Purification of Yolk Protein 2 of *Drosophila melanogaster* and Identification of Its Site of Tyrosine Sulfation*

(Received for publication, May 2, 1988)

Patrick A. Baeuerle†, Friedrich Lottspeich§, and Wieland B. Huttner¶

†From the Cell Biology Programme, European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, and the
§Max-Planck-Institute for Biochemistry, Genzentrum, Am Klopferspitz 18a, D-8033 Martinarder, Federal Republic of Germany

We have identified the site of tyrosine sulfation in an insect secretory protein, yolk protein 2 of *Drosophila melanogaster*. Yolk proteins were purified from [35S] sulfate-labeled flies, and yolk protein 2 was separated from yolk protein 1 and yolk protein 3 by preparative two-dimensional polyacrylamide gel electrophoresis. After digestion of yolk protein 2 with trypsin and reversed-phase high performance liquid chromatography, the sulfate label was recovered in two distinct sulfopeptides which, however, had identical NH2-terminal sequences and contained 3 tyrosine residues each. After chymotryptic digestion of the two tryptic sulfopeptides, the sulfate label was recovered in one sulfopeptide which contained a single tyrosine sulfate residue. NH2-terminal sequencing showed that this tyrosine residue corresponded to tyrosine 172 of the yolk protein 2 precursor (Hung, M.-C., and Wensink, P. C. (1983) J. Mol. Biol. 164, 487-492) in the sequence Glu-Thr-Thr-Asp-Tyr(S)-Ser-Asn-Val. This in vitro sulfation of tyrosine 172 by the vertebrate tyrosylprotein sulfotransferase.

Sulfation is the most abundant modification of tyrosine residues known. In multicellular eukaryotic organisms many secretory proteins and a few membrane proteins have been shown to undergo tyrosine sulfation (for reviews, see Refs. 1-3). Sulfate transfer is catalyzed by the enzyme tyrosylprotein sulfotransferase, an integral membrane protein of the trans Golgi (4, 5). Common structural features of the sequences surrounding sulfated tyrosine residues have been noted (1-3, 6). However, with the exception of hirudin, these sequences are all from vertebrate proteins, and little is known about the evolutionary conservation of tyrosine sulfation sites.

The function of protein tyrosine sulfation has so far only been elucidated for some of the small tyrosine-sulfated peptides, but not for any of the identified larger tyrosine-sulfated polypeptides. In the case of small tyrosine-sulfated peptides, studies on the role of tyrosine sulfation were facilitated by the fact that peptides can be chemically synthesized in sulfated and unsulfated form for comparative analyses. Using synthetic peptides it has been shown that tyrosine sulfation is required for the hormonal activity of cholecystokinin (for review, see Ref. 7). However, in the case of larger polypeptides, chemical synthesis of the sulfated and unsulfated form of a peptide is not feasible. Moreover, even if such synthesis were feasible, possible roles of tyrosine sulfation of proteins prior to their secretion could not be studied using synthetic polypeptides. A highly specific approach to investigating functional aspects of tyrosine sulfation of a protein before its secretion is to express the protein in cells after rendering it unsulfatable by site-directed mutagenesis. Prerequisites for this approach are the cloning of DNA coding for the protein and the identification of its tyrosine sulfation site(s).

We have previously shown that the three yolk proteins of *Drosophila melanogaster* are the major tyrosine-sulfated proteins in female flies (8). Yolk protein 2 (YP2)1 contains about 1 mol of tyrosine sulfate/mol of polypeptide, in contrast to YP1 and YP3 which have a higher content of tyrosine sulfate (8). Thus, YP2 is likely to contain only a single site of tyrosine sulfation. The sequence of YP2 is known from the sequence of the cloned DNA (9), which facilitates the identification of its site of tyrosine sulfation. In order to compare the structure of a tyrosine sulfation site of an insect secretory protein with that of the known tyrosine sulfation sites of vertebrate secretory proteins and as the first step towards functional studies involving site-directed mutagenesis of a tyrosine sulfation site, we report here the purification of YP2 from *Drosophila* flies and the identification of the site of tyrosine sulfation.

