The Demonstration of Selective Peptide Bond Formation in Clear Aqueous Solutions

Darshan Ranganathan* and Girij Pal Singh

Department of Chemistry, Indian Institute of Technology, Kanpur 208 016, India

The distribution of dipeptides upon treatment of equimolar amounts of glutamic acid, leucine, and $C_6H_{11}N{=}C{=}N(CH_2)_3{}^+NMe_3{}^-OSO_2C_6H_4Me$-*p*, in clear aqueous solution followed by *N,C*-protection, was found to be: Bz-Glu(γ-OMe)-Leu-OMe (72%), Bz-Glu(α-OMe)-Glu-di-OMe (15%), and Bz-Leu-Leu-OMe (8%); Bz-Leu-Glu-di-OMe, Bz-Glu(α-OMe)-Leu-OMe, and Bz-Glu(γ-OMe)Glu-di-OMe could not be detected (Bz = PhCO).

Recently, we have reported on the selectivity in peptide bond formation in a reverse micellar system harbouring water pools and using dioctadecylcarbodiimide as a co-surfactant as well as a condensing agent.[1] We present here a study of peptide bond formation with L-glutamic acid and L-leucine mediated by 1-cyclohexyl-3-(3-dimethylaminopropyl)carbodiimide metho-toluene-*p*-sulphonate (**1**),[2] in clear aqueous solutions.

Aqueous solutions of glutamic acid (2 mmol; 42 ml), leucine (2 mmol; 15 ml), and the carbodiimide (**1**) (2 mmol; 5 ml) were mixed and stirred for 48 h at 30 ± 2 °C. The pH of the solution, monitored at regular intervals, remained unchanged at 3 ± 0.1. The reaction mixture was treated with BzCl–2 M NaOH (Bz = PhCO), acidified with 1 M H_2SO_4, and extracted with EtOAc, and the extract dried, evaporated, dissolved in MeOH, and treated with ethereal diazomethane to afford 0.436 g of product consisting of *N,C*-protected dipeptides, Glu, and Leu (TLC).† Since the focus of the experiment was to determine preferences, if any, in peptide bond formation involving the polar amino acid Glu and the hydrophobic amino acid Leu, it was considered important to analyse the above product using all the six possible dipeptides as standards,‡ preparative details of which are presented in Table 1. An analytical HPLC–TLC complementary system was developed having capability for distinguishing all the possibilities, in addition to recognising higher peptides and *N,C*-protected Glu and Leu.§ The composition of the product, as determined by HPLC, was found to be Bz-Glu(γ-OMe)-Leu-OMe (72%), Bz-Glu(α-OMe)-Glu-di-OMe (15%), Bz-Leu-Leu-OMe (8%), higher peptides (1.5%), and

† The clarity of the NMR spectrum of the crude product in the high-field region precluded admixture with products arising from the reagent (**1**) to any significant extent.

‡ Satisfactory spectroscopic data and elemental analyses were obtained for compounds (**2**)—(**7**). Glu(α-OMe)-OH, needed for the synthesis of (**3**) and (**6**), was prepared in 65% yield by a novel pathway involving the Ru^{VIII} oxidation (S. Ranganathan, D. Ranganathan, and D. Bhattacharyya, *J. Chem. Soc., Chem. Commun.*, 1987, 1085) of Bz-Pro-OMe followed by the opening of the intermediate Bz-Pyroglu-OMe, with aqueous $NaHCO_3$. The dipeptide (**3**) obtained from this sample was found to be identical to that prepared by methylation with dimethyl sulphate of Bz-Glu (G. H. L. Nefkens and R. J. F. Nivard, *Recl. Trav. Chim. Pays-Bas*, 1964, **83**, 199), peptide formation with Glu-di-OMe, and separation by preparative TLC.

§ HPLC analyses were carried out using MeOH–H_2O (80:20) as solvent on a reverse phase column. The reaction mixture and the reference samples were analysed on the same day to avoid column aberrations. TLC (C_6H_6–EtOAc, 7:3) clearly distinguished between Bz-Leu-OMe, Bz-Glu-di-OMe, Bz-Glu(γ-OMe)-Leu-OMe (**2**), and Bz-Glu(α-OMe)-Leu-OMe (**6**). The latter two could also be distinguished by NMR spectroscopy.

