3-methyl-heptanoic acid incorporated into proteins is (2S, 3S), designated Ile. L-allo-Ile (2S, 3R) has the correct configuration at the α-carbon, but the opposite configuration at the β-position. It is not incorporated into proteins, although there is evidence that it is bound and activated by IleRS [19, 25-27].

To expand further the chemistries available for the modification of proteins, we prepared the unsaturated Ile analogues (2S, 3S and 2R, 3R)-2-amino-3-methyl-4-pentenoic acid (SS, RR-E-Ile), (2S, 3R and 2R, 3S)-2-amino-3-methyl-4-pentenoic acid (SR, RS-E-Ile), and 2-amino-3-methyl-4-pentynoic acid (as a mixture of the SS, RR, SR, and RS stereoisomers) (Figure II-1), which have been shown previously to inhibit growth of *E. coli* [28]. We are especially interested in unsaturated amino acid analogues because of

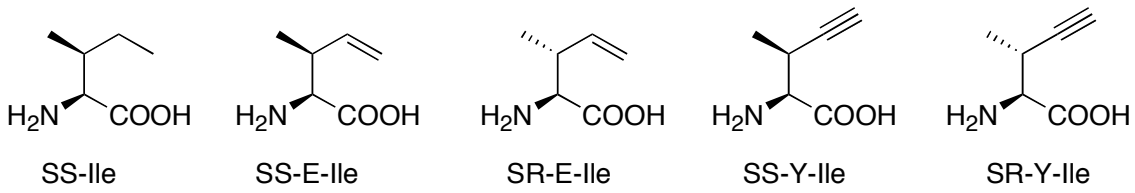


**Figure II-1.** Unsaturated isoleucine analogues SS-E-Ile and SR-E-Ile differ in the stereochemistry at the β-carbon. The stereoisomer of isoleucine incorporated into proteins (Ile) is 2S, 3S.

the versatile chemistry of alkenes and alkynes. For example, supramolecular structures made up of weakly hydrogen-bonded cyclic peptides can be stabilized through inter-peptide crosslinking [29] utilizing ruthenium-catalyzed ring-closing metathesis of pendant alkene moieties [30, 31] and Cu(I)-catalyzed azide-alkyne cycloaddition has been used to modify *E. coli* cells [32]. The stereoisomer pairs were evaluated with respect to incorporation into a test protein, murine dihydrofolate reductase (mDHFR), in an *E. coli* strain rendered auxotrophic for Ile. The kinetics of activation of SS and SR-E-Ile by the IleRS were also determined *in vitro* through ATP/$PP_i$ exchange.

## II.3 Methods

### *II.3.1 Synthesis of Analogues*

II.3.1.1 General procedures

Glassware was dried at 150 °C and cooled under argon prior to use. Tetrahydrofuran (THF) was freshly distilled from sodium/benzophenone. Other reagents were used as purchased. *Cis*- and *trans*-crotyl alcohol were purchased from Chemsampco, Trenton, NJ. $^1$H-NMR spectra were recorded on a Varian Mercury 300 MHz spectrophotometer. Column chromatography was performed on silica gel (300

Mesh, Baker) or alumina (80-200 Mesh, EM Science). Silica ($60F_{254}$ EM Science) was used for thin layer chromatography.

II.3.1.2 SS- and SR-2-amino-3-methyl-4-pentenoic acid (E-Ile)

SS- and SR-E-Ile were prepared according to Figure II-2; the stereochemistry at the β-carbon of the final product was controlled by choosing either *cis*- or *trans*-crotyl alcohol as the starting material (Figure II-3). *N*-benzyloxycarbonyl (Cbz)-protected glycine crotyl esters were prepared according to Hassner and Alexian [33], with slight modification. Cbz-glycine (17.44, 83.4 mmol) was dissolved in 100 mL THF under $N_2$. Either *cis*- or *trans*-crotyl alcohol (6.00 g, 83.2 mmol) was added, followed by 17.15 g (83 mmol) dicyclohexyl carbodiimide and 50 mg (6.1 mmol) dimethylaminopyridine. An exotherm was observed, and a precipitate immediately formed. The mixture was allowed to stir at room temperature for 2 days before being filtered through Celite. The

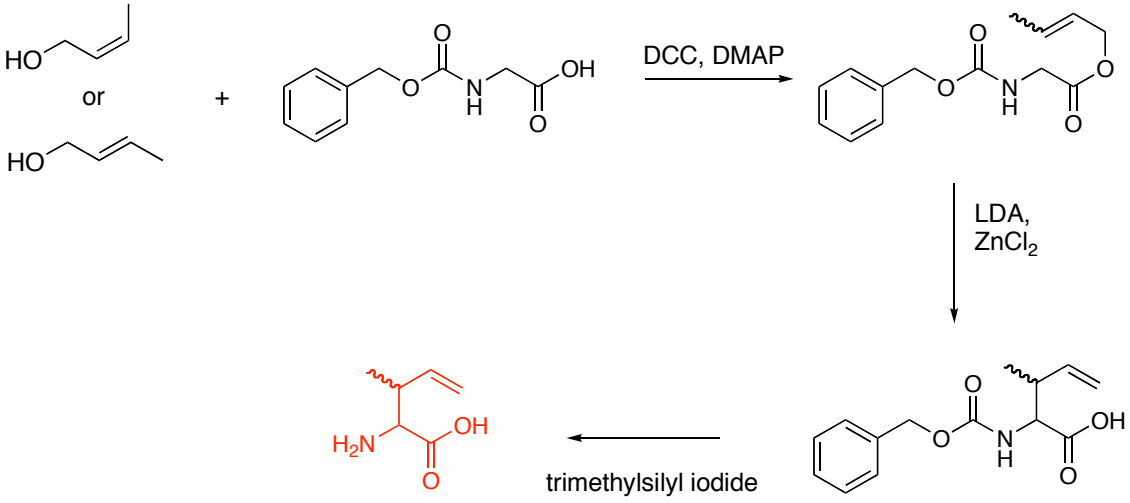


**Figure II-2**. Synthesis scheme for SS- and SR-E-Ile. The amine-protected crotyl esters undergo Claisen rearrangement; subsequent deprotection gives stereochemically enriched product.

solvent was evaporated, and the crude product was purified by flash chromatography ($CH_2Cl_2$). A viscous colorless oil was obtained by evaporation of the eluent (yield *trans*-isomer: 65 %, *cis*-isomer: 32 %). $^1$H-NMR ($CDCl_3$): δ 1.73 (d, 3H, $J$ = 6.9, C<u>H</u>$_3$), 3.98 (d, 2H, $J$ = 6.4, C<u>H</u>$_2$-CO), 4.57 (d, 2H, $J$ = 6.7, O-C<u>H</u>$_2$), 5.17 (s, 2H, Ph-C<u>H</u>$_2$), 5.56 (m, 1H, $J_E$ = 17.4, $J_Z$ = *10.2,* C=C<u>H</u>-CH$_3$), 5.81 (m, 1H, $J_E$ = 17.4, $J_Z$ = *10.2,* O-CH$_2$-C<u>H</u>=C), 7.35 (s, 5H, Ph).