**EXPERIMENTAL PROCEDURES AND RESULTS**

We have previously shown that YP2 of *D. melanogaster* contains 1 mol of tyrosine sulfate/mol of polypeptide and that all isoelectric variants of this protein are sulfated to the same extent (8). This means either that all YP2 molecules are sulfated on the same tyrosine residue or that the YP2 molecule contains 1 mol of tyrosine sulfate/mol of polypeptide, in contrast to YP1 and YP3 which have a higher content of tyrosine sulfate (8). Thus, YP2 is likely to contain only a single site of tyrosine sulfation.

**DISCUSSION**

We have previously shown that YP2 of *D. melanogaster* contains 1 mol of tyrosine sulfate/mol of polypeptide and that all isoelectric variants of this protein are sulfated to the same extent (8). This means either that all YP2 molecules are sulfated on the same tyrosine residue or that the YP2 molecule contains 1 mol of tyrosine sulfate/mol of polypeptide, in contrast to YP1 and YP3 which have a higher content of tyrosine sulfate (8). Thus, YP2 is likely to contain only a single site of tyrosine sulfation. The sequence of YP2 is known from the sequence of the cloned DNA (9), which facilitates the identification of its site of tyrosine sulfation. In order to compare the structure of a tyrosine sulfation site of an insect secretory protein with that of the known tyrosine sulfation sites of vertebrate secretory proteins and as the first step towards functional studies involving site-directed mutagenesis of a tyrosine sulfation site, we report here the purification of YP2 from *Drosophila* flies and the identification of the site of tyrosine sulfation.

1 The abbreviations used are: YP, yolk protein; PMSF, phenylmethylsulfonyl fluoride; HEPES, 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; TPCK, tosyl-L-phenylalanine chloromethyl ketone; HPLC, high performance liquid chromatography; TLCK, tosyl-L-lysine chloromethyl ketone.

2 Portions of this paper (including "Experimental Procedures," "Results," Figs. S1-S3, Tables S1-SII, and Footnotes 3-5) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full-size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
Tyrosine Sulfation Site in Drosophila Yolk Protein 2

FIG. 1. Sequence positions of sulfopeptides in YP2. Numbers refer to the sequence of the YP2 precursor (9). Bars A and B, sequence positions of tryptic sulfopeptides A and B, respectively; bar CA=B, sequence position of the tryptic/chymotryptic sulfopeptides CA=B with (solid part) or without (solid plus dashed part) additional tryptic cleavage at arginine 180.

cules are sulfated on various tyrosine residues in a mutually exclusive manner, each molecule containing only 1 sulfated residue. The present study shows that the former is the case since only 1 tyrosine residue of YP2, tyrosine 172, was found to be sulfated (see Fig. 1). Our results are consistent with tyrosine 172 being the exclusive site of sulfation in YP2, although they do not strictly exclude the presence of a very minor second site of sulfation. However, as will be reported elsewhere (27), the complete lack of sulfation of YP2, mutated to phenylalanine in position 172, confirms that tyrosine 172 is indeed the exclusive site of sulfation in YP2.

