Characterization of the cysteine residues and disulphide linkages in the protein crystal of *Bacillus thuringiensis*

Henri P. L. BIETLOT,* Indira VISHNUBHATLA,* Paul R. CAREY,† Marianne POZSGAY† and Harvey KAPLAN*‡

*Department of Chemistry, University of Ottawa, Ottawa, Ontario K1N 6N5, and †Division of Biological Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

Bacillus thuringiensis produces a 130–140 kDa insecticidal protein in the form of a bipyramidal crystal. The protein in the crystals from the subspecies *kurstaki* HD-1 and *entomocidus* was found to contain 16–18 cysteine residues per molecule, present primarily in the disulphide form as cystine. Evidence that all the cysteine residues form symmetrical interchain disulphide linkages in the protein crystal was obtained from the following results: (i) the disulphide diagonal procedure [Brown & Hartley (1966) *Biochem. J.* **101**, 214–228] gave only unpaired cysteic acid peptides in diagonal maps; (ii) the disulphide bridges were shown to be labile in dilute alkali and the crystal protein could be released quantitatively with 1 mM-2-mercaptoethanol; (iii) the thiol groups of the released crystal protein were shown by competitive labelling [Kaplan, Stevenson & Hartley (1971) *Biochem. J.* **124**, 289–299] to have the same chemical properties as exposed groups on the surface of the protein; (iv) the thiol groups in the released crystal protein reacted quantitatively with iodoacetate or iodoacetamide. The finding that all the disulphide linkages in the protein crystal are interchain and symmetrical accounts for its alkali-lability and for the high degree of conservation in the primary structure of the cystine-containing regions of the protein from various subspecies.

INTRODUCTION

During sporulation *Bacillus thuringiensis* produces a family of insecticidal proteins that form crystalline inclusion bodies. The major component of crystals toxic to lepidopteran larvae is a 130–140 kDa protein (protoxin) that, on ingestion by susceptible insect larvae, is proteolysed to a 60–70 kDa fragment (toxin). The details of the toxic mechanism have not been elucidated, but the toxin binds to receptors in the midgut epithelium, causing cell lysis and eventual larval death (Aronson *et al.*, 1986; Andrews *et al.*, 1987; Brousseau & Masson, 1988). Approx. 20 homologous genes coding for the 130–140 kDa protoxin have been identified, and many subspecies carry several of these genes (Höfte & Whitely, 1989).

Dissolution of the crystal and release of soluble protoxin can be effected by thiol reagents or alkali (pH > 10.5), indicating that interchain disulphide bridges are responsible for stabilizing the crystal structure (Nickerson, 1980). Jacquet *et al.* (1987) have shown that the solubilization of the crystal protein under the alkaline conditions of the insect gut is an important first step for the generation of insecticidal activity and that the disulphide structure may play a role in this process. The gene-deduced cysteine content varies from 13 to 19 residues per molecule, and almost all of these residues are located in the C-terminal half of the molecule (Brousseau & Masson, 1988). In the activation process the C-terminal region and the first 28 amino acid residues from the N-terminus of the protoxin are excised, leaving a cysteine-free toxin (Aronson & Arvidson, 1987; Bietlot *et al.*, 1989). Comparison of the gene-derived primary structures of the protoxins shows that the N-terminal region, which contains two cysteine residues, and C-terminal half, containing all the remaining cysteine residues, are highly conserved (Höfte & Whitely, 1989). At present, the significance of the unusual distribution of cysteine residues and their highly conserved nature is unclear.

The cysteine residues in the crystal protein appear to be present primarily in the disulphide form, although no direct cystine quantification has been reported (Dastidar & Nickerson, 1979). Nickerson (1980) has presented evidence that interchain disulphide bridges are an important factor in determining the crystal structure and properties. Interchain disulphide bridges are relatively uncommon in proteins, and the intrachain/interchain ratio in the *B. thuringiensis* crystal protein is unknown. Accordingly the present study was undertaken to quantify the cysteine/cystine content of the crystal protein, to determine the alignment of cysteine residues involved in interchain or intrachain linkages and to elucidate the role of the cysteine residues in crystal formation and generation of insecticidal activity.

MATERIALS AND METHODS

Materials

1-Fluoro-2,4-dinitro[³H]benzene ([³H]Dnp-F) (16.6 Ci/mmol) and [¹⁴C]Dnp-F (12 mCi/mmol) were obtained from Amersham Corp., Oakville, Ont., Canada. Alanine and N-acetyl-L-cysteine were supplied by Sigma Chemical Co., St. Louis, MO, U.S.A., Poropak Q was supplied by Waters Associates, Mississauga, Ont., Canada, Renografin by Squibb (Canada), Montreal, Que., Canada, and Sephadryl S-300 by Pharmacia, Baie d'Urté, Que., Canada. All other reagents, including h.p.l.c. solvents, were high-purity preparations obtained from commercial sources.