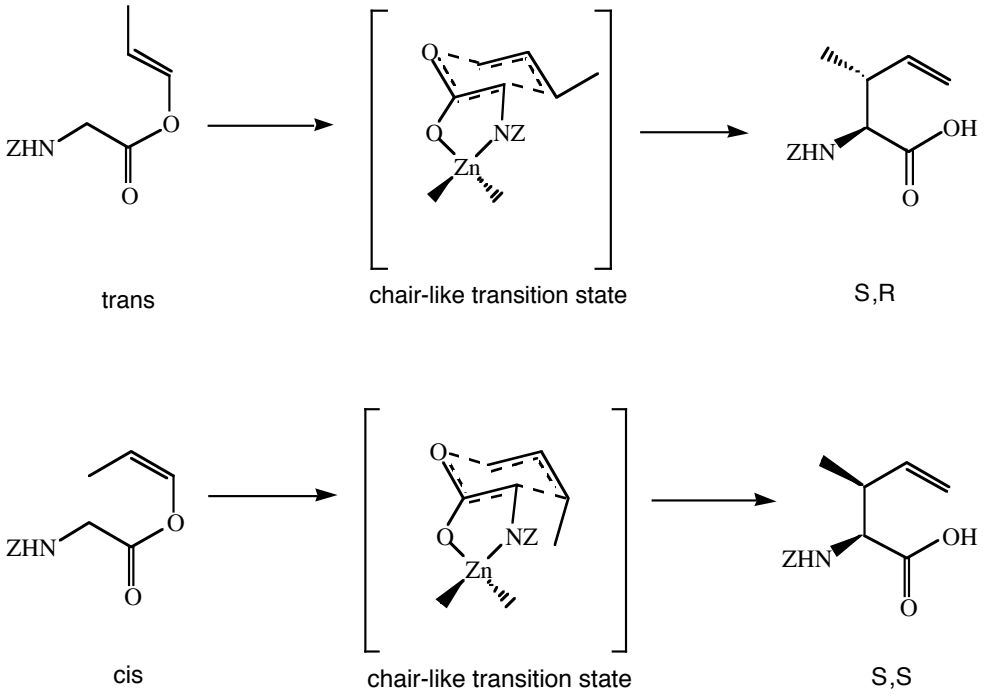


**Figure II-3.** Stereoselectivity in the preparation of E-Ile is a result of a preference for the chair-like transition state during the Claisen rearrangement, which determines the stereochemistry at the β-carbon in the product. The *trans*-crotyl alcohol yields SR (and RS)-E-Ile, while the *cis*-crotyl alcohol gives SS (and RR)-E-Ile.

*N*-benzyloxycarbonyl (Cbz)-protected glycine crotyl esters were rearranged following Kazmaier [34] to *N*-benzyloxycarbonyl-2-amino-3-methyl-4-pentenoic acid. Diisopropyl amine (6.9 mL, 45.8 mmol) was dissolved in 40 mL dry THF under argon.

The stirred solution was cooled to –20 °C and 26 mL (41.6 mmol) 1.6 M *n*-butyl lithium in *n*-hexane was added. After 20 minutes, the solution was cooled to –78 °C, and 5.4 g (20.5 mmol) Cbz-protected glycine crotyl ester (either *cis* or *trans*) in 20 mL THF and 25 mL of a 0.5 M $ZnCl_2$ solution in THF were added simultaneously over a 30 min period. The solution remained homogeneous. After an additional 30 min at –78 °C the solution was allowed to warm to room temperature. The rearrangement was monitored by thin layer chromatography (3: 7 ethyl acetate:dichloromethane). After 5 hours, 30 % (*trans*) to 50 % (*cis*) of the starting ester was still present in the reaction medium. Additional incubation time did not improve the yield of the rearrangement. The reaction was terminated by addition of 10 mL 1 M HCl. The ether phase was extracted with two 75 mL volumes of 1 M NaOH. The aqueous phase was neutralized with concentrated HCl to precipitate the acid product, which was extracted with diethyl ether (150 mL). The ether layer was dried over $MgSO_4$, and the solvent was evaporated to give a yellow oil. The acids were purified by flash chromatography (99:1 ethyl acetate:acetic acid).

Either SS, RR or RS, SR Cbz-protected 2-amino-3-methyl-4-pentenoic acid (0.8 g, 3 mmol) was dissolved in 10 mL dry $CHCl_3$ under nitrogen atmosphere. Trimethylsilyl iodide (1.2 mL, 8.8 mmol) was added. After 20 min stirring at room temperature, the reaction was quenched by addition of 1 mL methanol. The solvent was evaporated, the crude product was dissolved in 10 mL 30 % v/v acetic acid, and the solution was washed twice with 15 mL diethyl ether. The aqueous layer was evaporated, yielding a yellow oil. The SR, RS mixture of isomers crystallized upon cooling. Recrystallization from 4:1 isopropanol:water gave 0.18 g (yield 45 %) of pure amino acid (d.e.=94%). $^1$H-NMR

($D_2O$) δ: 1.08 (d, 3H, J=7.0, CH-C<u>H</u>$_3$), 2.8-2.9 (m, 1H, $J_1$=4.1, $J_2$=7.0, C<u>H</u>-$CH_3$,), 3.74 (d, 1H, J=4.1, $NH_2$-C<u>H</u>-COOH), 5.20-5.27 (m, 2H, $J_1$=5.9, $J_2$=11.2, CH=C<u>H</u>$_2$), 5.76-5.88 (m, 1H, $J_1$=6.2, $J_2$=11.6, C<u>H</u>=$CH_2$).

The crude SS, RR-isomer did not crystallize. It was dissolved in 6 N HCl; the hydrochloride was isolated by evaporation and dissolved in methanol. The amino acid was precipitated by slow addition of propylene oxide. The precipitate was crystallized from 4:1 isopropanol:water, yielding 0.116 g (29%) of pure SS, RR-E-Ile (d.e.=78%). $^1$H-NMR ($D_2O$) δ: 1.12 (d, 3H, J=7.0, CH-C<u>H</u>$_3$), 2.75-2.85 (m, 1H, $J_1$=6.9, $J_2$=5.8, C<u>H</u>-$CH_3$), 3.57 (d, 1H, J=5.7, $NH_2$-C<u>H</u>-COOH), 5.20-5.27 (m, 2H, $J_1$=5.1, $J_2$=11.1, CH=C<u>H</u>$_2$), 5.76-5.88 (m, 1H, $J_1$=7.3, $J_2$=10.1, C<u>H</u>=$CH_2$).

The reaction products were recrystallized 5 times from 5:1 isopropanol:water to give SS, RR-E-Ile (d.e. = 95%) and SR, RS-E-Ile (d.e. = 98%).

II.3.1.3 <u>2-amino-3-methyl-4-pentynoic acid (Y-Ile)</u>

The synthesis scheme for Y-Ile appears in Figure II-4. Diphenylmethylene glycine ethyl ester was prepared according to O'Donnell and Polt [35]. The $^1$H NMR spectrum was in accord with Aidene and coworkers [36].

To prepare 3-bromobutyne, 6.82 g (62 mol) $PBr_3$ containing 0.1 mg hydroquinone and a solution of 10 g (0.14 mol) 3-butyn-2-ol in 1 mL dry pyridine were added dropwise simultaneously over a 4-hour period to a 50 mL round bottom flask cooled to -15 °C, maintained under argon, and equipped with a stirrer, an addition funnel, and a gas inlet. After an additional 30 min of reaction, 20 mL of cold water were cautiously added to stop the reaction. The mixture was extracted with ether (3 x 20 mL). The combined ether layers were washed consecutively with water (3 x 20 mL), saturated sodium bicarbonate