It has been noted (4, 22, 23) that the sequences surrounding sulfated tyrosine residues contain multiple acidic amino acids. More recently, the sequences surrounding identified tyrosine sulfate residues in several, mostly vertebrate, proteins have been compared and analyzed in detail (1, 6). As a result of these studies, the role of acidic residues has been characterized, and additional consensus features of tyrosine sulfation sites have been outlined (1, 2, 6). It is of interest to note that the insect tyrosine sulfation site identified in the present study is very similar to the known vertebrate tyrosine sulfation sites in terms of amino acid composition and secondary structure. The hallmark of stoichiometrically sulfated tyrosines in vertebrate proteins is the presence of an acidic amino acid residue in position −1 and a total of at least 3 acidic amino acid residues between positions −5 and +5 of the sulfated tyrosine (1, 6). Tyrosine 172 is the only tyrosine residue in YP2 with this feature (see Table I). Three other tyrosine residues in YP2 (tyrosines 92, 377, 401; see Table I) are also preceded by an acidic amino acid residue in position −1 but lack the total of three acidic amino acid residues between positions −5 and +5. One tyrosine residue in YP2 (tyrosine 220; see Table I) is characterized by 3 acidic residues between positions −5 and +5, with one in position −2, but lacks an acidic residue in position −1. Since none of these other 4 tyrosine residues in YP2 are sulfated to any detectable extent, it appears that the requirements for the presence of acidic amino acids in tyrosine sulfation sites are very similar, if not identical, in vertebrates and insects. The lack of sulfation of tyrosine 220 in YP2 also supports the notion (see Fig. 7 in Ref. 1) that in the absence of an acidic residue in position −1, a total of 3 acidic residues between positions −5 and +5 may not be sufficient for sulfation and that 4 or more acidic residues are required, as is the case in factor X and gastrins (see Fig. 6 in Ref. 1).

An additional explanation as to why tyrosine 172 but not tyrosine 220 (despite the presence of 3 acidic residues from positions −5 to +5) is sulfated, concerns the secondary structure of tyrosine sulfation sites. The presence of turn-inducing amino acids has been proposed to be another consensus feature of tyrosine sulfation sites (1, 2). The sequence surrounding tyrosine 172 contains a total of 5 residues with significant turn-conformational preference (>1.3; Pro, Gly, Asp, Ser, Asn; Ref. 24) between positions −7 and +7 (with 2 such residues at positions −1 and +1), whereas that surrounding tyrosine 220 contains only 2 residues with turn-conformational preference between positions −7 and +7 (at positions −4 and +4). Recently, the Ω loop has been described as a novel category of protein secondary structure that occurs almost always at the protein surface (25). It is worth noting that in the sequence surrounding tyrosine 172 in YP2, all residues between positions −6 and +4 belong to the category of amino acids found preferentially in loops (f > 1.0; see Table 3 in Ref. 25). Since such amino acids are also frequently found in the sequences surrounding identified tyrosine sulfate residues in vertebrate proteins (see Fig. 6 in Ref. 1), we propose that loop structures are a typical feature of tyrosine sulfation sites that is conserved in vertebrates and insects.

The similarity in structure between the insect tyrosine sulfation site identified here and the known vertebrate tyrosine sulfation sites suggests a high evolutionary conservation of the interaction of protein substrates with tyrosylprotein sulfotransferase. In the accompanying paper (26), it is shown that tyrosine 172 of Drosophila YP2 is specifically and stoichiometrically sulfated in vivo by a mammalian tyrosylprotein sulfotransferase after expression of the protein in mouse fibroblasts.

Table I

Comparison of sequences surrounding "acidic" tyrosine residues in YP2

The sequences surrounding tyrosine residues (vertical lines, position 0) with acidic amino acid residues at positions −1 or −2 are shown from positions −7 to +7. The sequence positions of tyrosine residues refer to the YP2 precursor (9). Bold letters, acidic amino acids; underlined letters, amino acids with a turn-conformational preference >1.3 (Pro, Gly, Asp, Ser, Asn; Ref. 24).