Crystal preparation

B. thuringiensis subsp. *kurstaki* HD-1 and subsp. *entomocidus* were grown in half-strength trypticase soy broth. The cells were lysed in double-distilled water and the crystals were purified by using Renografin gradients as described previously (Carey *et al.*, 1986; Pozsgay *et al.*, 1987). Crystals were routinely checked for

Abbreviations used: Caps, 3-(cyclohexyl)propane-1-sulphonic acid; CM-cysteine, carboxymethylcysteine; Dnp, dinitrophenyl; Dnp-F, 1-fluoro-2,4-dinitrobenzene.

‡ To whom correspondence should be addressed.

purity by SDS/PAGE. Crystals were shown to be free of contaminating proteinases by a lack of proteolytic activity towards denatured [^{14}C]methylated haemoglobin (Rice & Means, 1971). Purified crystals were stored in distilled water at 4 °C.

Preparation of solubilized protoxin

Purified crystals (25 mg) were solubilized by incubation in 5 ml of 2% (v/v) 2-mercaptoethanol at pH 10. Any insoluble material was removed by centrifugation. The solubilized crystal protein was applied to a Sephacryl S-300 column (120 cm $\times$ 1 cm) and eluted with 1 mM-EDTA/0.1% (v/v) 2-mercaptoethanol/50 mM-Tris/HCl buffer, pH 8.0 (Yamamoto & McLaughlin, 1981). The flow rate was 15 ml/h, and 3 ml fractions were collected. The pooled fractions containing the protoxin were dialysed against 0.1 M-KCl/5 mM-N-methylmorpholine/5 mM-sodium borate buffer, pH 8.0. Homogeneity of the protoxin peak was verified by SDS/PAGE.

Amino acid analysis

Samples were hydrolysed in 6 M-HCl for 24 h and 48 h *in vacuo*, with norleucine added as an internal standard. Amino acids were quantified on a TSM Technicon amino acid analyser by using a ninhydrin detection system.

Protein quantification

The protein crystal was dissolved in KOH solution, pH 13. Estimates of protein concentration were made from u.v. absorbance at 280 nm using an absorption coefficient of 1.37 ml/mg (Bietlot *et al.*, 1989).

Determination of cystine content

Purified crystals (1 mg) were acid-hydrolysed, and cystine was quantified by amino acid analysis.

Determination of cysteine content

(i) Crystals (1 mg) were oxidized with performic acid according to the procedure of Hirs (1956). Cysteic acid was quantified by amino acid analysis after acid hydrolysis. (ii) Crystals (1 mg) were dissolved in 100 μl of 8 M-urea/0.2 M-Tris/HCl buffer, pH 8.6, containing 10 μl of 2-mercaptoethanol. Iodoacetate (26.8 mg in 100 μl) was added and allowed to react for 30 min. The samples were dialysed against distilled water and freeze-dried. Carboxymethylcysteine (CM-cysteine) was quantified by amino acid analysis after acid hydrolysis.

Reaction of crystals with iodoacetic acid

Crystals (25 mg) were suspended in 5.0 ml of 8 M-urea/0.2 M-Tris/HCl buffer, pH 8.6. Iodoacetic acid (0.268 g in 1.0 ml of 1 M-NaOH) was added. Samples (1.0 ml) were withdrawn after 0.5 h, 1 h, 2 h and 24 h, and 1 drop of 12 M-HCl was added to each. The samples were dialysed against distilled water and freeze-dried. CM-cysteine was determined by amino acid analysis after acid hydrolysis.

Reaction of crystals with iodo[^{14}C]acetate

Crystals (25 mg) were suspended in 5.0 ml of 8 M-urea/0.2 M-Tris/HCl buffer, pH 8.6. Iodo[^{14}C]acetate (5 μCi , sp. radioactivity 56 mCi/mmol) was added and allowed to react for 2 h with gentle shaking. The reaction was terminated by the addition of 1 drop of 12 M-HCl and the sample was dialysed and freeze-dried. The sample was then allowed to react with unlabelled iodoacetate in 8 M-urea/0.2 M-Tris/HCl buffer, pH 8.6, containing 2-mercaptoethanol as described above.

Reaction of solubilized protoxin with iodoacetic acid and iodoacetamide

The solubilized protoxin (1 mg) in 100 μl of 0.2 M-sodium metaborate buffer, pH 9.0, was allowed to react with iodoacetic acid (26.8 mg in 100 μl of 1 M-NaOH) or iodoacetamide (27 mg) for 5 min. CM-cysteine was quantified by amino acid analysis after dialysis and acid hydrolysis.

