ELSEVIER

FEBS Letters 337 (1994) 4–8

FEBS 13492

FEBS LETTERS

Hypothesis

Conformational rearrangements required of the V_3 loop of HIV-1 gp120 for proteolytic cleavage and infection

Michael E. Johnson[a,*], Zhaolan Lin[a], Kaillathe Padmanabhan[b], Alexander Tulinsky[b], Michael Kahn[c]

[a]*Center for Pharmaceutical Biotechnology and Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, 833 South Wood Street (mlc 781), Chicago, IL 60612-7230, USA*

[b]*Department of Chemistry, Michigan State University, East Lansing, MI 48824-1322, USA*

[c]*Department of Pathobiology, University of Washington, Seattle, WA 98195, USA*

Received 1 October 1993; revised version received 19 November 1993

Abstract

HIV gp120 is specifically cleaved at a single site in the V_3 loop between Arg^{315} and Ala^{316} by thrombin. Previous observations by others have indicated that binding to CD4 enhances the rate of V_3 loop cleavage, and that this cleavage is a prerequisite for HIV infection. Other observations also suggest that the cleavage site is in a type II β-turn centered at Pro^{313}–Gly^{314}. However, our docking experiments indicate that this conformation cannot dock to thrombin and other trypsin-like serine proteases. Thus, based on the thrombin-bound conformation of peptide substrates, we propose that CD4 binding, at a site remote from the V_3 loop, induces and stabilizes a *trans* to *cis* isomerization of the highly conserved residue Pro^{313}, and that this conformational shift is a prerequisite for cleavage by a 'thrombin-like' cellular protease and subsequent infection.

Key words: AIDS; Thrombin recognition motif; β-Turn, Molecular modeling; Factor Xa

1. Introduction

The human immunodeficiency virus infects cells following the binding of the viral envelope glycoprotein gp120 to the cell surface receptor CD4 [1,2]. However, the binding and infective processes can be separated [3,4]. Antibodies to the principal neutralizing determinant in the V_3 loop prevent infection, but do not inhibit binding to CD4 [5–7]. It has also been suggested that proteolytic cleavage of this loop, probably by a trypsin-like serine protease(s) at the cell surface, is a prerequisite to viral infection [8], and that cleavage is enhanced by sCD4 binding and inhibited by neutralizing antibodies to the V_3 loop [9].

Gp120 is cleaved specifically at a single site in the V_3 loop between R^{315} and A^{316} by thrombin [10]. Thrombin cleavage of gp120 is enhanced by sCD4 binding and is abrogated by transient exposure to nonionic detergent [9,11], underlining the importance of substrate conformation. Hattori and coworkers reported that trypstatin, a Kunitz-type proteinase inhibitor of coagulation factor Xa, prevented infection [8]. A strong homology exists between the active site inhibitory region of trypstatin and the highly conserved region of the gp120 V_3 loop [8], and certain structural features are conserved despite hypervariability in the loop. The GPGR sequence from residues 312–315 is found in 85–90% of HIV isolates [12]. Additionally, Gotoh et al. have shown that factor Xa from chick embryo is a critical determinant in paramyxovirus activation in chick embryos [13,14]. Mutations at G^{312}, G^{314} or R^{315} produce non-infectious virions, thus highlighting the importance of this region, although the P^{313} to A^{313} mutation did not prevent infection [15]. Secondary structure analysis [12], solution NMR [16], and X-ray crystallographic studies [17] suggest that this sequence, immediately preceding the thrombin cleavage site, forms a type II β-turn, with Pro^{313} at the $i+1$ position, and with the adjacent segments forming antiparallel β-strands [12,16].

Thrombin exhibits substantial secondary structure specificity for a β-sheet-turn-sheet motif, with the reverse turn appositely orienting substrates for binding to thrombin [18–20]. In this work, we have compared the predicted secondary structure for the consensus sequence ...RKSIHI---GPGR*AFY [12] and other high probability sequences for the principal neutralizing determinant (PND) region of the gp120 V_3 loop with the thrombin-

*Corresponding author. Fax: (1) (312) 413-8084.