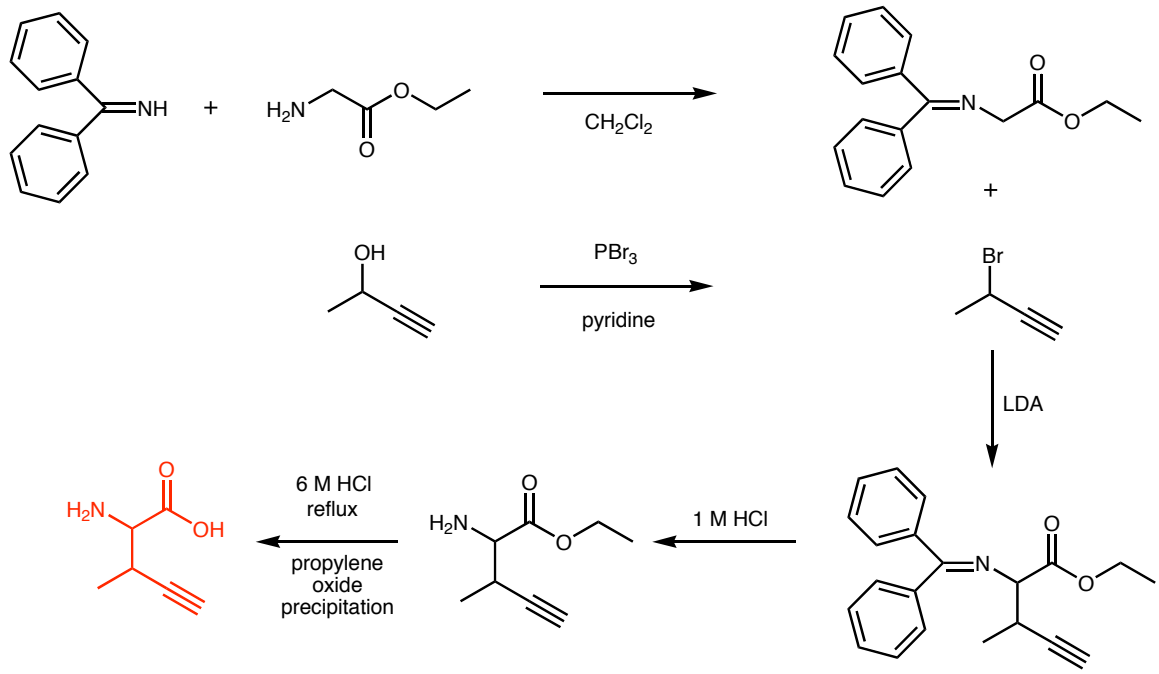


**Figure II-4.** Y-Ile was prepared from diphenylmethylene glycine ethyl ester, which was alkylated with 3-bromobutyne. The alkylated product was deprotected to yield a mixture of stereoisomers of Y-Ile.

(3 x 25 mL), and saturated sodium chloride (2 x 50 mL). The ether extract was dried with $MgSO_4$ and filtered. The dry ether phase was immediately distilled, affording 8.1 g (45%) of 3-bromobutyne. $^1$H-NMR ($CDCl_3$): δ 1.90 (d, 3H, $J$ = 7.1, C<u>H</u>$_3$-CHBr), 2.63 (d, 1H, CHBr-CC<u>H</u>), 4.55 (m, 1H, $J$ = 7.1, $CH_3$-C<u>H</u>Br) ppm.

A solution of 1.16 mL (8.22 mmol ) diisopropylamine in 15 mL dry tetrahydrofuran (THF) was cooled to –20 °C under an argon atmosphere. A 1.6 M solution of *n*-butyl lithium in *n*-hexane (5.13 mL, 8.22 mmol) was added by syringe, and the mixture was stirred for 10 min. The lithium diisopropylamide solution was then cooled to –70 °C, and a solution of 2 g (7.5 mmol) diphenylmethylene glycine ethyl ester in 3 mL THF was slowly added. After 30 min stirring, 1.0 g (7.5 mmol) 3-bromobutyne was added over a 15 min period. The reaction mixture was allowed to warm to room

temperature and stirred for an additional 3 hours. The solvent was removed by evaporation and the product purified by column chromatography (1:1 $CH_2Cl_2$:hexane). The purified product was dissolved in 15 mL diethyl ether, and 15 mL of 1 M HCl was added. The mixture was stirred vigorously for 3 hours at room temperature. The two phases were separated, and the aqueous phase was washed twice with 10 mL diethyl ether. The aqueous layer was concentrated, and the residue was redissolved in 15 mL of 6 M HCl and heated to reflux for 6 hours. The water was removed, and the residue was taken up in 10 mL methanol and 5 mL propylene oxide and stirred for 12 hours at room temperature. The precipitate that formed was filtered and dried, yielding 0.25 g (27 %) of Y-Ile. The $^1$H-NMR spectrum was in agreement with that reported previously [36]. Statistical mixtures of all 4 stereiosmers were obtained; the product was recrystallized four times from 5:1 isopropanol:water to give SS, RR-Y-Ile (d.e. = 81%) and SR, RS-Y-Ile (d.e. = 60%).

### *II.3.2 Purification of IleRS*

*E. coli* strain MV1184 transformed with the multicopy plasmid pkS21[37] encoding the IleRS was kindly provided by Valerie de Crécy-Lagard and Paul Schimmel at the Scripps Research Institute. IleRS was expressed in 1 L (2xYT) cultures, and the collected protein was purified by ammonium sulfate fractionation, cold water dialysis, and anion exchange chromatography following an established protocol [38]. The concentration of IleRS stock was determined by the Bradford method to be 0.13 mM.

### *II.3.3 ATP/$PP_i$ Exchange*

ATP-$PP_i$ exchange assays were performed as previously described [39] in 150 mL of reaction buffer (pH 7.6, 20 mM imidazole, 0.1 mM EDTA, 10 mM β-mercaptoethanol, 7 mM $MgCl_2$, 2 mM ATP, 0.1 mg/ml BSA, and 2 mM $PP_i$ [$^{32}P$ sodium pyrophosphate with a specific activity of 0.1 mCi/ml]) with 75 nM IleRS and concentrations of analogues from 10 mM to 1 mM. Aliquots of 15-20 mL of each reaction were removed at various time points and quenched in 0.5 ml of a solution of 200 mM $PP_i$, 7 % v/v $HClO_4$, and 3% w/v activated charcoal. The charcoal was washed twice with 0.5 ml of a solution of 10 mM $PP_i$ with 0.5% v/v $HClO_4$ and resuspended in 0.5 ml of this solution. Each charcoal suspension was transferred to a 20-mL scintiallation vial, and 10 mL Safety-Solve liquid scintillation cocktail (Research Products Institute, Inc.) was added before counting on a Beckman Coulter liquid scintillation counter.

*II.3.4 Protein expression*

To test for analogue incorporation, a 50 mL culture of M9AA medium supplemented with ampicillin (200 mg/L), chloramphenicol (35 mg/L), 1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 0.2 % glucose, and 1 mg/L thiamine was inoculated with a single colony of the isoleucine auxotrophic expression system AI-IQ[PQE15] [40]. After overnight growth at 37 °C, a 5 μL aliquot of culture was used to inoculate 50 mL of supplemented M9AA medium (for small-scale expressions) or a 1 mL aliquot was used to inoculate 1 L of medium (for large-scale expressions). When the culture reached an $OD_{600}$ of 0.9-1.0, the cells were sedimented (5000*g*, 10 min, 4°C), washed twice with 0.9 % NaCl, and resuspended in 50 mL fresh supplemented M9 medium containing 19 natural amino acids (20 mg/L) but lacking isoleucine. For tests of incorporation, the cultures were divided

into aliquots, to which were added water (negative control), L-isoleucine (positive control), L-valine, SS-E-Ile, SR-E-Ile, SS-Y-Ile, or SR-Y-Ile. Tests of incorporation using once-recrystallized SS-E-Ile (d.e. = 68 %) were performed in 10 mL cultures; the concentrations of L-amino acid in the medium were 25, 50, and 125 mg/L. Studies of SS-E-Ile (d.e. = 95 %) and SR-E-Ile (d.e. = 98 %) were performed at 130 mg/L of the L-amino acid in 5 mL cultures, while studies of SS-Y-Ile (d.e. = 81%) and SR-Y-Ile (d.e. = 60%) were performed at 100 mg/L in 5 mL cultures. For large-scale expressions of protein samples for $^{1}$H-NMR studies, the L-amino acid concentrations were 25 mg/L and 125 mg/L SS-E-Ile (d.e. = 76%) and 70 mg/L SR-E-Ile (d.e. = 91%). After 10 min of growth, mDHFR expression was induced by addition of 1 mM IPTG. After 4 h growth at 37 °C, the cells were sedimented (5000*g*, 10 min, 4°C), resuspended in 4 M urea, and frozen at –20 °C overnight. The cells were thawed, sonicated, and incubated for 30 min at 37 °C with 10 mg/mL DNase, 10 mg/mL RNase, and 10 mM $MgCl_2$ prior to purification. Protein expression was monitored by SDS-PAGE and Western blotting with antibodies recognizing the histidine tag of mDHFR (Qiagen, Inc., Santa Clarita, CA, USA) [41].

