37 Bell, S.A. and Goodnow, C. (1994) *EMBO J.* 13, 816–826
38 Chen, C., Radic, M.Z., Erikson, J. *et al.* (1994) *J. Immunol.* 152, 1970–1982
39 Bensimon, C., Chastagner, P. and Zouali, M. (1994) *EMBO J.* 13, 2951–2962
40 Silverman, G.J., Sasano, M. and Wormsley, S.B. (1993) *J. Immunol.* 151, 5840–5855
41 Levy, J.A. (1993) *Microbiol. Rev.* 57, 183–289
42 Berberian, L., Valles-Ayoud, Y., Sun, N., Martinez-Maza, O. and Braun, J. (1991) *Blood* 78, 175–179
43 Müller, S., Wang, H., Silverman, G.J. *et al.* (1993) *Scand. J. Immunol.* 38, 327–334
44 Berberian, L., Shukla, J., Jefferis, R. and Braun, J. (1994) *J. AIDS* 7, 641–646
45 Goodglick, L. and Braun, J. (1994) *Am. J. Pathol.* 144, 623–636
46 Barbas, C.F., III, Collet, T.A., Amberg, W. *et al.* (1993) *J. Mol. Biol.* 230, 812–823
47 Kreier, J.P. and Mortensen, R.F. (1990) *Infection, Resistance and Immunity*, Harper & Row
48 Wilson, A.C., Carlson, S.S. and White, T.J. (1977) *Annu. Rev. Biochem.* 46, 573–639
49 Sasso, E.H., van Dijk, K.W., Bull, A.P. and Milner, E.C.B. (1993) *J. Clin. Invest.* 91, 2358–2367
50 Zouali, M., Madaio, M., Canoso, R. and Stollar, B.D. (1989) *Eur. J. Immunol.* 19, 1757–1760
51 Zachau, H.G. (1993) *Gene* 135, 167–173
52 Bentley, G.A., Boulot, G., Riottot, M-M. and Poljack, R.J. (1990) *Nature* 348, 254–257
53 Utsumi, S. and Karush, F. (1964) *Biochemistry* 3, 1329–1338
54 Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T. and Winter, G. (1989) *Nature* 341, 544–546
55 Hamers-Casterman, C., Atarhouvh, T., Muyldermans, S. *et al.* (1993) *Nature* 363, 446–448
56 Wu, T.T. and Kabat, E.A. (1992) *Mol. Immunol.* 29, 1141–1146
57 Schroeder, H.W., Jr *Semin. Immunol.* (in press)
58 Tutter, A. and Riblet, R. (1989) *Proc. Natl Acad. Sci. USA* 86, 7460–7464
59 Schroeder, H.W., Jr, Hillson, J.L. and Perlmutter, R.M. (1990) *Int. Immunol.* 2, 41–50
60 Friedman, S.M., Posnett, D.N., Tumang, J.R., Cole, B.C. and Crow, M.K. (1991) *Arthritis Rheum.* 34, 468–480
61 Andris, J.S., Johnson, S., Zolla-Pazner, S. and Capra, J.D. (1991) *Proc. Natl Acad. Sci. USA* 88, 7783–7787
62 Marasco, W.A., Bagley J., Zani, C. *et al.* (1992) *J. Clin. Invest.* 90, 1467–1478
63 Moran, M.J., Andris, J.S., Matsumato, Y.I., Capra, J.D. and Hersh, E.M. (1993) *Mol. Immunol.* 30, 1543–1551
64 Bagley, J., Dillon, P.J., Rosen, G. *et al.* (1994) *Mol. Immunol.* 31, 1149–1169
65 van der Donck, E.M.M., Schutten, M., Osterhaus, A.D.M.E. and van der Heuden, R.W.J. (1994) *AIDS Res. Hum. Retroviruses* 10, 1639–1648
66 Andris, J.S., Ehrlich, P.H., Ostberg, L. and Capra, J.D. (1992) *J. Immunol.* 149, 4053–4059
67 Andris, J.S., Brodeur, B.R. and Capra, J.D. (1993) *Mol. Immunol.* 30, 1601–1616
68 Adderson, E.E., Shackelford, P.G., Quinn, A. and Carroll, W.L. (1991) *J. Immunol.* 147, 1667–1674
69 Adderson, E.E., Shackelford, P.G., Insel, R.A., Quinn, A., Wilson, P.M. and Carroll, W.L. (1992) *J. Clin. Invest.* 89, 729–738
70 Huang, D.F., Olee, T., Masuho, Y. *et al.* (1992) *J. Clin. Invest.* 90, 2197–2208
71 Newkirk, M.M., Gram, H., Heinrich, G.F. *et al.* (1988) *J. Clin. Invest.* 81, 1511–1518
72 Ikematsu, H., Harindranath, N., Ueki, Y., Notkins, A.L. and Casali, P. (1993) *J. Immunol.* 150, 1325–1337
73 Zouali, M. and Theze, J. (1991) *J. Immunol.* 146, 2855–2864