SSDI 0014-5793(93)E1409-F

bound conformation of fibrinopeptide [21,22], and offer a hypothesis and working model for examination of the CD4 induced conformational changes involved in the requisite proteolysis of the gp120 V_3 loop prior to fusion. Our results suggest that the enhancement of the cleavage rate observed upon CD4 binding to gp120 results from CD4-induced *trans/cis* isomerization and stabilization of the *cis* Pro^{313} configuration. In the β-turn $i + 2$ position, the *cis* form of Pro^{313} would be compatible with a type VI reverse turn at the tip of the V_3 loop; this conformational shift is required to permit cleavage by thrombin, factor Xa or a related thrombin-like cellular protease.

2. Materials and methods

Models for the bound structures of the V_3 loop peptides from gp120 were constructed using the InsightII modeling program (Biosym Technologies Inc.), using BPTI and the bound structure of fibrinopeptide A within the thrombin binding site [19–22] as a template. Thrombin coordinates were obtained from the crystal structure of the hirudin–thrombin complex [18]. Factor Xa coordinates were those of a recently solved structure at 2.2 Å resolution [23]. Criteria for construction of the models included the requirements that: (i) the Arg residue side chain at the P1 cleavage site was fixed within its S1 subsite, (ii) the ϕ,φ angles for the gp120 substrates were similar to the P2 to P1 ϕ,φ angles for BPTI, (iii) key hydrogen bonds and hydrophobic interactions critical for substrate binding in the thrombin [21,22] and factor Xa active sites were maintained, and (iv) the insertions and deletions in the various HIV gp120 isolates could be accommodated. Following manual construction of various models, the structures were energy minimized, and examined to ensure that ϕ,φ dihedral angles were within allowed regions.

3. Results

Construction of a type II turn with Pro at the $i + 1$ position, and placement within the thrombin active site is shown in Fig. 1, from which it can be seen that the Pro residue exhibits substantial clashing interactions with thrombin His^{57}, and with the insertion loop $Tyr^{60}A$ and $Trp^{60}D$ residues. The active site of factor Xa exhibits similar collisions, also shown in Fig. 1. Attempts to reorient the positioning of the peptide to remove these unfavorable interactions, while maintaining the type II β-turn with Pro in the $i + 1$ position, were uniformly unsuccessful. Although secondary structure analysis [12], NMR-based conformational analysis [16], and X-ray crystallographic analysis of a V3 loop peptide antigen with an HIV-1 neutralizing antibody [17] would suggest that the gp120 peptide should exhibit a type II β-turn centered at Pro^{313} and Gly^{314}, it is impossible to dock this conformation in a productive mode within either the thrombin or factor Xa active site regions.

Following our previous work with fibrinopeptide A [20], we have used the P^2 to P^1 residues of the fibrinopeptide-thrombin crystal structure [21] as a template for setting the ϕ,φ angles of the P^{313} to R^{315} residues, while using the *cis*-Pro Type VI turn in the REI protein [24] as a template for the turn. This then places the P^{313} residue in a *cis* configuration in the $i + 2$ position of a type VI turn. This turn motif then permits a favorable interaction of the Ile^{311} residue with the thrombin hydrophobic cluster, a secondary binding interaction that has been shown to contribute greatly to the conformational specificity of thrombin cleavage [19,20,25–27]. The overlay of our proposed gp120 consensus sequence conformation with the thrombin active site is shown in Fig. 2, while Fig. 3 shows the proposed gp120 conformation within the factor Xa binding site, again showing very reasonable docking. An analogous proposed gp120 conformation, with an Arg^{311} insertion in the sequence (the most common insertion [12]) can also be readily accommodated within this basic structural motif, with the hydrocarbon chain moiety of the R side chain in contact with the thrombin hydrophobic cluster, and the guanidinium group making a salt bridge with the thrombin Glu^{192} side chain carboxyl group (not shown).