### *II.3.5 Mass spectrometry*

Expressed mDHFR was purified from cell lysates by Ni affinity chromatography using a Ni-NTA Spin Kit (Qiagen). Purified mDHFR was incubated with trypsin in 50 mM ammonium bicarbonate buffer overnight at room temperature. The proteolysis product was purified by C18 ZipTip (Millipore), deposited in a matrix of α-cyano-4-

hydroxycinnamic acid, and analyzed by MALDI-TOF mass spectrometry on an Applied Biosystems Voyager DE Pro instrument.

### *II.3.6 $^{1}$H-NMR spectroscopy*

For $^{1}$H-NMR, samples of purified protein were dissolved in 0.1% DCl in $D_2O$. Samples of SS-E-Ile (6 mg/L) were prepared in standard tubes (700 μL volume), while SR-E-Ile samples (4 mg/mL) were prepared in low volumes (300 μL) in solvent-matched tubes (Shigemi, Inc.). Spectra were collected on a Varian Inova NMR spectrometer with proton acquisition at 599.69 MHz and with water suppression by presaturation.

## II.4 Results

### *II.4.1 Synthesis of unsaturated amino acids*

Kazmaier has shown that *N*-benzyloxycarbonylglycine *trans*-crotyl ester undergoes a [3,3]-sigmatropic rearrangement [34] to yield Cbz-protected 2-amino-3-methyl-4-pentenoic acid (E-Ile). The reaction is stereoselective, giving mainly the SR, RS pair of stereoisomers (95%). Because it was of interest to compare the *in vivo* incorporation of the two diastereomers SS-E-Ile and SR-E-Ile, we applied the Kazmaier method also to Cbz-glycine *cis*-crotyl ester to obtain SS, RR-E-Ile (hereafter referred to as SS-E-Ile). The efficiency of rearrangement was lower than for the *trans*-ester, and the work-up requires an additional step because of a lower tendency of the SS, RR-E-Ile to crystallize, which contributed to a lower overall yield for the SS-analogue. The amino

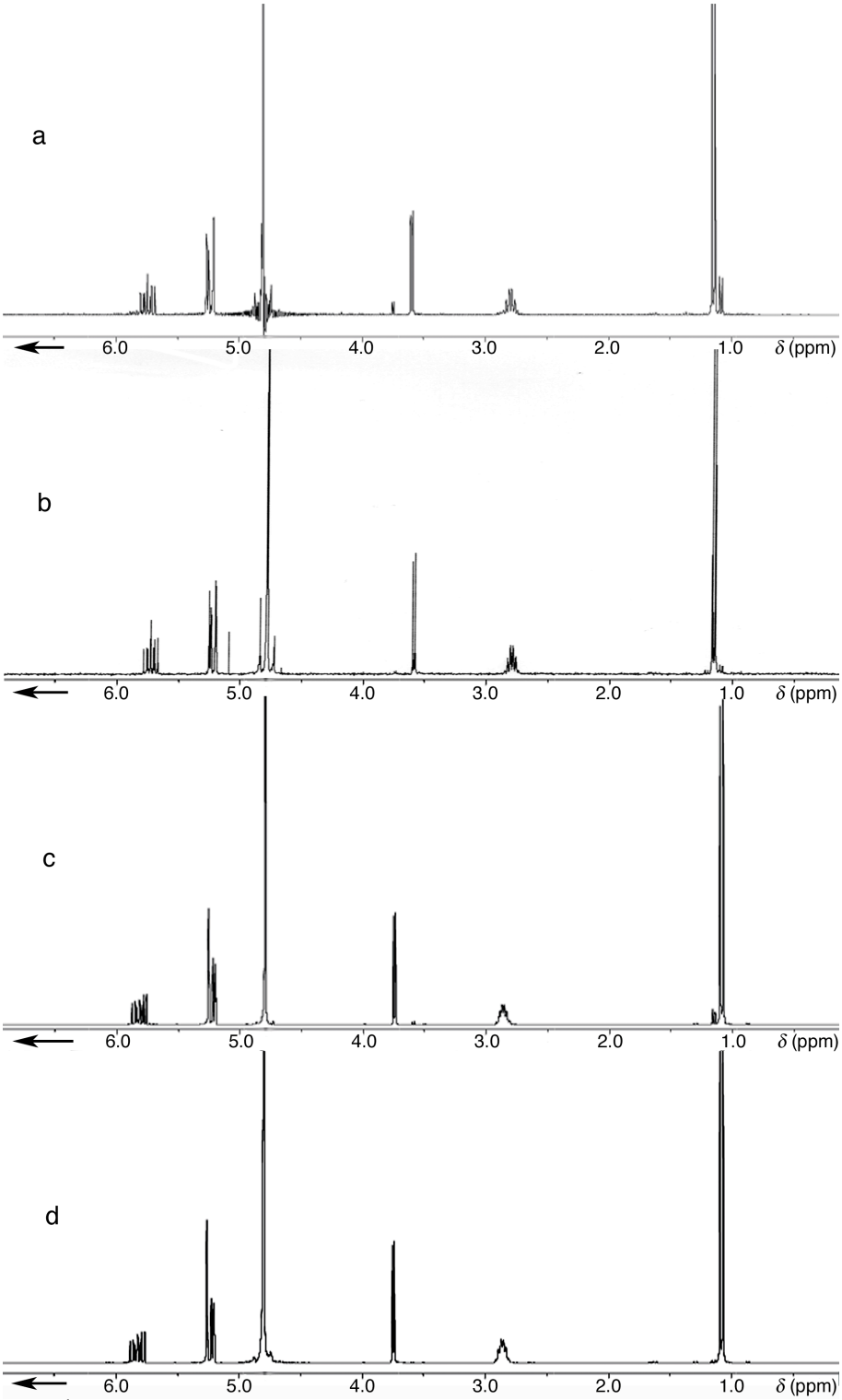


**Figure II-5.** $^{1}$H-NMR spectra of a) SS-E-Ile after 1 recrystallization, b) SS-E-Ile after 6 recrystallizations, c) SR-E-Ile after 1 recrystallization, and d) SR-E-Ile after 6 recrystallizations.

acids were further purified by multiple recrystallizations to give SS-E-Ile (d.e. = 95 %) and SR-E-Ile (d.e. = 98 %) as determined by $^{1}$H-NMR spectroscopy (Figure II-5).

The preparation of 2-amino-3-methyl-4-pentynoic acid was not stereoselective and gave a complex mixture of the SS, RR, SR, and RS stereoisomers. Recrystallization did not succeed in sufficiently separating SS, RR and SR, RS pairs; final products were ss-Y-Ile (d.e. = 81%) and ss-Y-Ile (d.e. = 60%) (Figure II-6).