letters

T-cell receptor signal sequences

Future studies of the physical and structural chemistry of the T-cell receptor (TCR) will be greatly facilitated by prokaryotic, or other heterologous, expression systems. In such studies, no matter whether only the mature portion is expressed and refolded[1] or a prokaryotic signal sequence is fused[2–5], the assignment of the original signal cleavage site becomes crucial. The exact length of the TCR is also important in all model-building studies (e.g. see Ref. 6).

However, the position of signal cleavage sites of TCRs is not yet known. Therefore, most publications (e.g. Refs 3,7–12) use the alignment to antibody variable domains as a guide, or relate it to other TCRs, themselves assigned by alignments. In contrast to the older literature, which acknowledges this fact, current sequence databases do not explicitly indicate the hypothetical nature of such approaches. Thus, the signal cleavage sites have wrongly assumed the status of fact. Furthermore, the assignment of the signal cleavage site is not always reasonable from a biochemical point of view.

Using biochemical and statistical criteria[13], some of the cleavage sites proposed in the literature, which would give rise to very long signal sequences, appear very implausible. Briefly, a typical signal peptide as determined experimentally for many proteins comprises three parts[13]: a short (1–3 residues) positively charged N-terminal region (which, in eukaryotes, includes the α-amino group as a positive charge); followed by a hydrophobic core (7–15 residues); and completed by the C-terminal signal peptide cleavage site (3–7 residues). Predictive schemes for cleavage sites have been developed using data from proteins with experimentally determined cleavage sites[14]. Using the algorithm of the program SIGSEQ2, cleavage sites were predicted for the α chains of six TCRs with long signal sequences in the database (Fig. 1). The sites predicted by SIGSEQ2 were found to be different from the ones given in the literature based on antibody homology, which reeeive exceptionally poor scores in the prediction. Even a superficial application of the biochemical criteria makes these sites improbable. In every case, cleavage would occur too far from the hydrophobic core

letters

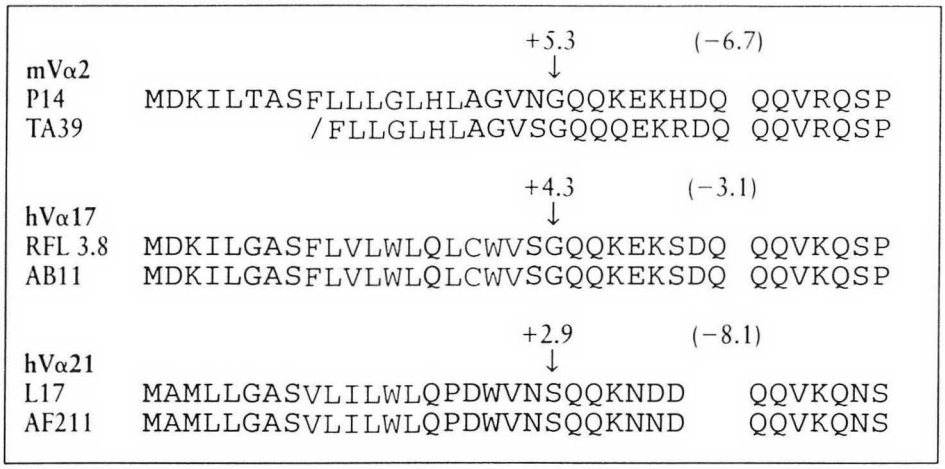

Fig. 1. *The beginning of the open reading frame of six T-cell receptor (TCR) clones from three human subfamilies[12]. The sequences are derived from translation of cDNA sequences. The gap illustrates the cleavage site assumed in the cited references (P14, Ref. 7; TA39, Ref. 8; RFL 3.8, Refs 3,9; AB11, Ref. 10; L17, Ref. 11; AF211, Ref. 10). The sequence of TA39 is only known as the given fragment (the solidus indicates that this sequence is known only from this point on). The signal sequence of RFL 3.8 has been assigned twice, and differently each time (since Ref. 9 does not give any reason for its assignment, this figure follows the assignment used in Ref. 3). Hydrophobic regions are shown in red. Cleavage sites predicted by the program SIGSEQ2 are given by arrows, and the score of this site and the one calculated for the cited cleavage site are indicated (a positive number indicates a high probability of cleavage at this site). Abbreviations: m, mouse; h, human.*

of the signal peptide (Fig. 1). For this reason, our studies on the expression of TCR fragments in *Escherichia coli*[5] utilize signal cleavage sites based on such biochemical and statistical criteria.