| Sequence position of tyrosine | "Acidic" tyrosine residues in YP2 | Acidic residue at −1 | At least 3 acidic residues between −5 and +5 | At least 3 residues with turn-conformational preference between −7 and +7 | Sulfated in vivo |
|-----------------------------|----------------------------------|---------------------|---------------------------------------------|---------------------------------------------|----------------|
| −7                          |                                  |                     |                                             |                                             | Yes            |
| 0                           |                                  |                     |                                             |                                             | Yes            |
| +7                          |                                  |                     |                                             |                                             | Yes            |
| 172                         | Q P Y E T T D | Y | S N E E Q S Q | Yes | Yes | Yes | Yes |
| 92                          | N H V F K P D | Y | T P E P S Q | I | Yes | No | Yes | No |
| 220                         | A I E D F E Q | Y | A T L N I E R | No | Yes | No | No | No |
| 377                         | A A S Y Q E | Y | K Q N K G Y G | Yes | No | Yes | No | No |
| 401                         | D F D L G D | Y | I L Q V N S K | Yes | No | Yes | No | No |
Tyrosine Sulfation Site in Drosophila Yolk Protein 2

Acknowledgments—We thank Professors G. Hartmann and H. Thoenen for support of the doctoral thesis of P. A. Baeuerle, H.-W. Mewes for help with HPLC, Drs. J. Hesse and T. Rau kindly supplied us with flies; Dr. E. Friederich and J. Stinchcombe made helpful comments on the manuscript, and R. Wainwright typed the manuscript.

REFERENCES

1. Huttner, W. B., and Baeuerle, P. A. (1988) Mod. Cell Biol. 6, 97-140
2. Huttner, W. B. (1987) Trends Biochem. Sci. 12, 361-363
3. Huttner, W. B. (1988) Annu. Rev. Physiol. 50, 363-376
4. Lee, R. W. H., and Huttner, W. B. (1988) Proc. Natl. Acad. Sci. U. S. A. 82, 6143-6147
5. Baeuerle, P. A., and Huttner, W. B. (1987) J. Cell Biol. 105, 2655-2664
6. Hortin, G., Folz, R., Gordon, J. I., and Strauss, A. W. (1986) Biochem. Biophys. Res. Commun. 141, 326-333
7. Mutt, V. (1986) in Gastrointestinal Hormones (Glass, G. B. J., ed) pp. 169-221, Raven Press, New York.
8. Baeuerle, P. A., and Huttner, W. B. (1985) J. Biol. Chem. 260, 6434-6439
9. Huttner, M.-C., and Wensink, P. C. (1983) J. Mol. Biol. 164, 481-492
10. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275
11. Laemmli, U. K. (1970) Nature 227, 680-685
12. O’Farrell, P. H. (1975) J. Biol. Chem. 250, 4007-4021
13. Lottspeich, F. (1985) J. Chromatogr. 326, 321-327
14. Tsgut, A., Uchida, T., Mewes, H.-W., and Ataka, T. (1987) J. Biol. Chem. 262, 1593-1597
15. Godel, H., Graser, T., Foldi, P., and Fürst, P. (1984) J. Chromatogr. 297, 49-61
16. Mahowald, A. P., and Kamyszewski, M. P. (1989) in The Genetics of Drosophila (Ashburner, M., and Wright, T. R., F., eds) Vol. 2, pp. 185-188, Academic Press, London
17. Warren, T. G., and Mahowald, A. P. (1979) Dev. Biol. 78, 130-139
18. Hovemann, B., Galler, R., Wältler, V., Küpper, H., and Bautz, E. K. P. (1981) Nucleic Acids Res. 9, 4725-4731
19. Huttner, M.-C., and Wensink, P. C. (1981) Nucleic Acids Res. 9, 6407-6419
20. Hortin, G., Sims, H., and Strauss, A. W. (1986) J. Biol. Chem. 261, 1786-1793
21. Gregory, R. A., Hardy, P. M., Jones, D. S., Kenner, G. W., and Sheppard, R. C. (1984) Nature 309, 533-533
22. Lee, R. W. H., and Huttner, W. B. (1983) J. Biol. Chem. 258, 11326-11334
23. Huttner, W. B., Baeuerle, P. A., Benedum, U. B., Friederich, F., Hille, A., Lee, R. W. H., Rosa, P., Seydel, U., and Suchanek, C. (1986) Colloq. INSERM (Inst. Natl. Sante Rech. Med.) 199, 197-217
24. Argos, P., Hanei, M., Wilson, J. M., and Kelley, W. N. (1983) J. Biol. Chem. 258, 6450-6457
25. Leszcynski, J. F., and Rose, C. D. (1986) Science 234, 849-855
26. Friederich, E., Baeuerle, P. A., Garoff, H., Hovemann, B., Huttner, W. B. (1988) J. Biol. Chem. 263, 14930-14938
27. Friederich, E., Fritz, H.-J., and Huttner, W. B. (1988) J. Cell Biol., in press