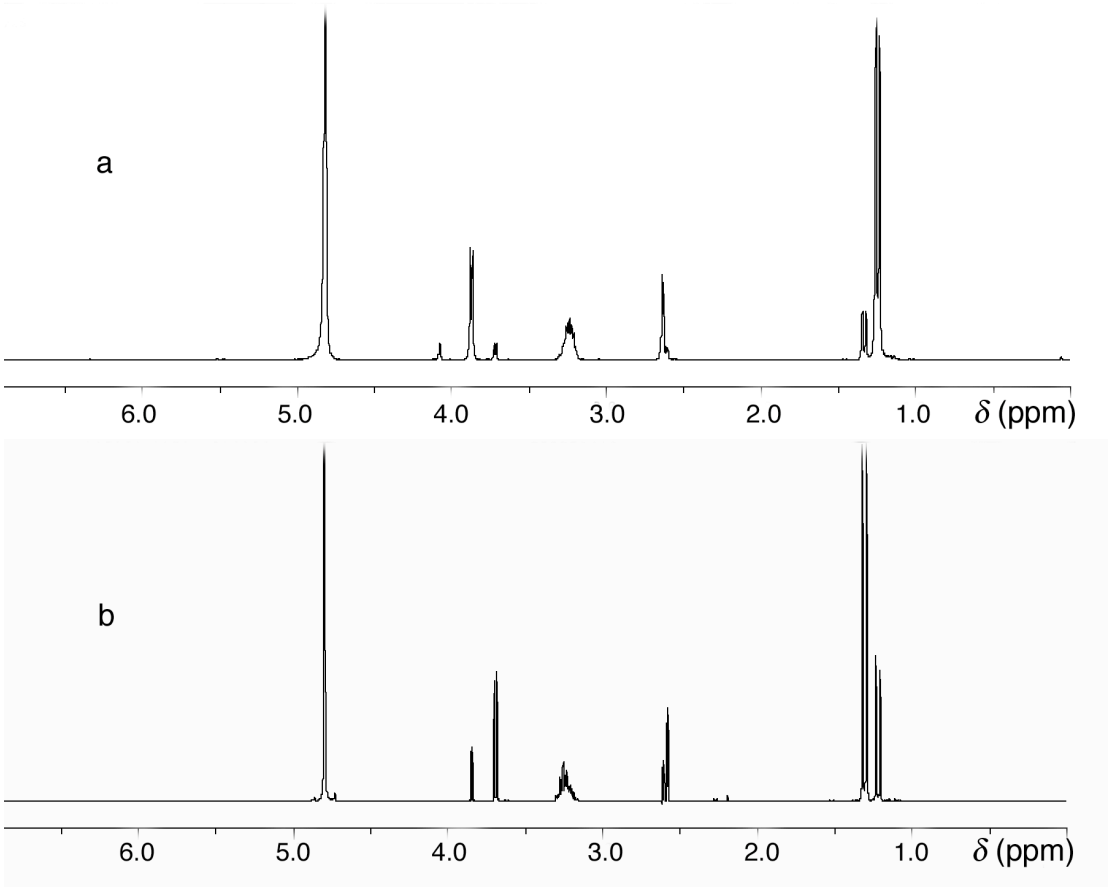


**Figure II-6.** $^{1}$H-NMR spectra of a) SS-Y-Ile (d.e. = 81 %) and b) SR-Y-Ile (d.e. = 60%) after recrystallization.

*II.4.2 Analogue incorporation*

II.4.2.1 SS- and SR-Y-Ile

An *E. coli* strain rendered auxotrophic for isoleucine was used to assay the extent of *in vivo* incorporation of isoleucine analogues into mDHFR, a test protein readily expressed in bacterial cultures. SDS-PAGE and Western blotting of the total cellular protein produced in cultures supplemented with both SS-Y-Ile (d.e. = 81%) and SR-Y-Ile (d.e. = 60%) indicate expression of target protein (Figure II-7).

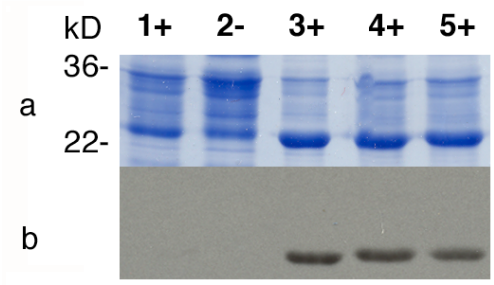


**Figure II-7.** a) SDS-PAGE and b) Western blot of proteins produced in Ile auxotrophic *E. coli* cultures supplemented with the 19 amino acids (lacking Ile) and 1) nothing, or 100 mg/L of 2) Ile, 3) Ile, 4) SR-Y-Ile, or 5) SS-Y-Ile. A + indicates induction of mDHFR expression; expression was not induced in lane 2. A similar amount of mDHFR is produced in cultures supplemented with SS-Y-Ile (d.e. = 81 %) and in those containing SR-Y-Ile (d.e. = 60%).

In the case of both SS- and SR-Y-Ile, MALDI-TOF mass spectra on tryptic fragments of mDHFR produced in medium supplemented with the analogue show peaks with the expected mass difference of -4 for each isoleucine in the peptide (Figure II-8). It is clear that we are able to incorporate at least one alkyne analogue into proteins *in vivo*, but without more careful purification of the stereoisomers it is impossible to say whether both or only one of these amino acids is translationally active.

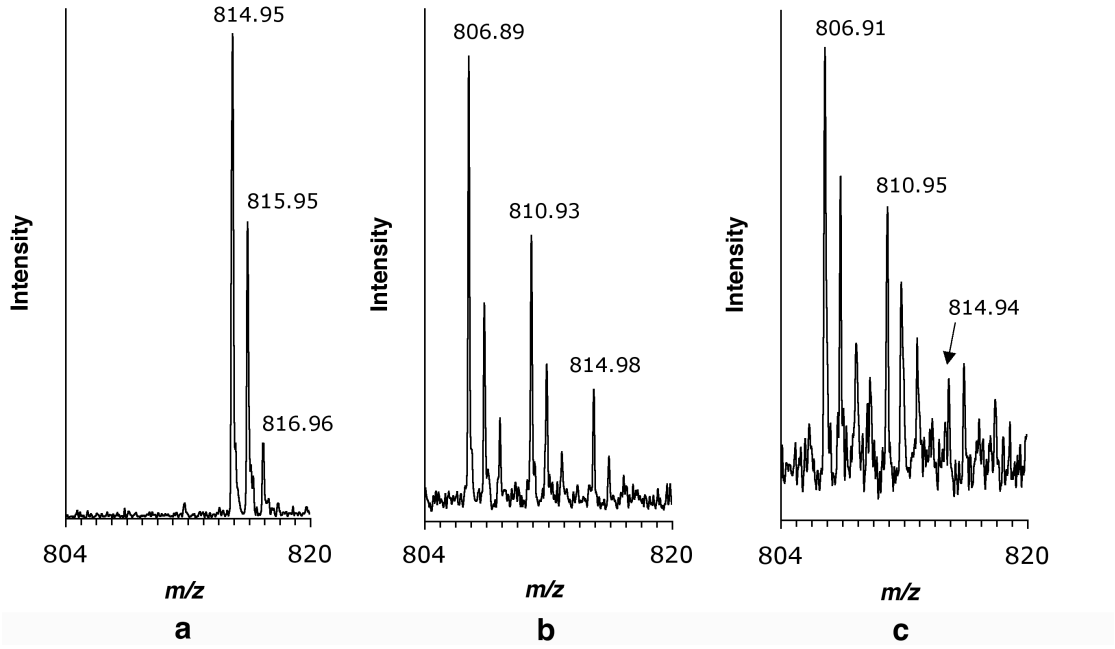


**Figure II-8.** MALDI-TOF spectra of tryptic fragments of mDHFR produced in medium supplemented with a) Ile, b) SS-Y-Ile (d.e. = 81 %), and c) SR-Y-Ile (d.e. = 60 %). Mass shifts of –4 per Ile residue in the peptide INIVLSR (residues 86 – 92, m/z = 814.5) in panels b) and c) indicates incorporation of an alkynyl analogue.