In the proposed gp120 conformations, the Pro^{313} to Arg^{315} segment is juxtaposed with the thrombin segment Ser^{214}–Gly^{216}, forming an antiparallel β-pleated sheet. The Pro^{313} carbonyl oxygen and the Arg^{315} amide are hydrogen-bonded to the Gly^{216} amide and the Ser^{214} car-

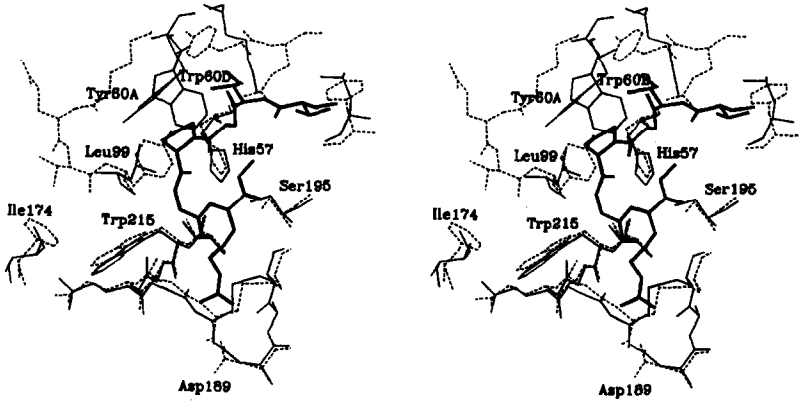

Fig. 1. Stereo-diagram showing docking of the gp120 V_3 loop primary consensus sequence with Pro^{313} in the $i + 1$ turn position (▬) within the thrombin (——) and factor Xa (·····) binding sites. Note the collision of the gp120 residues with those of thrombin and factor Xa, making it impossible to productively dock this gp120 model structure within the active site of either enzyme in this conformation.

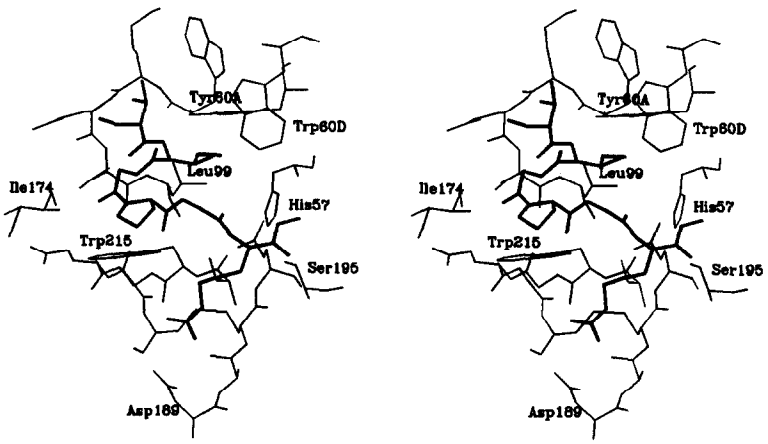

Fig. 2. Stereo-overlay of the gp120 V_3 loop primary consensus sequence with *cis* Pro^{313} in the $i + 2$ turn position (—) within the thrombin binding site site (——). Note the positioning of the Ile residue within the hydrophobic cluster formed by residues $Tyr^{60}A$ and $Trp^{60}D$, within the thrombin binding site.

bonyl oxygen, respectively, with N···O separations of 2–3 Å and angles of 140–160°. These hydrogen bonds are commonly observed with the serine proteases and their substrates and inhibitors.