EXPERIMENTAL PROCEDURES

PURIFICATION OF YOLK PROTEIN FROM FLIES OF DROSOPHILA MELANOGASTER AND IDENTIFICATION OF ITS N-TERTIARY SEQUENCE OF THE TYROSINE SULFATE GROUP

Tyrosinesulfation site in Drosophila yolk protein 2 was investigated in detail. We have used the respective bacterial expression product of Drosophila yolk protein 2, which contains a cysteine-rich domain, in order to determine the tyrosine sulfate group in the yolk protein complex in vitro. The purified yolk protein complex was extensively fractionated using various chromatographic techniques and the sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry.

Preparation of tyrosine sulfates

The tyrosine sulfates were purified from the yolk protein complex by a combination of affinity chromatography and reverse phase HPLC. The tyrosine sulfates were identified by mass spectrometry and the sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry.

Preparation of tyrosine sulfates

The tyrosine sulfates were purified from the yolk protein complex by a combination of affinity chromatography and reverse phase HPLC. The tyrosine sulfates were identified by mass spectrometry and the sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry. The sulfate group was further analyzed by reverse phase HPLC and identified by mass spectrometry. The sulfate group was found to be covalently bound to the tyrosine residue in the yolk protein complex. The tyrosine sulfate group was identified by mass spectrometry.
Tyrosine Sulfation Site in Drosophila Yolk Protein 2

Allots of fractions A-D obtained during the purification of 35S-sulfate-labeled yolk proteins (see Fig. 1 of miniprint section) were either subjected to SDS-PAGE followed by determination of the 35S-radioactivity present in the sum of YPI, YP2 and YP3 (yolk protein-specific 35S-radioactivity), or analyzed for total protein. Values are expressed for each entire fraction; yields are given in brackets. The purification factor for the sum of the three yolk proteins is given on the right.

| Fraction | Yolk protein-specific 35S-radioactivity (cpm x 10^-6) | Total protein (mg) | Purification factor |
|----------|-----------------------------------------------------|--------------------|-------------------|
| A        | 13.5 (100)                                          | 381 (100)          | 1.0               |
| B        | 6.7 (51.2)                                          | 127 (39.3)         | 1.5               |
| C        | 5.4 (41.5)                                          | 14.2 (5.7)         | 11.1              |
| D        | 4.0 (32.8)                                          | 4.4 (1.2)          | 24                |

Trypsin-sulfopptides of YP2: After preparative two-dimensional PAGE, sulfopptides were generated by tryptic treatment of gel pieces containing YP2. This resulted in an essentially quantitative elution (≈ 95%) of the 35S-radioactivity. In order to remove gel debris and Coomassie blue, and to concentrate sulfopptides, the tryptic eluate was subjected to chromatography on a C18 SEPAK cartridge. A small portion of the 35S-radioactivity did not bind to the resin and was recovered in the flow-through (Table II of miniprint section). Most of the 35S-radioactivity bound to the resin and was almost quantitatively eluted with 20% methanol, whereas only insignificant amounts of 35S-radioactivity (but all of the Coomassie blue) were eluted in a subsequent step using 100% methanol (Table II).