Solubilization of *kurstaki* HD-1 crystals with 2-mercaptoethanol

Crystals (500 μg) were suspended in 0.1 M-Caps/NaOH buffer, pH 10.5, containing amounts of 2-mercaptoethanol ranging from 0 to 2% (v/v) and were incubated for 5 min at 25 °C. A 5 μl portion of iodoacetate (10 mg/ml, pH 10.5) was added and the sample was subjected to non-reducing PAGE.

Gel electrophoresis

Gels were run on a Pharmacia Phast electrophoresis system with preformed gels and other materials supplied by Pharmacia. Samples were dissolved in sample buffer [2.5% (w/v) SDS/5% (v/v) 2-mercaptoethanol/1 mM-EDTA/10 mM-Tris/HCl buffer, pH 8.3], placed in a boiling-water bath, cooled and then applied to 10–15% gradient gels.

Diagonal electrophoresis

Enzymic digestion. Crystals were suspended in 10% (v/v) formic acid (2.5 mg/ml), and pig pepsin (Sigma Chemical Co.) (1:50, w/w) was added. The samples were incubated at 37 °C with gentle agitation. The digests clarified within 5 min. They were left overnight (18 h) and then freeze-dried. Where a second digestion was carried out with elastase, chymotrypsin, trypsin or thermolysin (all from Sigma Chemical Co.), the freeze-dried peptic digest was dissolved in 30 mM-NH₄HCO₃ (2.5 mg/ml) and the appropriate enzyme (1:50, w/w) was added. The samples were digested for 18 h and then freeze-dried.

Electrophoresis. The procedure employed was essentially the two-dimensional paper-electrophoretic procedure described by Brown & Hartley (1966), but required a modification because of the large size of the crystal protein. The enzymic digest was applied to Whatman 3 MM paper along a 10 cm band at 1 mg of digest/cm for electrophoresis in the first dimension at pH 6.5 (60 V/cm for 40 min). Usually a 3 cm guide-strip is removed for performic acid oxidation; however, because of the high molecular mass of the protein, the entire 10 cm band was used in order to have sufficient peptide for detection. The 10 cm strip was sewn on to a larger sheet and buffer was applied to each side, so that it concentrated the peptides into the centre of the strip. After drying, a 3 cm strip was cut from the centre, oxidized with performic acid, and then run in the second dimension and stained with Cd²⁺/ninhydrin reagent (Darbre, 1986). The entire 10 cm neutral band from the pH 6.5 electrophoresis was run at pH 2.1 (60 V/cm for 50 min). The peptides were concentrated as described above, oxidized and then run at pH 6.5.

Peptide isolation

Peptides were isolated according to the procedure described by Brown & Hartley (1966) with the modification described above. After elution from paper, peptides were further purified by reverse-phase h.p.l.c.

Competitive labelling

Preparation of [^{14}C]Dnp derivatives. S-[^{14}C]Dnp-cysteine and [^{14}C]Dnp-alanylalanine were prepared by treating N^ε-acetylcysteine and alanylalanine with [^{14}C]Dnp-F as described previously (Hefford *et al.*, 1985). The purification procedure was the same except that the final purification step was C₁₈ reverse-

phase h.p.l.c. The [^{14}C]Dnp-alanylalanine was eluted by using the following solvents: 0.01 M-HCl/17% (v/v) acetonitrile for 6 min, then 0.01 M-HCl/25% (v/v) acetonitrile for 4 min followed by 0.01 M-HCl/60% (v/v) acetonitrile. The *S*-[^{14}C]Dnp-cysteine was eluted isocratically with 0.01 M-HCl/60% (v/v) acetonitrile.

Competitive labelling of protoxin. Portions (1.0 ml) of a stock solution containing 0.113 mg/ml (0.357 μM) of protoxin and alanylalanine (75.0 μM) in 5 mM-*N*-methylmorpholine/5 mM-sodium borate buffer were equilibrated at 25 °C in a thermostatically controlled water bath. The pH was adjusted to the desired value between 7 and 10 with either 1 M-NaOH or 1 M-HCl. A 50 μl portion of acetonitrile containing [^3H]Dnp-F (20.8 nmol, sp. radioactivity 16.6 Ci/mmol) was added with vigorous stirring and the reaction was allowed to proceed for 18 h in the dark. Conc. HCl was added to adjust the pH to 2.0.

The following was added to each sample: a 1.0 ml portion of a 25% (v/v) acetone solution containing 2500 d.p.m. of [^{14}C]Dnp-alanylalanine and *S*-[^{14}C]Dnp-cysteine and 0.03 mg of each unlabelled Dnp derivative as carrier. The acetone was removed by evaporation, and the samples were hydrolysed in 6 M-HCl for 18 h at 110 °C *in vacuo*.