4. Discussion

Notwithstanding a particularly high variability in the external envelope protein gp120, analysis of 245 different HIV-1 isolates of the PND in the V_3 loop has shown a greater than 80% conservation among 9 out of 14 central positions [12]. This conservation may be due to a requirement to maintain an essential β-strand-turn-strand structural motif. The G-P-G-R segment, which is part of the consensus sequence ...RKSIHI---GPGR*AFY... (with insertions of Q and R being most common at the dashed positions and * indicating the thrombin cleavage site) has been predicted to adopt a type II β-turn in the crown of the V_3 loop [12,16,17,28]. However, with the aid of computer-assisted molecular modeling, we have determined that this conformer can not realistically and productively bind in either the deep and narrow thrombin cleft or even in the relatively more exposed factor Xa active site (Fig. 1), and that a conformational shift must occur prior to proteolytic cleavage. Thus, we propose that binding of gp120 to CD4 at a site remote from the V_3 loop induces a conformational shift in the V_3 loop, which either stabilizes a *cis*-proline configuration or enhances the rate of proline *trans-cis* isomerization. Energy calculations suggest a $\Delta G°$ on the order of 1 to 2 kcal/mol for trans and cis proline [29] and an activation barrier of 13 kcal/mol (compared with 20 kcal/mol at other peptide bonds) [30]. The presence of Gly residues preceding and following Pro^{313} would further stabilize the *cis* Pro configuration [31]. This would favor a frameshift of the reverse turn from the presumably favored type II β-turn incorporating a *trans*-proline in the $i + 1$ position to a type VI turn [32] with a *cis*-proline in the $i + 2$ site as shown in Fig. 4. Precedence for this type of isomerism has been observed by X-ray crystallography for Gly^{42}–Pro^{43} of calbindin D_{9k} [33]. This CD4 induced shift would provide

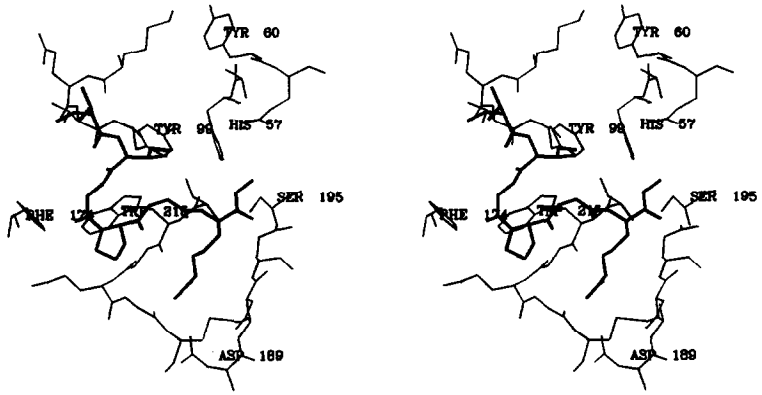

Fig. 3. Stereo-diagram of the gp120 V_3 loop primary consensus sequence with *cis* Pro^{313} in the $i + 2$ turn position (—) docked within the factor Xa binding site (——). The factor Xa binding cleft, unimpeded by the 60A-D insertion loop, permits binding of the model structure with even greater ease.

a mechanism for presentation of the V_3 loop as a competent 'trypsin-like' substrate to a cellular protease for activation for subsequent membrane fusion.

