Technical Comment

Similarity of Protein G and Ubiquitin

A. M. Gronenborn *et al.* (*1*) describe the three-dimensional structure of the immunoglobulin G (IgG) binding domain of protein G, a streptococcal multidomain protein on the cell surface. I have found an unexpected structural similarity of the IgG binding domain of protein G to ubiquitin (*2*), a protein that is thought to be part of the intracellular protein degradation pathway. The similarity can be seen by comparing the fold of the IgG binding domain of protein G with that of ubiquitin (Fig 1).

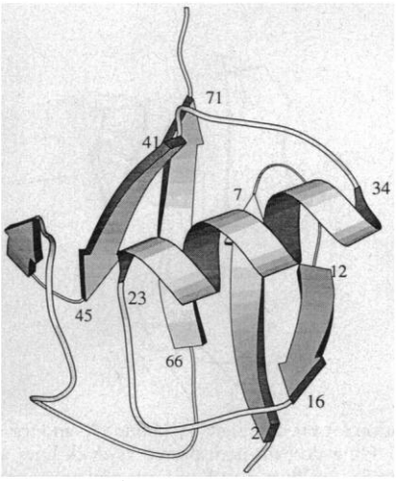

Fig. 1. Schematic ribbon drawing of ubiquitin.

Both proteins share the unusual crossover motif of the outer strands of the β sheet and the central α helix. The topology of the fold is identical in both proteins except for a short additional β strand in ubiquitin. There is no discernible sequence similarity between the two proteins. Although neither protein contains any disulfide bridges, both are extremely stable to denaturation.

Ubiquitin has been identified as part of a lymphocyte-homing receptor, and it has been proposed (*3*) that ubiquinated cell-surface molecules could play a role in cell-cell interaction and adhesion. This raises the question of whether ubiquitin and the IgG binding domain of protein G are evolutionarily as well as structurally related.

Such a relation would not be unprecedented, as it has been shown that the chaperone protein PapD, which mediates the assembly of pili in *Escherichia coli*, contains the same structural fold as the immunoglobulin-type domain. It has also been proposed that "genes involved in eukaryotic cell-cell interactions may have been recruited by bacteria to aid in their attachment to eukaryotic cells" (*4*). It is possible that this hypothesis also applies to the IgG binding domain of protein G and ubiquitin.

PER J. KRAULIS
Department of Biochemistry,
University of Cambridge,
Tennis Court Road,
Cambridge, CB2 1QW,
United Kingdom

REFERENCES

1. A. M. Gronenborn *et al.*, *Science* **253**, 657 (1991).
2. S. Vijay-Kumar, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **194**, 531 (1987); Brookhaven Protein Data Bank (PDB accession number 1 UBQ).
3. M. Siegelman *et al.*, *Science* **231**, 823 (1986).
4. A Holmgren and C.-I. Branden, *Nature* **342**, 248 (1989).

21 August 1991; accepted 10 September 1991

Response: We thank P. J. Kraulis for his comment. The structures of the IgG binding domain of protein G and ubiquitin are examples of a growing number of protein structures that have been found with unrelated sequences but similar folding motifs (*1*).

The IgG binding domain (Fig. 1) has a four-stranded sheet with a $-1, +3x, -1$ topology, while ubiquitin has a five-strand ed sheet with a $-1, +3x, +1, -2x$, topology. A least squares best match of the two structures with the program O (*2*) reveals that 41 residues of the two proteins can be superimposed with a backbone atomic root-mean-square difference of 2.2 Å (Fig. 2). The resulting sequence

AAAS–Newcomb Cleveland Prize

To Be Awarded for an Article or a Report Published in *Science*

The AAAS–Newcomb Cleveland Prize is awarded to the author of an outstanding paper published in *Science*. The value of the prize is $5000; the winner also receives a bronze medal. The current competition period began with the 7 June 1991 issue and ends with the issue of 29 May 1992.

Reports and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the competition period, readers are invited to nominate papers appearing in the Reports or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS–Newcomb Cleveland Prize, AAAS, Room 924, 1333 H Street, NW, Washington, D.C. 20005, and **must be received on or before 30 June 1992.** Final selection will rest with a panel of distinguished scientists appointed by the editor of *Science*.

The award will be presented at the 1993 AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.

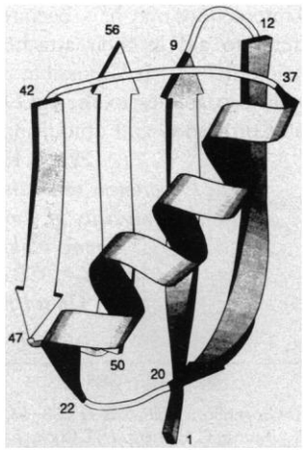

Fig. 1. Schematic ribbon drawing of the IgG binding domain of protein G, produced with the program Molscript (*3*).

alignment based on the structural superposition (Fig. 3) indicates that within the 41 boxed residues with backbones that can be superimposed, there is only 12% sequence identity (5 out of 41 residues).