Purification of [^3H]Dnp and [^{14}C]Dnp derivatives

The [^3H]Dnp and [^{14}C]Dnp derivatives were isolated and purified as described previously (Hefford *et al.*, 1985), with the only change being that the final purification step was carried out by C_{18} reverse-phase h.p.l.c.

Liquid-scintillation counting

Samples were dissolved in 0.1 ml of 0.01 M-HCl and added to 10 ml of Aquasol-2. Scintillation counting was carried out on a programmable LKB 1215 Rack Beta scintillation counter equipped with automatic quench correction and a d.p.m. converter.

RESULTS AND DISCUSSION

Cysteine content and free thiol groups

Cysteine analysis (Table 1) of intact crystal protein was carried out by the following three procedures: (i) hydrolysis of crystal in 6 M-HCl and quantification of cystine; (ii) performic acid oxidation and quantification of cysteic acid after acid hydrolysis; (iii) reduction with 2-mercaptoethanol, reaction with iodoacetate and quantification of CM-cysteine after acid hydrolysis. The free thiol group content (Table 1) was determined by treating crystals in 8 M-urea with iodoacetate and quantifying the CM-cysteine produced on acid hydrolysis. Urea was used to swell the crystals and maximize the accessibility of the reagent to any reactive

groups. It was found that the same results were obtained whether the reaction was carried out for 30 min or 24 h, indicating that the free thiol groups were completely accessible to the iodoacetate.

The cysteine contents obtained for the crystal proteins from the *kurstaki* HD-1 and *entomocidus* subspecies are consistent with the gene-deduced values of 17 and 16 cysteine residues for the *CryIA(a)* (4.5 kb gene) and *CryIA(c)* (6.6 kb gene) genes respectively (Brousseau & Masson, 1988; Höfte & Whitely, 1989). Our results therefore indicate that one or both of these gene products are major components of these crystals. The cystine content obtained confirms the deductions made in earlier reports that most of the cysteine residues present in the crystal protein form disulphide bonds (Nickerson, 1980). Some free thiol groups were detected in the crystals from both the *kurstaki* HD-1 and *entomocidus* subspecies, approx. 2 and 1 mol/mol of protoxin respectively. This free thiol group content could result from one or two specific cysteine residues in the crystal protein existing in the free thiol form or from a small proportion of each cysteine residue being present in the free thiol form. When the crystals were treated with iodo[^{14}C]acetate and digested with pepsin, a large number of radiolabelled peptides were detected by radioautography of the acidic peptides after electrophoresis at pH 6.5 and in the neutral peptides after electrophoresis at pH 2.1.

Fig. 1. Diagonal maps of the *B. thuringiensis* subsp. *kurstaki* HD-1 protein crystal digested with pepsin

(a) Diagonal electrophoresis was carried out at pH 6.5. (b) Electrophoresis of the neutral peptide band obtained at pH 6.5 was carried out at pH 2.1 followed by oxidation and electrophoresis at pH 6.5. Strongly staining peptides are indicated with continuous lines and faintly staining peptides with broken lines.

Table 1. Cysteine contents of the crystal proteins of *B. thuringiensis* subsp. *kurstaki* HD-1 and subsp. *entomocidus*

Values are given as means $\pm$ S.D. for five analyses.

Procedure	Cysteine content (mol/mol of 130 kDa protein)	
	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> HD-1	<i>B. thuringiensis</i> subsp. <i>entomocidus</i>
Cystine	8.0 $\pm$ 1.0	7.0 $\pm$ 1.5
Cysteic acid	18.1 $\pm$ 1.7	15.3 $\pm$ 2.3
CM-cysteine	15.9 $\pm$ 1.0	18.6 $\pm$ 0.8
Free thiol groups	1.9 $\pm$ 0.1	1.2 $\pm$ 0.2

No evidence for the incorporation of radioactive label into one or two specific cysteine residues was obtained. It therefore appears that even the mild alkaline conditions at pH 8.6 are sufficient to generate some free thiol from each of the disulphide bridges present in the crystal protein.

Alignment of disulphide bridges in the crystal protein

The most effective enzyme for digestion of the crystal protein was found to be pepsin, which rapidly solubilizes crystals leaving no insoluble core. Fig. 1(a) shows a pH 6.5 diagonal map of a peptic digest of the subspecies *kurstaki* HD-1 crystal. Most of the cystine-containing peptides lie in the neutral band, with one acidic peptide (HD-1-1) and several faint spots in the acidic and basic regions. When the neutral band is run at pH 2.1, concentrated, oxidized with performic acid and then re-run at pH 6.5, the chromatogram in Fig. 1(b) is obtained. The notable feature of these diagonal maps is that none of the strongly staining peptides is paired. The crystal protein from subspecies *entomocidus* gives a very similar pH 6.5 diagonal map (Fig. 2a), with an acidic peptide (ENT-1) running in the same position as that observed with the protein crystal from the subspecies *kurstaki* HD-1. The neutral peptides give the peptide map in Fig. 2(b). Again, no pairing of the cysteine acid-containing peptides is observed.