II.4.2.2 SS- and SR-E-Ile

In cultures supplemented with SS-E-Ile (d.e. = 95 %), SDS-PAGE and Western blotting again indicate that target protein is expressed in amounts comparable to the positive control (Ile); in contrast, cultures supplemented with SR-E-Ile (d.e. = 98 %) yielded significantly less protein (Figure II-9). This result is consistent with either a low level of translational activity of SR-E-Ile or with incorporation of residual SS-E-Ile in the amino acid sample. As discussed below, we believe the latter interpretation is correct.

MALDI-TOF mass spectra on tryptic fragments of mDHFR produced in medium supplemented with SS-E-Ile (d.e. = 95%) show signals shifted by the expected mass difference of -2 for each isoleucine residue in the peptide (Figure II-10). The MALDI

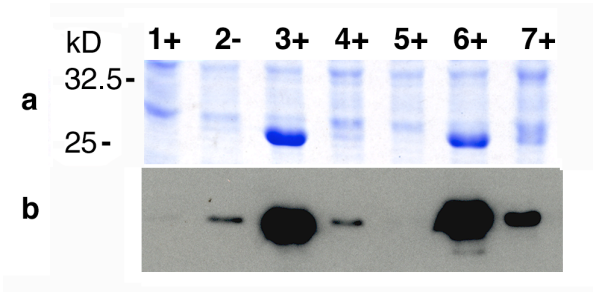


**Figure II-9**. a) SDS-PAGE and b) Western blot of proteins produced in Ile auxotrophic E. coli cultures supplemented with the 19 amino acids other than Ile and 1) nothing, or 130 mg/L of 2) Ile, 3) Ile, 4) norvaline, 5) Val, 6) SS-E-Ile, 7) SR-E-Ile. A + indicates induction of mDHFR expression; expression was not induced in lane 2. Significantly more mDHFR is produced in cultures supplemented with SS-E-Ile (d.e. = 95%) than in those containing SR-E-Ile (d.e. = 98%).

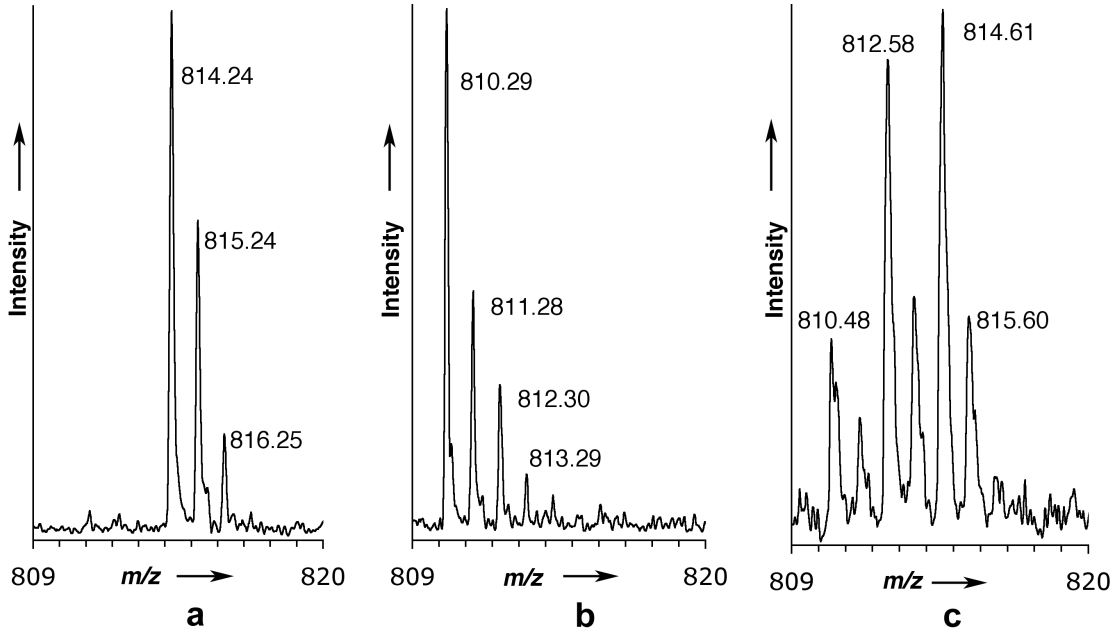


**Figure II-10.** MALDI-TOF spectra of tryptic fragments of mDHFR produced in medium supplemented with a) Ile, b) SS-E-Ile and c) SR-E-Ile. Mass shifts of –2 per Ile residue in the peptide INIVLSR (residues 86 – 92, m/z = 814.5) in panel b) indicate incorporation of SS-E-Ile. A lesser extent of incorporation of an unsaturated amino acid is also evident in proteins produced in culture supplemented with c) SR-E-Ile (d.e. = 98%).

spectra of fragments of target protein expressed in medium containing SR-E-Ile (d.e. = 98%) also show evidence of incorporation of an amino acid with a mass difference of -2. It is apparent that an unsaturated analogue did replace a fraction of isoleucine in each of these proteins, but of course it is not possible to distinguish between incorporation of SS-E-Ile and incorporation of SR-E-Ile from these data.

To determine the identity of the unsaturated amino acid that was incorporated into mDHFR in experiments with the SR-isomer, we purified protein from large-scale expressions conducted in medium supplemented with 70 mg/L SR-E-Ile (d.e. = 91%); protein yield was 4.1 mg/L. The alkene region of the $^1$H-NMR spectrum of this protein (Figure II-11a) is identical to that of a protein expressed in medium supplemented with SS-E-Ile (d.e. = 76%) (Figure II-11b). In the amino acid spectra, the multiplet assigned to the internal alkene proton, $H_i$, is sensitive to stereochemistry; the chemical shift of this

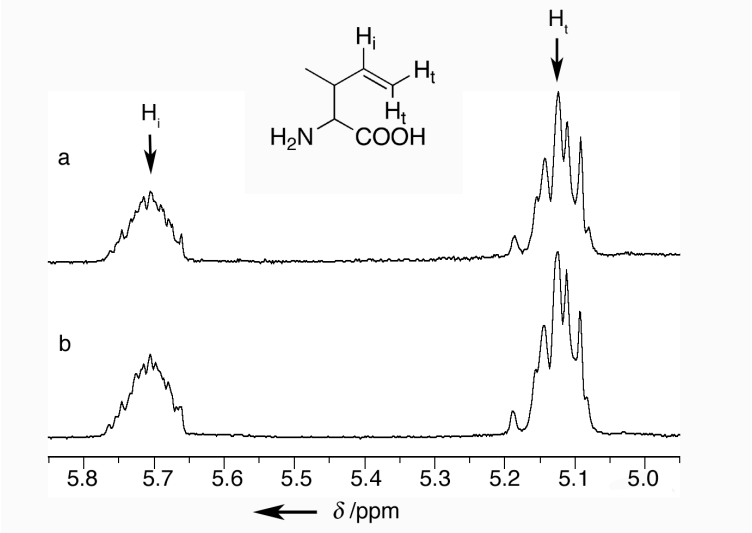


**Figure II-11.** $^1$H-NMR spectra of mDHFR produced in medium supplemented with a) SS-E-Ile (d.e. = 76%) and b) SR-E-Ile (d.e. = 91%). Spectra of mDHFR containing only canonical amino acids show no peaks in this region.

proton differs by 0.1 ppm for the SS- and SR-isomers. We find it unlikely that this resonance would lose its sensitivity to stereochemistry in the protein context. We conclude that SR-E-Ile is a poor substrate for IleRS and that it is unable to compete with residual SS-isomer with respect to *in vivo* incorporation into proteins under the conditions used here.