Independently of the considerations raised above concerning expression systems for soluble, recombinant TCRs (reviewed in Ref. 15), reconstructing the correct N-terminus of the protein will also be important. By pointing out the apparent contradiction between the widely used signal sequence assignments, and the biochemical and statistical plausibility of such alignments, we hope to provoke experimental determinations of dubious cleavage sites.

Christoph Wülfing
Andreas Plückthun

Universität Zürich,
Biochemisches Institut,
Winterthurerstr. 190,
CH-8057 Zürich, Switzerland.

References
1 Kurucz, I., Jost, C.R., George, A.J.T., Andrew, S.M. and Segal, D.M. (1993) *Proc. Natl Acad. Sci. USA* 90, 3830–3834
2 Ward, E.S. (1992) *J. Mol. Biol.* 224, 885–890
3 Novotny, J., Ganju, R.K., Smiley, S.T. *et al.* (1991) *Proc. Natl Acad. Sci. USA* 88, 8646–8650
4 Soo Hoo, W.F., Lacy, M.J., Denzin, L.K., Voss, E.W., Hardman, K.D. and Kranz, D.M. (1992) *Proc. Natl Acad. Sci. USA* 89, 4759–4763
5 Wülfing, C. and Plückthun, A. (1994) *J. Mol. Biol.* 242, 655–669
6 Chothia, C., Boswell, D.R. and Lesk, A.M. (1988) *EMBO J.* 7, 3745–3755
7 Pircher, H., Michalopoulos, E.E., Iwamoto, A. *et al.* (1987) *Eur. J. Immunol.* 17, 1843–1846
8 Arden, B., Klotz, J.L., Siu, G. and Hood, L.E. (1985) *Nature* 316, 783–787
9 Diamond, D.J., Szalay, P., Symer, D. *et al.* (1991) *J. Exp. Med.* 174, 229–241
10 Klein, M.H., Concannon, P., Everett, M., Kim, L.D., Hunkapiller, T. and Hood, L.E. (1987) *Proc. Natl Acad. Sci. USA* 84, 6884–6888
11 Leiden, J.M., Fraser, J.D. and Strominger, J.L. (1986) *Immunogenetics* 24, 17–23
12 Wilson, R.K., Lai, E., Concannon, P., Barth, R.K. and Hood, L.E. (1988) *Immunol. Rev.* 101, 149–172
13 von Heijne, G. (1990) *J. Membr. Biol.* 115, 195–201
14 Folz, R.J. and Gordon, J.L. (1987) *Biochem. Biophys. Res. Commun.* 146, 870–877
15 Wülfing, C. and Plückthun, A. (1995) *Immunologist* 3, 59–66

Sequence similarity between β-cell autoantigens

Most humans and nonobese diabetic (NOD) mice with insulin-dependent diabetes mellitus (IDDM) manifest autoimmune reactivity to insulin and/or isoforms of the enzyme glutamic acid decarboxylase (GAD-65 and -67). Whether any of these molecules actually trigger or drive a pathogenic immune response against pancreatic β cells remains controversial, and no basis for association among them has previously been proposed. We have identified a 13 amino acid sequence of high similarity (with six identical and three conserved residues) between human GAD-65 (residues 506–518) and human pro-insulin (24–36). This similarity extends to human GAD-67 and both of the mouse pro-insulins and GADs.

Other than the similarity between pro-insulin and GAD, this sequence has several interesting features. It lies within a region of human GAD-65 (473–543) shown to be immunodominant for T cells of IDDM patients[1], and largely overlaps one of the two mouse GAD-65 peptides (509–528) reported to be the first recognized by NOD mouse T cells[2]. The sequence contains a (+)XXX(−) motif characteristic of most I-A^{g7} (I-A^{NOD})-binding peptides reported to date[3–6], as well as an XYXXXXXXXV motif for H-$2K^{d}$ (Ref. 7), the major histocompatibility complex (MHC) class I of NOD implicated in the pathogenesis of IDDM (Ref. 8). Both of these motifs overlap the cleavage site in pro-insulin (Arg32-Glu33), which separates the C-peptide from the C-terminus of the insulin B chain, and would therefore be absent from the mature insulin molecule.

Evidence exists for 'T-cell epitope mimicry'[9] and 'determinant spreading'[10] in other autoimmune disorders. Reactivity to this pro-insulin