Fig. 2. Diagonal maps of the *B. thuringiensis* subsp. *entomocidus* protein crystal digested with pepsin

(a) Diagonal electrophoresis was carried out at pH 6.5. (b) Electrophoresis of the neutral peptide band obtained at pH 6.5 was carried out at pH 2.1 followed by oxidation and electrophoresis at pH 6.5. Strongly staining peptides are indicated with continuous lines and faintly staining peptides with broken lines.

Diagonal peptide maps were also run on the following enzymic digests of the subspecies *kurstaki* HD-1 and subspecies *entomocidus* crystals: pepsin + chymotrypsin, pepsin + elastase, pepsin + thermolysin and pepsin + trypsin. Figs. 3 and 4 show the diagonal peptide maps obtained for the pepsin + elastase digests of protein crystals from these two subspecies. Comparison of Figs. 3(a) and 4(a) with Figs. 1(a) and 2(a) shows that there is an acidic peptide lying off the diagonal in approximately the same position as the acidic peptides, HD-1-1 and ENT-1, from the peptic digest. A second, slower-moving, acidic cystine-containing peptide is generated by the second digestion, but most of the cystine-containing peptides still lie in the neutral band. The neutral peptides (Figs. 3b and 4b) give rise to more peptides than with the peptic digestion alone (Figs. 1b and 2b), but again no pairing of the cysteine acid-containing peptides is observed. All the other digests gave the same results, with no pairing among the major cysteine acid-containing peptides for both crystal types.

Unpaired cysteine acid-containing peptides can arise with the diagonal method in three instances: (i) a single polypeptide with an intrachain disulphide linkage; (ii) two identical cysteine-containing peptides joined by an interchain disulphide bond; (iii) intrachain disulphide bonds between repeating sequences. As the *B. thuringiensis* crystal protein has been shown by gene sequencing not to contain any repeating sequences, the last

Fig. 3. Diagonal maps of the *B. thuringiensis* subsp. *kurstaki* HD-1 protein crystal digested with pepsin + elastase

(a) Diagonal electrophoresis was carried out at pH 6.5. (b) Electrophoresis of the neutral peptide band obtained at pH 6.5 was carried out at pH 2.1 followed by oxidation and electrophoresis at pH 6.5. Strongly staining peptides are indicated with continuous lines and faintly staining peptides with broken lines.

Fig. 4. Diagonal maps of the *B. thuringiensis* subsp. *entomocidus* digested with pepsin + elastase

(a) Diagonal electrophoresis was carried out at pH 6.5. (b) Electrophoresis of the neutral peptide band obtained at pH 6.5 was carried out at pH 2.1 followed by oxidation and electrophoresis at pH 6.5. Strongly staining peptides are indicated with continuous lines and faintly staining peptides with broken lines.

possibility can be eliminated. The former two possibilities may be distinguished by treating the peptic digest with another enzyme in order to cleave peptides with an intrachain disulphide linkage and observing if any paired peptides appear on the diagonal maps. The fact that, with crystals from two different *B. thuringiensis* subspecies, no pairing of the major peptides was observed with any enzyme digest, namely pepsin, pepsin + chymotrypsin, pepsin + elastase, pepsin + thermolysin and pepsin + trypsin, strongly suggests that the vast majority of the disulphide bridges that link the polypeptide chains in the crystal protein are between the same cysteine residues on each chain, i.e. the cysteine residues form predominantly symmetrical interchain bonds in the protein crystal.

The peptides HD-1-1 and ENT-1 (Figs. 1a and 2a) are unusually acidic and seemed likely candidates for peptides that contained an intrachain disulphide linkage. These peptides were isolated and both gave an analysis (Table 2) for a peptide with the following composition: Asp₃Cys(SO₃H)₁Glu₁Gly₁Ile₁Leu₁Thr₁Val₁. Since only one cysteine residue is present, this peptide cannot arise from the cleavage of an intrachain disulphide linkage. The highly acidic nature of the peptides is accounted for by the presence of glutamic acid and aspartic acid residues. The gene-deduced sequences for the *CryIA(a)*-gene-, *CryIA(b)*-gene- and *CryIA(c)*-gene-product types contain the sequence Ile-Asp-

Table 2. Composition of the acidic off-diagonal peptides HD-1-1 and ENT-1 of peptide maps of peptic digests of the crystal proteins of *B. thuringiensis* subsp. *kurstaki* HD-1 and subsp. *entomocidus* respectively

Values were calculated from the amino acid analysis as mole fraction of amino acid $\times 10$. Nearest integers are given in parentheses.