Table 1. Synthesis of authentic dipeptides expected from the Glu, Leu combination.

$$Bz\text{-}aa_1 + aa_2\text{-}OMe \xrightarrow{a} Bz\text{-}aa_1\text{-}aa_2\text{-}OMe \ (aa = \text{amino acid})$$

		Bz-aa$_1$-aa$_2$-OMe			
aa$_1$	aa$_2$	No.	% Yield	M.p., *t*/°C	$[\alpha]_D^{30}$ ($CHCl_3$)/°
Glu(γ-OMe)	Leu	(**2**)	53	104—106	−5.3 (*c* 3.3)
Glu(α-OMe)	Glu(γ-OMe)	(**3**)	45	Gum	—
Leu	Leu	(**4**)	39	199—201	−40.3 (*c* 0.46)
Leu	Glu(γ-OMe)	(**5**)	48	102—104	+10.7 (*c* 1.9)
Glu(α-OMe)	Leu	(**6**)	61	122—124	−10.0 (*c* 3.3)
Glu(γ-OMe)	Glu(γ-OMe)	(**7**)	68	119—120	+2.6 (*c* 0.3)

[a] *Reagents*: dicyclohexylcarbodiimide and hydroxybenzotriazole in CH_2Cl_2.

negligible amounts of other compounds (total: 3.5%). Thus, on the basis of HPLC analysis, the yield of the major product, Bz-Glu(γ-OMe)-Leu-OMe is 40%, which agrees quite well with the 32% yield that was isolated on preparative TLC of a portion of the product.¶

$C_6H_{11}N{=}C{=}N(CH_2)_3\overset{+}{N}Me_3{}^{-}OSO_2C_6H_4Me\text{-}p$
(**1**)

Bz-Glu(γ-OMe)-Leu-OMe
(**2**)

Bz-Glu(α-OMe)-Glu-di-OMe
(**3**)

Bz-Leu-Leu-OMe
(**4**)

Bz-Leu-Glu-di-OMe
(**5**)

Bz-Glu(α-OMe)-Leu-OMe
(**6**)

Bz-Glu(γ-OMe)-Glu-di-OMe
(**7**)

Bz = PhCO

The overwhelming preference for the formation of the Glu-(γ-OH)-Leu peptide bond, involving the more hindered α-carboxy function of Glu, could not easily have been predicted. The observed preference could be rationalized on the basis of interaction of the most acidic α-carboxy unit of Glu with the carbodiimide, a process that could be assisted by protonation of the reagent by the $^+NH_3$ group present, followed by opening of the resulting activated ester with the more basic amino function of leucine.

Regardless of mechanistic details, the experimental findings reported here support a preference in the choice of neighbours in peptide bond formation in water, which, when extrapolated, is in agreement with the notion that specific short sequences of peptides could have existed in the very early stages of protein evolution.‖

We thank Professor S. Ranganathan for valuable advice, Dr. Sircar, R.S.I.C., C.D.R.I., Lucknow for HPLC analyses, and U.G.C. and D.S.T., New Delhi, for financial assistance.

Received, 16th August 1989; Com. 9/03833G

References

1 D. Ranganathan, G. P. Singh, and S. Ranganathan, *J. Am. Chem. Soc.*, 1989, **111**, 1144.
2 J. C. Sheehan, P. A. Cruickshank, and G. Boshart, *J. Org. Chem.*, 1961, **26**, 2525.

¶ This represents the best yield. The experiment has been performed several times. Whilst the HPLC profile remained practically unaltered in every case, the yield showed variation.

‖ Support for such a view has come from early results of a computer based analysis of a number of functional protein sequences, the salient features of which are: (i) sequence analysis of 50 functional proteins shows that the neighbour slots are filled on the basis of a pattern that diverges widely from the estimated non-preference value; (ii) sequence analysis of 150 functional proteins shows that whilst the Ser–Ser sequence is quite rare, it is frequent in selected classes of proteins, such as carbonic anhydrase, proteases (trypsin, pepsin, subtilisin, carboxypeptidase, elastase), trypsin inhibitors, ribonuclease, and deoxyribonuclease; (iii) analysis for persistent occurrence of dipeptide elements across an evolutionary span covering *Escherichia coli* to man in the case of cytochrome shows that the Lys–Lys unit is nearly uniformly maintained (G. P. Singh, PhD. Thesis, I.I.T.K., 1988).