When the 35S-radioactivity recovered in the sum of the HPLC fractions was compared to that subjected to HPLC, significant losses (10-30%) of 35S-radioactivity were noted. These losses were not the result of desulfation of sulfopptides prior to, or during, HPLC since no 35S-radioactivity was found in the first HPLC fractions (where sulfate would be recovered) and since tryrosine sulfate was found to be stable in the HPLC buffers used. Rather, these losses appeared to be due to unspecific adsorption of sulfopptides during HPLC since similar losses were observed upon rechromatography of sulfopptides A and B.

The recovery of 35S-radioactivity in the various fractions obtained after SEPAK C18 chromatography and HPLC, and the profile of 35S-radioactivity after HPLC, were very similar to those observed when 35S-sulfate-labeled YP2 obtained from a two-dimensional gel of the total protein of 35S-sulfate-labeled female flies, was digested with trypsin in the presence of unlabeled carrier YP2 which had been purified from unlabeled fly by the method described in Fig. 1 (see Table II, data not shown). This indicated that the sulfation of YP2 did not change during the purification.
Tyrosine Sulfation Site in Drosophila Yolk Protein 2

TABLE II
Quantitation of the recovery of the $^{35}$S-sulfate label in the various tryptic/chymotryptic peptides of YP2

Values are expressed for each entire fraction. The first value in brackets gives the recovery of $^{35}$S-radioactivity with respect to the total $^{35}$S-radioactivity recovered after the respective chromatography. The second value gives the recovery with respect to the total radioactivity recovered after the respective chromatography. The starting material for experiment 1 was YP2 purified from a mixture of unlabelled and $^{35}$S-labeled flies as described in Fig. 1. The starting material for experiment 2 was a mixture of $^{35}$S-labeled YP2 obtained from two-dimensional gels of whole fly homogenate and YP2 purified from unlabelled flies as described in Fig. 1. MeOH: methanol; total peak: sum of $^{35}$S-radioactivity in the fractions containing sulfopeptides A and B plus the adjacent fractions (see Fig 2, top panel).

| Tryptic eluate | SEP-PAK C18 chromatography | HPLC of 0-20% MeOH fraction | HPLC of chymotryptic digests of sulfopeptides A and B |
|----------------|-----------------------------|-----------------------------|----------------------------------------------------|
|                | 0/20% MeOH (x)              | 0/20% MeOH (x)              |                                                    |
|                | flow-through (y)            | 42,250 (78/85)              | 14,000 (33/49)                                     |
|                | CPM                          | total peak                  | sulfopeptide A                                    |
|                | recovery (%)                 | 25,300 (60/89)              | sulfopeptide A                                    |
| Experiment 1   | 54,000                       | 7,360 (78/85)               | not determined                                    |
|                | sulfopeptide A               | 14,000 (33/49)              |                                                   |
|                | sulfopeptide B               | 7,450 (19/28)               |                                                    |
|                | total peak                   | 25,300 (60/89)              |                                                    |
|                | 20-100% MeOH (z)             | 10 (0/0)                    |                                                    |
| Experiment 2   | 63,000                       | 9,760 (15/18)               | 13,500 (32/34)                                    |
|                | flow-through (y)            | 42,200 (67/76)              | sulfopeptide A                                    |
|                | CPM                          | total peak                  | sulfopeptide A                                    |
|                | recovery (%)                 | 35,560 (74/85)              | sulfopeptide A                                    |
|                | sulfopeptide A               | 13,500 (32/34)              | not determined                                    |
|                | sulfopeptide B               | 16,936 (45/51)              |                                                    |
|                | total peak                   | 35,560 (74/85)              |                                                    |
|                | 20-100% MeOH (z)             | 3,300 (5/6)                 |                                                    |

![Fig. 2. Separation of tryptic peptides of YP2 by HPLC.](image)

![Fig. 3. Chymotryptic digestion of sulfopeptides A and B.](image)