Amino acid	Amino acid composition (mol of residue/10 mol of total residues)	
	Peptide HD-1-1	Peptide ENT-1
Aspartic acid	3.02 (3)	2.83 (3)
Cysteic acid	0.80 (1)	0.95 (1)
Glutamic acid	1.25 (1)	1.26 (1)
Glycine	1.15 (1)	1.11 (1)
Isoleucine	0.89 (1)	0.82 (1)
Leucine	1.01 (1)	0.80 (1)
Threonine	0.91 (1)	0.91 (1)
Valine	0.95 (1)	1.02 (1)

Val-Gly-Cys-Thr-Asp-Leu-Asn-Glu, which corresponds to the composition obtained for these peptides (Schnepf & Whitely, 1985; Adang *et al.*, 1985; Geiser *et al.*, 1986).

In a few cases pairing of some of the faint spots (e.g. Fig. 2a) is observed. However, these are minor products that vary with each chromatogram and could arise after digestion from oxidation of trace free thiol groups present or from a small amount of non-symmetrical bridges present in the crystal as a result of disulphide interchange. We cannot absolutely rule out the possibility that they may be derived from an intrachain disulphide linkage that, because of its particular structure, is generated in low yield. Other than this observation no evidence was obtained for the existence of any intrachain disulphide bridges in the crystal protein.

Stability of disulphide linkages

Creighton (1986) has demonstrated on the basis of kinetic and thermodynamic considerations that intrachain disulphide linkages have a much greater stability than interchain disulphide linkages. In order to test whether there are two such classes of disulphide linkages within the crystal protein, various amounts of 2-mercaptoethanol were added to the crystals from the *kurstaki* HD-1 subspecies. This was followed by reaction with iodoacetate to prevent re-formation of disulphide linkages, and then the protein was run on non-reducing SDS/polyacrylamide gels. Some 130 kDa protoxin protein and 65 kDa mosquitocidal protein (Yamamoto & McLaughlin, 1981) are solubilized at pH 10.5 in the absence of 2-mercaptoethanol (Fig. 5, lane 1), but the bulk of the protein remains at the top of the stacking gel. As the 2-mercaptoethanol concentration is increased, more of the crystal protein goes into the gel, and a small amount of an intermediate with molecular mass 210 kDa is observed (Fig. 5, lanes 2 and 3). The entire crystal protein is solubilized with only 0.01% (v/v) (approx. 1 mM) 2-mercaptoethanol (Fig. 5, lane 6), indicating a very high lability of the disulphide linkages. Couche *et al.* (1987) observed a similar effect with dithioerythritol and the protein crystal from the subspecies *israelensis*. The free thiol group content observed at pH 8.6 (Table 1) most probably arises from the alkali-lability of the interchain disulphide bridges in the protein crystal.

The protoxin would be expected to have a higher mobility in non-reducing SDS/polyacrylamide gels than in reducing SDS/polyacrylamide gels, if any intrachain disulphide linkages were present (Creighton, 1989). However, crystal protein run on

Fig. 5. Solubilization of the protein crystal from *B. thuringiensis* subsp. *kurstaki* with 2-mercaptoethanol

An SDS/10–15%-gradient-polyacrylamide gel stained with Coomassie Blue is shown. Crystals were incubated with various concentrations of 2-mercaptoethanol: lane 1, 0%; lane 2, 0.001%; lane 3, 0.002%; lane 4, 0.005%; lane 6, 0.01%; lane 7, 0.02%; lane 8, 0.055%. Lane 5 shows molecular-mass standards.

reducing gels gives a protoxin band with the same mobility as the protoxin in the non-reducing gels in Fig. 5, indicating that no intrachain disulphide linkages are present. The results obtained therefore support the conclusion, drawn from the diagonal mapping, that all the disulphide linkages are interchain.

Competitive labelling

Competitive labelling (Kaplan *et al.*, 1971; Young & Kaplan, 1989) of the solubilized protoxin was carried out at pH values between 7.0 and 10.0 with alanylalanine added as an internal standard. The reactivity of the side chains of the cysteine residues of the solubilized protoxin, relative to the α -amino group of alanylalanine, was determined by reaction with a trace amount of [^3H]Dnp-F. In order to determine the relative reactivities of the functional groups, it is necessary to quantify the attachment of the [^3H]Dnp label at each group. This was achieved by adding equal amounts of *S*-[^{14}C]Dnp-cysteine and [^{14}C]Dnp-alanylalanine to each reaction mixture, extracting the [^3H]Dnp-alanylalanine/[^{14}C]Dnp-alanylalanine and then hydrolysing the mixture in 6 M-HCl. The $^3\text{H}/^{14}\text{C}$ ratios of the purified *S*-Dnp-cysteine and Dnp-alanine were quantified by scintillation counting. This $^3\text{H}/^{14}\text{C}$ ratio for *S*-Dnp-cysteine was corrected for the amount of cysteine present in the protoxin (Table 1).