Although there is a high degree of conservation of the GPGR sequence among viral isolates, the residues flanking this region are hypervariable. In particular, the residue immediately N-terminal to the conserved region is most commonly either Arg (in SF-2, NL4-3 and IIIB which are T cell trophic isolates) or Ile (in BaL and SF162 which are monocyte-macrophage trophic isolates) [34]. To examine the role of the sequence variability at this site, we docked the two consensus sequences within the active sites of thrombin and factor Xa. It can be seen in Figs. 2 and 3 that the gp120 primary sequence incorporating the type VI turn can be readily accommodated in either the thrombin or factor Xa active sites. By comparison with the FPA bound structure [21,22], the turn region in the gp120 sequence is more deeply buried in the narrow active site of the enzyme. In MT trophic viral isolates the consensus residue in the *i* position of the turn is Ile. This residue occupies a site analogous to Leu^9 in the FPA bound structure. Strong hydrophobic interactions exist between the Ile side chain and the characteristic thrombin insertion loop around $Tyr^{60}A$–$Pro^{60}B$–$Pro^{60}C$-$Trp^{60}D$. In factor Xa, the hydrophobic pocket formed by Tyr^{99}, Phe^{174} and Trp^{215} is unhindered by the 60A-D insertion loop found in thrombin, thus providing more flexible accommodation of the Ile side chain. The gp120 conformation proposed for the thrombin active site can be docked into factor Xa without any difficulty (Fig. 3), and the more open hydrophobic pocket in factor Xa allows further adjustment of orientation to provide a better fit. The presence of large hydrophobic residues in the P3–P4 pockets of thrombin substrates is quite common [25]. In contrast, the conserved residue at this position is Arg in the T cell trophic viruses (IIIB, SF2) [34]. However, this non-conservative replacement can be readily accommodated and has precedence for occupying this site within thrombin-susceptible sequences [25]. The hydrophobic portion of the Arg side chain traverses the hydrophobic pocket created by the thrombin insertion loop and the Arg guanidinium group is within 3.5 Å of the abnormally acidic (by comparison with trypsin) Glu^{192} (residue 192 in factor Xa is Gln). The extended Pro^{313} to Arg^{315} region of the main chain is juxtaposed with the thrombin segment Ser^{214}–Gly^{216} in an antiparallel, slightly twisted manner; the $Pro^{313}O$ and the $Arg^{315}N$ form favorable hydrogen bonds with the $Gly^{216}N$ and $Ser^{214}O$, respectively, forming characteristic interactions between thrombin and its substrates and inhibitors. Further details of the interaction between the proposed consensus sequences and thrombin are shown in Fig. 3. Extensive efforts to dock the GPGR turn with Pro^{313} at the $i + 1$ position with Arg^{315} positioned in the S1 subsite and the scissile amide appositely oriented for interaction with Ser^{195}, proved to be unsuccessful with both thrombin and factor Xa (Fig. 1), principally due to the incompatible positioning of the amino-terminal residues (311–313) deeply within the active site, and causing major steric clashes with the thrombin insertion loop (Fig. 1). In factor Xa, Pro^{313} is abutting Tyr^{99} in addition to constraints on the entrance of the amino-terminal to the active site.

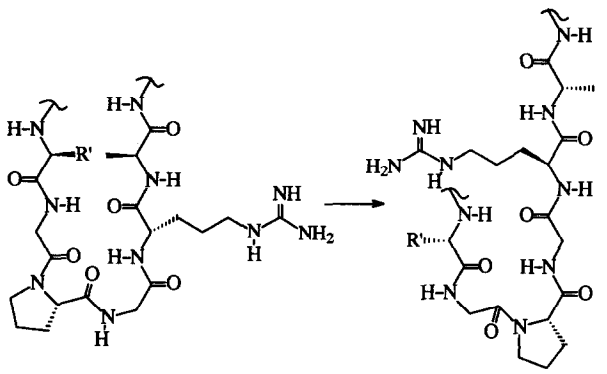

Fig. 4. Scheme showing the frame-shift moving the Pro^{313} from a *trans* configuration in the $i + 1$ position of a type II β-turn to a *cis* configuration in the $i + 2$ position of a type VI β-turn

The binding of CD4 to HIV-1 virions and to the surface of HIV-1 infected cells induces conformational changes in the viral envelope that are believed to prime the fusogenic activity of gp41. We have suggested a working hypothesis for the initial CD4 induced conformational changes involved in the proteolysis of the gp120 V_3 loop, which is apparently a requisite step in the infective process. Although there is a particularly high variability in the external envelope protein gp120, analysis of 245 different HIV-1 isolates of the PND in the V_3 loops has shown greater than 80% conservation among 9 out of 14 central positions [12]. This conservation may be due to a requirement to maintain an essential β-strand-turn-strand structural motif. Additionally, point mutations at G^{312}, G^{314} or R^{315} produce non-infectious virions [15].