Although speculations about a possible evolutionary relation between these two proteins may be tempting, there is little evidence to support this notion. Indeed, it may be that this particular structural motif represents an energetically favorable folding unit that is not derived from a common ancester.

ANGELA M. GRONENBORN
G. MARIUS CLORE
Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

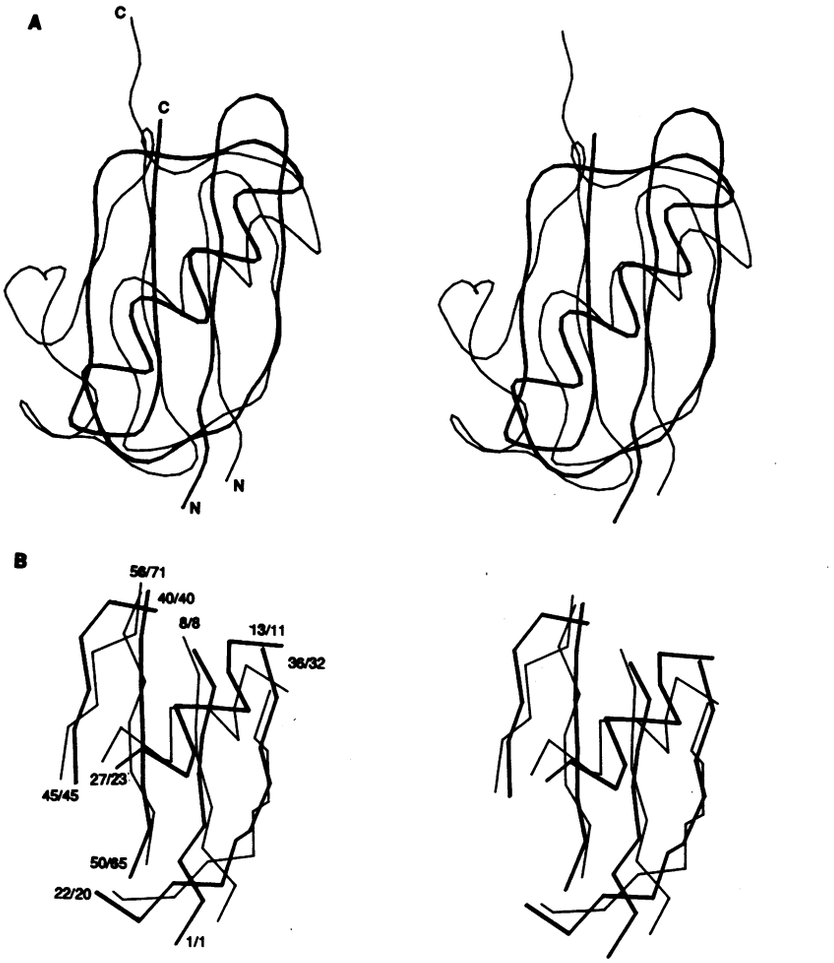

Fig. 2. Stereograms of backbone atoms of the superposition for the complete backbone (**A**) and for the C^{α} atoms (**B**) of the matched 41 residues. The IgG binding domain is displayed as thick lines and ubiquitin as thin lines. These 41 residues make up strand β1 (residues 1 to 8 of both proteins), strand β2 (residues 13 to 22 of the IgG binding domain and 11 to 20 of ubiquitin), part of the helix (residues 27 to 36 of the IgG binding domain and 23 to 32 of ubiquitin), strand β3 (residues 40 to 45 of both proteins), strand β4 of the IgG binding domain (residues 50 to 56), and strand β5 of ubiquitin (residues 65 to 71).

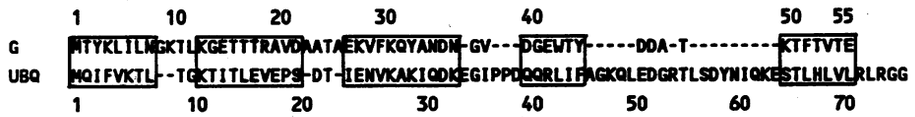

Fig. 3. Sequence alignment of the IgG binding domain of protein G (G) and ubiquitin (UBQ); boxed residues indicate backbone λ atoms that can be superimposed. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

REFERENCES

1. G. Vriend and C. Sander, *Proteins* **11**, 52 (1991).
2. T. A. Jones and M. Kjeldgaard, *O Version 5 Manual* (Univ. of Uppsala, Uppsala, Sweden, 1990).
3. P. J. Kraulis, *Appl. Cryst.*, in press.

4 September 1991; accepted 10 September 1991