The data were analysed by using the expression (Young & Kaplan, 1989):

$$\alpha_x r = \alpha_s (^3\text{H}/^{14}\text{C})_x / (^3\text{H}/^{14}\text{C})_s$$

where α_x is the degree of ionization of the functional group under study, α_s is the degree of ionization of the internal standard alanylalanine (pK 8.31) at 25 °C, r is the pH-independent second-order velocity constant for the reaction of the functional group under study relative to the amino group of alanylalanine, and $(^3\text{H}/^{14}\text{C})_x$ and $(^3\text{H}/^{14}\text{C})_s$ are the radioactivity ratios determined by scintillation counting for the functional group and alanylalanine respectively. A plot of $\alpha_x r$ versus pH will give a titration curve with pK equal to the pH at the inflexion point and a limiting value of $\alpha_x r = r$ if the groups under study titrate normally over the entire pH range employed.

The solubilized protoxin from subspecies *kurstaki* HD-1 has

Fig. 6. pH-reactivity profile for the reaction of the thiol groups of the solubilized protoxin from *B. thuringiensis* subsp. *kurstaki* HD-1 with Dnp-F

The continuous line is a theoretical titration curve with pK 8.45 and $r = 39.6$. Solvent conditions were 0.1 M-KCl/5 mM-*N*-ethylmorpholine/5 mM-sodium borate buffer at 25 °C.

a molecular mass of 130 kDa, and we have observed a tendency to form insoluble high-molecular-mass aggregates during the purification on Sephacryl S-300 columns. It also has a relatively low solubility at neutral pH values. The properties of the protoxin therefore make it necessary to study it in very dilute solution. Modifications in the competitive labelling methodology have made it possible to study proteins in very dilute solution (Hefford *et al.*, 1985). The present study was carried out at a protoxin concentration of 0.35 μM , at which the protoxin is soluble over the entire pH range from 7 to 10.

No cysteine residues are present in the toxin, and therefore the reactivity of the thiol groups reflects properties of the C-terminal half of the protoxin molecule, which is removed during the activation process. A non-linear least-squares regression (Enzfitter; Biosoft, Milton, NJ, U.S.A.) of the reactivity data gives a titration curve with a pK of 8.45 ± 0.13 and a reactivity (r) 39.6 ± 2.7 times that of the α -amino group of alanylalanine (Fig. 6). The protoxin has 16–18 residues of cysteine per molecule, and the fact that the average reactivity of these residues can be fitted to a regular titration curve with a small standard error in pK and reactivity indicates that the thiol groups in the solubilized protoxin have similar environments, and hence similar pK values and reactivities. The average pK value obtained for the thiol groups is close to the pK value of approx. 8.5 observed for exposed thiol groups in model compounds and proteins (Shaked *et al.*, 1980). This result, along with the high reactivity observed, strongly suggests that the majority if not all the thiol groups are fully exposed on the surface of the protoxin.

In order to quantify the extent of exposure of the thiol groups, the native solubilized protoxin isolated from the *kurstaki* HD-1 crystal was treated with iodoacetate and iodoacetamide in the absence of any denaturant. The following CM-cysteine content was obtained on amino acid analysis: (a) with iodoacetate, 16.5 mol of CM-cysteine/mol of protoxin; (b) with iodo-

acetamide, 16.6 mol of CM-cysteine/mol of protoxin. Therefore in the native protoxin the thiol groups can be quantitatively modified, which supports the conclusion that all these groups are exposed on the surface of the protoxin.

Conclusion

In the protein crystal from *B. thuringiensis* subsp. *kurstaki* HD-1 and subsp. *entomocidus*, the cysteine residues form predominantly symmetrical interchain disulphide bridges. The thiol groups of the solubilized 130 kDa crystal protein (protoxin) are exposed on the surface of the molecule, and the disulphide linkages that are formed in the crystallization process are also readily accessible. A mechanism that specifically orients the protein so that a regular disulphide-linked crystal lattice is formed must be involved. As there is evidence that cloned protoxin forms crystals in *Bacillus subtilis* (Calogero *et al.*, 1989), a specific association of the protoxin molecules may be sufficient for the formation of the symmetrical interchain disulphide bridges. The exposed disulphide linkages in the crystal protein make them susceptible to alkaline cleavage, which appears to be important in the generation of toxin in the highly alkaline (pH 10–10.5) lepidopteran insect gut (Jaquet *et al.*, 1987).