In summary, we propose that upon binding to CD4, a conformational shift of the Gly^{312}–Pro^{313} bond from *trans* to *cis* is either induced or stabilized. Stabilization of this conformation might be effected by the interaction of conserved acidic residues (particularly Glu^{89}, in the CDR-3 like region of CD4) with a region of N-terminal basic residues in the V_3 loop [35]. This shift would favor the proline residue occupying the $i + 2$ position of a type VI β-turn [32]. Importantly, this hypothesis is consistent with the relative unimportance of mutagenesis of Pro^{313} to Ala^{313} [13], in that a type II β turn with Gly^{312} in the $i + 1$ and Ala^{313} in the $i + 2$ position would also be expected to be a competent thrombin-like proteolytic substrate. Alternatively, the type VI turn could be maintained in the $P^{313}A$ variant, as recently observed for a

carbonic anhydrase II variant [36]. We suggest that this shift is essential for V_3 loop cleavage by thrombin, factor Xa, or related proteases. Furthermore, this CD4 induced conformational shift would provide the virus with a timing mechanism which would prevent premature proteolytic cleavage and exposure of gp41 fusion domains [37]. Model studies with conformationally constrained peptides may be used to evaluate the mechanistic aspects of this hypothesis.

Acknowledgements: We thank Professor W.C. Still (Columbia) for a copy of MACROMODEL and Professor Brian Edwards (Wayne State) for the fibrinopeptide-thrombin structure. This work was supported by grants from the National Institutes of Health and the American Heart Association of Metropolitan Chicago, and by a National Science Foundation Presidential Young Investigator Award, Camille and Henry Dreyfus Teacher-Scholar Award, and an American Heart Association for an Established Investigator award (M.K.).