To determine the extent of incorporation of ss-E-Ile, we prepared mDHFR in cultures supplemented with different levels of the analogue. SDS-PAGE and Western blotting of the total cellular protein produced in cultures supplemented with increasing concentrations of once-recrystallized SS-E-Ile (d.e. = 68 %) indicate increasing levels of target protein expression (Figure II-12). Integration of $^{1}$H-NMR spectra of mDHFR produced in large-scale expressions in medium supplemented with 25 mg/L or 125 mg/L SS-E-Ile (d.e. = 76%) indicate levels of replacement of isoleucine by analogue of 63% and 72%, respectively. Protein yields are 12 and 22 mg/L, respectively.

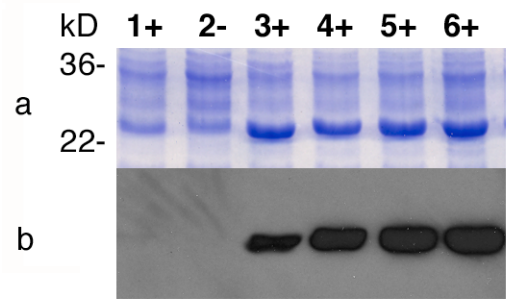


**Figure II-12.** a) SDS-PAGE and b) Western blot of proteins produced in Ile auxotrophic *E. coli* cultures supplemented with the 19 amino acids other than Ile and 1) nothing, 2) 25 mg/L Ile, 3) 25 mg/L Ile, 4) 25 mg/L ss-E-Ile, 5) 50 mg/L ss-E-Ile, and 6) 125 mg/L SS-E-Ile. A + indicates induction of mDHFR expression; expression was not induced in lane 2. mDHFR is produced in cultures supplemented with all concentrations of SS-E-Ile (d.e. = 68 %) in this range.

*II.4.3 Kinetics of Activation of E-Ile by IleRS*

The activation of the analogues by IleRS *in vitro* was investigated by ATP/$PP_i$ exchange. Interestingly, SS-E-Ile was activated as fast as (or somewhat faster than) the natural substrate, Ile ($k_{cat}/K_M$ = 0.30 $\mu M^{-1}$ $s^{-1}$) [42], while SR-E-Ile was activated 40-fold more slowly than Ile (Table II-1). Because of the ~100-fold difference in the rates of activation of the SS- and SR-isomers, it not surprising that residual SS-E-Ile in the sample of the SR-isomer is responsible for the small amount of mDHFR produced in medium supplemented with SR-E-Ile (d.e. = 98 %).

**Table II-1. Kinetics of Activation E-Ile by IleRS**

| analogue | $V_{max}$ (M $s^{-1}$) | $K_M$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_M$ ($\mu M^{-1}$ $s^{-1}$) | Relative to Ile |
|---|---|---|---|---|---|
| SS-E-Ile | 2.04 x $10^{-7}$ | 32.9 ± 13 | 20.35 ± 0.23 | 0.756 ± 0.32 | 2.5 |
| SR-E-Ile | 4.54 x $10^{-8}$ | 432 ± 101 | 4.53 ± 0.51 | 0.026 ± 0.0013 | 1/40 |
| Ile | 2.5 x $10^{-7}$ | 58.2 ± 7.6 | 17.5 ± 4.5 | 0.3 ± 0.14 | 1 |

## II.5 Discussion

Incorporation of amino acids analogues into proteins *in vivo* requires that the analogues (i) cross the cellular membrane; (ii) be charged to one or more tRNA(s); (iii) and be delivered to the growing end of the polypeptide chain. Our work indicates that none of these steps precludes efficient incorporation of at least one stereoisomer of both 2-amino-3-methyl-4-pentynoic acid (Y-Ile) and 2-amino-3-methyl-4-pentenoic acid (E-Ile). In the case of the alkene analogue, we show that the SS-isomer is greatly preferred

over the SR-isomer by the natural biosynthetic machinery of *E. coli*; indeed, we find no conclusive evidence of translational activity for SR-E-Ile.

Following cellular uptake and activation, the analogue must circumvent the editing pathways that normally limit misacylation of tRNAs. The selectivity (s) of an aaRS toward an amino acid is defined as the ratio of the rate of editing to the rate of activation [21]. The editing mechanism of *E. coli* isoleucyl-tRNA synthetase (IleRS) has been extensively studied [43, 44], and its selectivity for natural amino acids is high, ranging from s = 6000 for valine to s = $8.5 \times 10^6$ for alanine [21]. IleRS possesses two active sites: a synthetic site for binding of the amino acid prior to activation through formation of the aminoacyl adenylate and an editing site for removal of amino acids smaller than isoleucine (which fit into the editing pocket) [22, 23]. The SS-analogues tested in this study appears to circumvent the editing mechanism of IleRS, possibly because they are too large to fit into the editing site.

Our results show that IleRS is sensitive to stereochemistry at the β-carbon of E-Ile; only the SS-isomer of the isoleucine analogue is incorporated into protein at a measurable rate, and it is activated by the IleRS ~100-fold more rapidly than SR-E-Ile. This result is in agreement with previous binding studies that demonstrated that L-2-amino-3S-methylhexanoic acid binds preferentially to IleRS ($K_a$ = 20 $mM^{-1}$); its diastereomer L-2-amino-3R-methylhexanoic acid binds to the enzyme with much a lower affinity ($K_a$ = 0.6 $mM^{-1}$) [45]. It is also consistent with the fact that IleRS distinguishes L-isoleucine from L-allo-isoleucine [25, 27, 46]. Our data do not preclude the possibility, however, that discrimination between the *SS*- and *SR*-analogues occurs not in

the synthetic active site of IleRS but rather during some other translational step, such as editing by IleRS or binding to elongation factor-Tu or the ribosome.

Finally, the efficiency of substitution of SS-E-Ile and Y-Ile for Ile in recombinant proteins provides a simple and useful method for the incorporation of terminal double and triple bonds into proteins, giving the chemist access to versatile functional groups in proteins and protein-based materials.

## II.6 References


1. van Hest JCM, Tirrell DA: **Protein-based materials, toward a new level of structural control**. *Chemical Communications* 2001:1897-1904.

2. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR: **Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation**. *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**:19-24.

3. Bae JH, Rubini M, Jung G, Wiegand G, Seifert MHJ, Azim MK, Kim JS, Zumbusch A, Holak TA, Moroder L, Huber R, Budisa N: **Expansion of the genetic code enables design of a novel "gold" class of green fluorescent proteins**. *Journal of Molecular Biology* 2003, **328**:1071-1081.

4. Budisa N: **Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire**. *Angewandte Chemie-International Edition* 2004, **43**:6426-6463.

5. Bentin T, Hamzavi R, Salomonsson J, Roy H, Ibba M, Nielsen PE: **Photoreactive bicyclic amino acids as substrates for mutant Escherichia coli phenylalanyl-tRNA synthetases**. *Journal of Biological Chemistry* 2004, **279**:19839-19845.

6. Furter R: **Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in Escherichia coli**. *Protein Science* 1998, **7**:419-426.

7. Hendrickson TL, de Crecy-Lagard V, Schimmel P: **Incorporation of nonnatural amino acids into proteins**. *Annual Review of Biochemistry* 2004, **73**:147-176.

8. Kiick KL, Weberskirch R, Tirrell DA: **Identification of an expanded set of translationally active methionine analogues in Escherichia coli**. *Febs Letters* 2001, **502**:25-30.

9. Kim WY, George A, Evans M, Conticello VP: **Cotranslational incorporation of a structurally diverse series of proline analogues in an Escherichia coli expression system**. *Chembiochem* 2004, **5**:928-936.

10. Link AJ, Mock ML, Tirrell DA: **Non-canonical amino acids in protein engineering**. *Current Opinion in Biotechnology* 2003, **14**:603-609.

11. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the genetic code of Escherichia coli**. *Science* 2001, **292**:498-500.

12. Wang L, Schultz PG: **Expanding the genetic code**. *Angewandte Chemie-International Edition* 2005, **44**:34-66.

13. Doring V, Mootz HD, Nangle LA, Hendrickson TL, de Crecy-Lagard V, Schimmel P, Marliere P: **Enlarging the amino acid set of Escherichia coli by infiltration of the valine coding pathway**. *Science* 2001, **292**:501-504.

14. Ibba M, Hennecke H: **Relaxing the substrate-specificity of an aminoacyl-transfer-rna synthetase allows in-vitro and in-vivo synthesis of proteins containing unnatural amino-acids**. *Febs Letters* 1995, **364**:272-275.

15. Datta D, Wang P, Carrico IS, Mayo SL, Tirrell DA: **A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins**. *Journal of the American Chemical Society* 2002, **124**:5652-5653.

16. Sharma N, Furter R, Kast P, Tirrell DA: **Efficient introduction of aryl bromide functionality into proteins in vivo**. *Febs Letters* 2000, **467**:37-40.

17. Tang Y, Tirrell DA: **Attenuation of the editing activity of the Escherichia coli Leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo**. *Biochemistry* 2002, **41**:10635-10645.

18. Pauling L: **The Probability of Errors in the Process of Synthesis of Protein Molecules**. In *Festschrift Prof. Dr. Arthur Stoll*. Birkhauser Verlag; 1957:597-602.

19. Loftfield RB: **Frequency of Errors in Protein Biosynthesis**. *Biochemical Journal* 1963, **89**:82-&.

20. Jakubowski H, Fersht AR: **Alternative Pathways for Editing Non-Cognate Amino-Acids by Aminoacyl-Transfer Rna-Synthetases**. *Nucleic Acids Research* 1981, **9**:3105-3117.

21. Jakubowski H, Goldman E: **Editing of Errors in Selection of Amino Acids for Protein Synthesis**. *Microbiological Reviews* 1992, **56**:412-429.

22. Nureki O, Vassylyev DG, Tateno M, Shimada A, Nakama T, Fukai S, Konno M, Hendrickson TL, Schimmel P, Yokoyama S: **Enzyme structure with two catalytic sites for double-sieve selection of substrate**. *Science* 1998, **280**:578-582.

23. Silvian LF, Wang JM, Steitz TA: **Insights into editing from an Ile-tRNA synthetase structure with tRNA(Ile) and mupirocin**. *Science* 1999, **285**:1074-1077.

24. Schmidt E, Schimmel P: **Mutational Isolation of a Sieve for Editing in a Transfer-Rna Synthetase**. *Science* 1994, **264**:265-267.

25. Flossdorf J, Kula MR: **Ultracentrifuge Studies on Binding of Aliphatic Amino-Acids to Isoleucyl-Transfer-Rna Synthetase from Escherichia-Coli Mre 600**. *European Journal of Biochemistry* 1973, **36**:534-540.

26. Polet H, Conrad ME: **Influence of 3 Analogs of Isoleucine on in Vitro Growth and Protein Synthesis of Erythrocytic Forms of Plasmodium Knowlesi**. *Proceedings of the Society for Experimental Biology and Medicine* 1969, **130**:581-&.

27. Loftfield RB, Eigner EA: **Specificity of Enzymic Reactions . Aminoacyl-Soluble Rna Ligases**. *Biochimica Et Biophysica Acta* 1966, **130**:426-&.

28. Parker ED, Shive W, Skinner CG: **Biological Specificities of 4,5-Dehydro Analogues of Isoleucine and Alloisoleucine**. *Journal of Biological Chemistry* 1961, **236**:3267.

29. Clark TD, Ghadiri MR: **Supramolecular Design by Covalent Capture - Design of a Peptide Cylinder Via Hydrogen-Bond-Promoted Intermolecular Olefin Metathesis**. *Journal of the American Chemical Society* 1995, **117**:12364-12365.

30. Schafmeister CE, Po J, Verdine GL: **An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides**. *Journal of the American Chemical Society* 2000, **122**:5891-5892.

31. Nguyen ST, Grubbs RH, Ziller JW: **Syntheses and Activities of New Single-Component, Ruthenium- Based Olefin Metathesis Catalysts**. *Journal of the American Chemical Society* 1993, **115**:9858-9859.

32. Link AJ, Tirrell DA: **Cell surface labeling of Escherichia coli via copper(I)-catalyzed [3+2] cycloaddition**. *Journal of the American Chemical Society* 2003, **125**:11164-11165.

33. Hassner A, Alexanian V: **Synthetic Methods .12. Direct Room-Temperature Esterification of Carboxylic-Acids**. *Tetrahedron Letters* 1978:4475-4478.

34. Kazmaier U: **Synthesis of Unsaturated Amino-Acids by 3,3 -Sigmatropic Rearrangement of Chelate-Bridged Glycine Ester Enolates**. *Angewandte Chemie-International Edition in English* 1994, **33**:998-999.

35. Odonnell MJ, Polt RL: **A Mild and Efficient Route to Schiff-Base Derivatives of Amino-Acids**. *Journal of Organic Chemistry* 1982, **47**:2663-2666.

36. Aidene M, Barbot F, Miginiac L: **Reactivity of alpha-unsaturated organozinc compounds with N- (phenylsulfanyl)iminoesters - Application to the synthesis of monosubstituted or disubstituted unsaturated alpha-aminoacids**. *Journal of Organometallic Chemistry* 1997, **534**:117-127.

37. Shiba K, Schimmel P: **Functional Assembly of a Randomly Cleaved Protein**. *Proceedings of the National Academy of Sciences of the United States of America* 1992, **89**:1880-1884.

38. Shepard A, Shiba K, Schimmel P: **Rna-Binding Determinant in Some Class-I Transfer-Rna Synthetases Identified by Alignment-Guided Mutagenesis**. *Proceedings of the National Academy of Sciences of the United States of America* 1992, **89**:9964-9968.

39. Lawrence F, Blanquet S, Poiret M, Robertge.M, Waller J: **Mechanism of Action of Methionyl-Transfer Rna-Synthetase .3. Ion Requirements and Kinetic Parameters of Atp-Ppi Exchange and Methionine-Transfer Reactions Catalyzed by Native and Trypsin-Modified Enzymes**. *European Journal of Biochemistry* 1973, **36**:234-243.

40. Wang P, Tang Y, Tirrell DA: **Incorporation of trifluoroisoleucine into proteins in vivo**. *Journal of the American Chemical Society* 2003, **125**:6900-6906.

41. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual* edn 2nd. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory; 1989.

42. While the $k_{cat}$ value measured for Ile (17 $s^{-1}$) is consistent with those reported in the literature [20], the $K_M$ value (58 μM) was higher than reported values (4 – 5 μM) [20, 24].

43. Fersht A: *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. New York: W.H. Freeman & Company; 1999.

44. Fersht AR, Dingwall C: **Evidence for the Double-Sieve Editing Mechanism in Protein- Synthesis - Steric Exclusion of Isoleucine by Valyl-Transfer Rna-Synthetases**. *Biochemistry* 1979, **18**:2627-2631.

45. Flossdorf J, Pratorius HJ, Kula MR: **Influence of Side-Chain Structure of Aliphatic Amino-Acids on Binding to Isoleucyl-Transfer-Rna Synthetase from Escherichia- Coli Mre 600**. *European Journal of Biochemistry* 1976, **66**:147-155.

46. Loftfield TB, Hecht LI, Eigner EA: **Measurement of Amino Acid Specificity of Transfer Ribonucleic Acid**. *Biochimica Et Biophysica Acta* 1963, **72**:383-&.