Comparison of the gene nucleotide sequences of the known *B. thuringiensis* genes (Höfte & Whitely, 1989) shows that the cysteine-containing regions of the protoxin are highly conserved. In families of proteins where such conservation is observed the precise structure is usually essential for biological activity. Höfte & Whitely (1989) have speculated that the structure in the C-terminal half is conserved because it is required for crystal formation. The present results, which show that all the thiol groups are located on the surface of the protoxin and form symmetrical interchain disulphide bridges, support the hypothesis that a precise orientation of the thiol groups is essential for crystal formation and ultimately its insecticidal activity.

This work was supported by a strategic grant from the Natural Sciences and Engineering Research Council of Canada. We thank Dr. Paul Fast for his advice and support for this work, Tim Lessard for preparation of the purified crystals and Howard Lazarus and Jim Webster for their assistance in carrying out the experimentation.

REFERENCES

- Adang, M. L., Staver, M. J., Rocheleau, T. A., Leighton, J., Baker, R. F. & Thompson, D. V. (1985) *Gene* **36**, 289–300.
- Andrews, R. E., Jr., Faust, R. M., Wabiko, H. & Raymond, K. C. (1987) *CRC Crit. Rev. Biotechnol.* **6**, 163–232.
- Aronson, A. I. & Arvidson, H. C. (1987) *Appl. Environ. Microbiol.* **53**, 416–421.
- Aronson, A. I., Beckman, W. & Dunn, P. (1986) *Microbiol. Rev.* **50**, 1–24.
- Bietlot, H., Carey, P. R., Choma, C., Kaplan, H., Lessard, T. & Pozsgay, M. (1989) *Biochem. J.* **260**, 87–91.
- Brousseau, R. & Masson, L. (1988) *Biotechnol. Adv.* **6**, 697–724.
- Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228.
- Calogero, S., Albertini, A. M., Fogher, C., Marzari, R. & Galizzi, A. (1989) *Appl. Environ. Microbiol.* **55**, 446–453.
- Carey, P. R., Fast, P., Kaplan, H. & Pozsgay, M. (1986) *Biochim. Biophys. Acta* **872**, 169–176.
- Couche, G. A., Pfannenstiel, M. A. & Nickerson, K. W. (1987) *J. Bacteriol.* **169**, 3281–3288.
- Creighton, T. E. (1986) *Methods Enzymol.* **131**, 83–106.
- Creighton, T. E. (1989) in *Protein Structure: A Practical Approach* (Creighton, T. E., ed.), pp. 147–172, IRC Press, Oxford.
- Darbre, A. (1986) in *Practical Protein Chemistry* (Darbre, A., ed.), pp. 230–325, John Wiley and Sons, New York.
- Dastidar, P. G. & Nickerson, K. W. (1979) *FEBS Lett.* **108**, 411–414.
- Geiser, M., Schweitzer, S. & Grimm, C. (1986) *Gene* **48**, 109–118.
- Hefford, M. A., Evans, R. M., Oda, G. & Kaplan, H. (1985) *Biochemistry* **24**, 867–874.
- Hirs, C. H. W. (1956) *J. Biol. Chem.* **219**, 611–621.
- Höfte, H. & Whitely, H. R. (1989) *Microbiol. Rev.* **53**, 242–255.
- Jaquet, F., Hütter, R. & Lüthy, P. (1987) *Appl. Environ. Microbiol.* **53**, 500–504.
- Kaplan, H., Stevenson, K. J. & Hartley, B. S. (1971) *Biochem. J.* **124**, 289–299.
- Nickerson, K. W. (1980) *Biotechnol. Bioeng.* **12**, 1305–1335.
- Pozsgay, M., Fast, P., Kaplan, H. & Carey, P. R. (1987) *J. Invertebr. Pathol.* **50**, 246–253.
- Rice, R. H. & Means, G. E. (1971) *J. Biol. Chem.* **246**, 831–832.
- Schnepf, H. E. & Whitely, H. R. (1985) *J. Biol. Chem.* **260**, 6273–6285.
- Shaked, Z., Szajewski, P. & Whitesides, G. (1980) *Biochemistry* **19**, 4156–4166.
- Yamamoto, T. & McLaughlin, R. E. (1981) *Biochem. Biophys. Res. Commun.* **103**, 414–421.
- Young, N. M. & Kaplan, H. (1989) in *Protein Structure and Function: A Practical Approach* (Creighton, T. E., ed.), pp. 225–244, IRC Press, Oxford.

Received 3 August 1989/26 October 1989; accepted 3 November 1989