References

[1] Dalgleish, A.G., Beverley, P.C.L., Clapham, P.R., Crawford, D.H., Greaves, M.F. and Weiss, R.A. (1984) Nature 312, 763–767.
[2] Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C. and Montagnier, L. (1984) Nature 312, 767–768.
[3] Celada, F., Cambiaggi, C., Maccari, J., Burastero, S., Gregory, T., Patzer, E., Porter, J., McDanal, C. and Matthews, T. (1990) J. Exp. Med. 172, 1143–1150.
[4] Healey, D., Dianda, L., Moore, J.P., McDougal, J.S., Moore, M.J., Estess, P., Buck, D., Kwong, P.D., Beverley, P.C.L. and Sattentau, Q.J. (1990) J. Exp. Med. 172, 1233.
[5] Rusche, J.R., Javaherian, K., McDanal, C., Petro, J., Lynn, D.L., Grimaila, R., Langlois, A., Gallo, R.C., Arthur, L.O., Fischinger, P.J., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1988) Proc. Natl. Acad. Sci. USA 85, 3198–3202.
[6] Capon, D.J. and Ward, R.H.R. (1991) Annu. Rev. Immunol. 9, 649–678.
[7] Javaherian, K., Langlois, A.J., McDanal, C., Ross, K.L., Lawrence, I.E., Jellis, C.L., Profy, A.T., Rusche, J.R., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1989) Proc. Natl. Acad. Sci. USA 86, 6768–6772.
[8] Hattori, T., Koito, A., Takatsuki, K., Kido, H. and Katunuma, N. (1989) FEBS Lett. 248, 48–52.
[9] Clements, G.J., Price-Jones, M.J., Stephens, P.E., Sutton, C., Schulz, T.F., Clapham, P.R., McKeating, J.A., McClure, M.O., Thomson, S., Marsh, M., Kay, J., Weiss, R.A. and Moore, J.P. (1991) AIDS Res. Hum. Retrovir. 7, 3–16.
[10] Stephens, P.E., Clements, G. and Yarranton, G.T. (1990) Nature 343, 219.
[11] Koito, A., Hattori, T., Murakami, T., Matsushita, S., Maeda, Y., Yamamoto, T. and Takatsuki, K. (1989) Int. Immunol. 1, 8178–8189.
[12] LaRosa, G.J., Davide, J.P., Weinhold, K., Waterbury, J.A., Profy, A.T., Lewis, J.A., Langlois, A.J., Dreesman, G.R., Boswell, R.N., Shadduck, P., Holley, L.H., Karplus, M., Bolognesi, D.P., Matthews, T.J., Emini, E.A. and Putney, S.D. (1990) Science 249, 932–935.
[13] Gotoh, B., Yamauchi, F., Ogasawara, T. and Nagai, Y. (1992) FEBS Lett. 296, 274–278.
[14] Gotoh, B., Ogasawara, T., Toyoda, T., Inocencio, N.M., Hamaguchi, M. and Nagai, Y. (1990) EMBO J. 9, 4189–4195.
[15] Ivanoff, L.A., Dubay, J.W., Morris, J.F., Roberts, S.J., Gutshall, L., Sternberg, E.J., Hunter, E., Matthews, T.J. and Petteway Jr.,S.R. (1992) Virology 187, 423–432.
[16] Chandrasekhar, K., Profy, A.T. and Dyson, H.J. (1991) Biochem. 30, 9187–9194.
[17] Rini, J.M., Stanfield, R.L., Stura, E.A., Salinas, P.A., Profy, A.T. and Wilson, I.A. (1993) Proc. Natl. Acad. Sci. USA 90, 6325–6329.
[18] Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C. and Fenton II, J.W. (1990) Science 249, 277–280.
[19] Ni, F., Meinwald, Y.C., Vàsquez, M. and Scheraga, H.A. (1989) Biochemistry 28, 3094–3105.
[20] Nakanishi, H., Chrusciel, R.A., Shen, R., Bertenshaw, S., Johnson, M.E., Rydel, T.J., Tulinsky, A. and Kahn, M. (1992) Proc. Natl. Acad. Sci. USA 89, 1705–1709.
[21] Martin, P.D., Robertson, W., Turk, D., Huber, R., Bode, W. and Edwards, B.F.P. (1992) J. Biol. Chem. 267, 7911–7920.
[22] Stubbs, M.T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S.R. and Bode, W. (1992) Eur. J. Biochem. 206, 187–195.
[23] Padmanabhan, K., Padmanabhan, K.P., Tulinsky, A., Park, C.H., Bode, W., Huber, R., Blankenship, D.T., Cardin, A.D. and Kisiel, W. (1993) J. Mol. Biol. 232, 947–966.
[24] Huber, R. and Steigemann, W. (1974) FEBS Lett. 48, 235–237.
[25] Chang, J.-Y. (1985) Eur. J. Biochem. 151, 217–224.
[26] Ni, F., Konishi, Y., Bullock, L.D., Rivetna, M.N. and Scheraga, H.A. (1989) Biochemistry 28, 3106–3119.
[27] Ni, F., Konishi, Y., Frazier, R.B. and Scheraga, H.A. (1989) Biochemistry 28, 3082–3094.
[28] Wilmot, C.M. and Thornton, J.M. (1988) J. Mol. Biol. 203, 221–232.
[29] Wüthrich, K. and Grathwohl, C. (1974) FEBS Lett. 43, 337–340.
[30] Schultz, G.D. and Schirmer, R.H. (1978) Principles of Protein Structure, Springer-Verlag, New York.
[31] MacArthur, M.W. and Thornton, J.M. (1991) J. Mol. Biol. 218, 397–412.
[32] Lewis, P.N., Momany, F.A. and Scheraga, H.A. (1973) Biochim. Biophys. Acta 303, 211–229.
[33] Svensson, L.A., Thulin, E. and Forsen, S. (1992) J. Mol. Biol. 223, 601–606.
[34] Hwang, S.S., Boyle, T.J., Lyerly, H.K. and Cullen, B.R. (1991) Science 253, 71–74.
[35] Peterson, A. and Seed, B. (1988) Cell 54, 65–72.
[36] Tweedy, N.B., Nair, S.K., Paterno, S.A., Fierke, C.A. and Christianson, D.W. (1993) Biochemistry 32, 10944–10949.
[37] Sattentau, Q.J. and Moore, J.P. (1991) J. Exp. Med. 174, 407–415.