Volume 215, number 2, 285-290 FEB 04674 May 1987

The amino acid sequence of the aliphatic amidase from *Pseudomonas aeruginosa*

R.P. Ambler, A.D. Auffret* and P.H. Clarke+°

Department of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland and +*Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, England*

Received 22 January 1987; revised version received 19 February 1987

Amino acid sequence studies show that the aliphatic amidase (EC 3.5.1.4) from *Pseudomonas aeruginosa* PAC142 consists of a single polypeptide chain of 346 residues, giving an M_r of 38400. The evidence from the amino acid studies is in complete agreement with that deduced from the DNA sequence of the *amiE* gene. Studies of the protein from *Pseudomonas putida* A87 show that it differs from the *Ps. aeruginosa* protein by about 30 amino acid substitutions. It now becomes possible to relate changes in the enzyme which result in altered specificity to structural changes in the protein.

Aliphatic amidase; Primary structure; *(Pseudomonas aeruginosa)*

1. INTRODUCTION

The aliphatic amidase (EC 3.5.1.4) of *Pseudomonas aeruginosa* enables the organism to use acetamide as both a carbon and a nitrogen source. The system has been under investigation for many years [1,2]. The control of enzyme synthesis, the substrate specificity, and the subunit structure have all been well studied, and the system has been used for exploring protein evolution by seeing how mutation and selection can be used to produce enzymes with new substrate specificities [3]. To complement these studies, the amino acid sequence of the wild-type enzyme has been under investigation since 1973 [4]. Most of the sequence has been known since 1978, but it proved very difficult to get definitive evidence to establish the connexions between the three main blocks of sequence. The *amiE* gene, which codes for the enzyme, has been cloned [5], and the N-terminal sequence of the protein aligned within the cloned fragment by restriction mapping [6]. Work then started on sequencing the DNA of the gene, and we are happy that by the combination of evidence from amino acid and nucleotide sequencing we have now established the primary structure of the protein. The evidence and results of the DNA experiments are reported in the adjoining paper [7].

Correspondence address: R.P. Ambler, Dept of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland

* Present address: LKB Biochrom Ltd., Cambridge Science Park, Milton Road, Cambridge CB4 4BH, England

° Present address: Department of Chemical and Biochemical Engineering, University College London, Torrington Place, London WC1E 7JE, England

2. EXPERIMENTAL

Amidase was isolated from *Ps. aeruginosa* PAC142 (L10R) [3] and from *Ps. putida* A87 [8,9] as described previously [2,10]. The amino acid sequences were investigated by the methods we have used for the study of bacterial cytochromes *c* [11,12], blue copper proteins [13] and β-lactamases [14]. The amidases contain several cysteine residues [2], and satisfactory proteolytic digestion was not obtained unless the protein was first ox-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

idized with performic acid, or aminoethylated under reducing conditions [13]. The sequence of one difficult peptide was determined in a gas-phase sequencer. Early attempts to degrade the intact protein in a spinning-cup liquid-phase sequencer were unrewarding, as the protein did not form a satisfactory film.

3. RESULTS

The evidence for the amino acid sequence proposed for *Ps. aeruginosa* amidase is summarized in fig.1. The amino acid composition deduced from the sequence (table 1) is in good agreement with previous estimates from analyses of hydrolysates [2,4] when the old results are recalculated to fit a 346 residue protein. The differences that have been detected between the *Ps. aeruginosa* and the *Ps. putida* proteins are listed in table 2.

The initial experiments [4] determined the sequences of many of the small peptides formed by enzymic hydrolysis and CNBr cleavage of the *Ps. aeruginosa* protein, but in each digest much of the material was in large insoluble fragments that could not be purified. These experiments were followed by what was intended as a complementary study of the *Ps. putida* protein, and the problems associated with the insoluble core were overcome by using thermolysin and pepsin for primary digests. In the course of these experiments it was eventually realized that the large batch of *Ps. putida* amidase used for most of the work was contaminated with about 20% of *Ps. aeruginosa* enzyme. We presume that this must have come about by a batch of *Ps. aeruginosa* cells having been inadvertantly pooled with the several batches of *Ps. putida* cells needed for the large preparation. The two amidases are of very similar size and charge, and would have copurified. The contamination would not have been detected by the criteria of purity used for amidase preparations [10]. The problem was eventually recognized when several cases were found of digests containing two homologous but different peptides covering the same region of sequence, with the sequence present in low yield being already known to occur in *Ps. aeruginosa*. The amount of sequence information derived from these digests was extensive, and enabled a better strategy to be used for the next round of experiments on pure *Ps. aeruginosa* amidase. They were also complete enough so that when the final *Ps. aeruginosa* sequence was available (fig.1), the *Ps. putida* differences shown in table 2 could be deduced.

Limited experiments with a 'pure' batch of *Ps. putida* amidase did not result in the detection of any of the characteristic *Ps. aeruginosa* peptides found in the earlier experiments. In particular the two C-terminal *Ps. aeruginosa* tryptic peptides (EA and LPYEGLEK) are very easily recognized, and quite different to the peptides from the proposed *Ps. putida* sequence, and no trace of them was seen in the digest of the 'pure' batch of protein. These results make it unlikely that the earlier results were the effect of *Ps. putida* possessing two amidase genes, one of which is very similar to that of *Ps. aeruginosa*.

The *Ps. aeruginosa* sequence was finally derived by consideration of the peptides isolated from five separate enzymic digests. These were (a) 3.7 μmol oxidized protein digested with thermolysin, (b) 3.7 μmol oxidized protein digested with chymotrypsin, (c) 3.6 μmol oxidized protein digested with staphylococcal protease, (d) 4.5 μmol oxidized protein digested with pepsin, (e) 4.5 μmol aminoethylated protein digested with trypsin. The protein contains 15 methionine residues, but the preliminary experiments had shown that their distribution was unfavourable for using CNBr cleavage to order sets of overlapping peptides. There are two -Met-Met- sequences, as well as other looser clusters of methionine residues (fig.1),

Fig.1. Amino acid sequence of *Pseudomonas aeruginosa* amidase. Peptides derived by staphylococcal protease (F), thermolysin (H) or trypsin (T) digestion are shown above the sequence, and by chymotrypsin (C) or pepsin (P) digestion below the sequence. Solid lines indicate peptides that have been purified and analysed quantitatively. Substandard analyses [11] are marked *, and particularly bad ones ** (0.3–0.45 mol/mol of an unexpected amino acid, usually glycine and serine). A second line under the peptide line indicates sequence determination by phenyl isothiocyanate degradation, with this second line dashed when the degradation results were substandard. The 'difficult peptide' (residues 135–151), marked §, was degraded in a gas-phase sequencer, and is shown as ⟨═══⟩. Peptides examined by carboxypeptidase digestion are marked †. Peptides marked ¶ were also sequenced by mass spectroscopy [20].

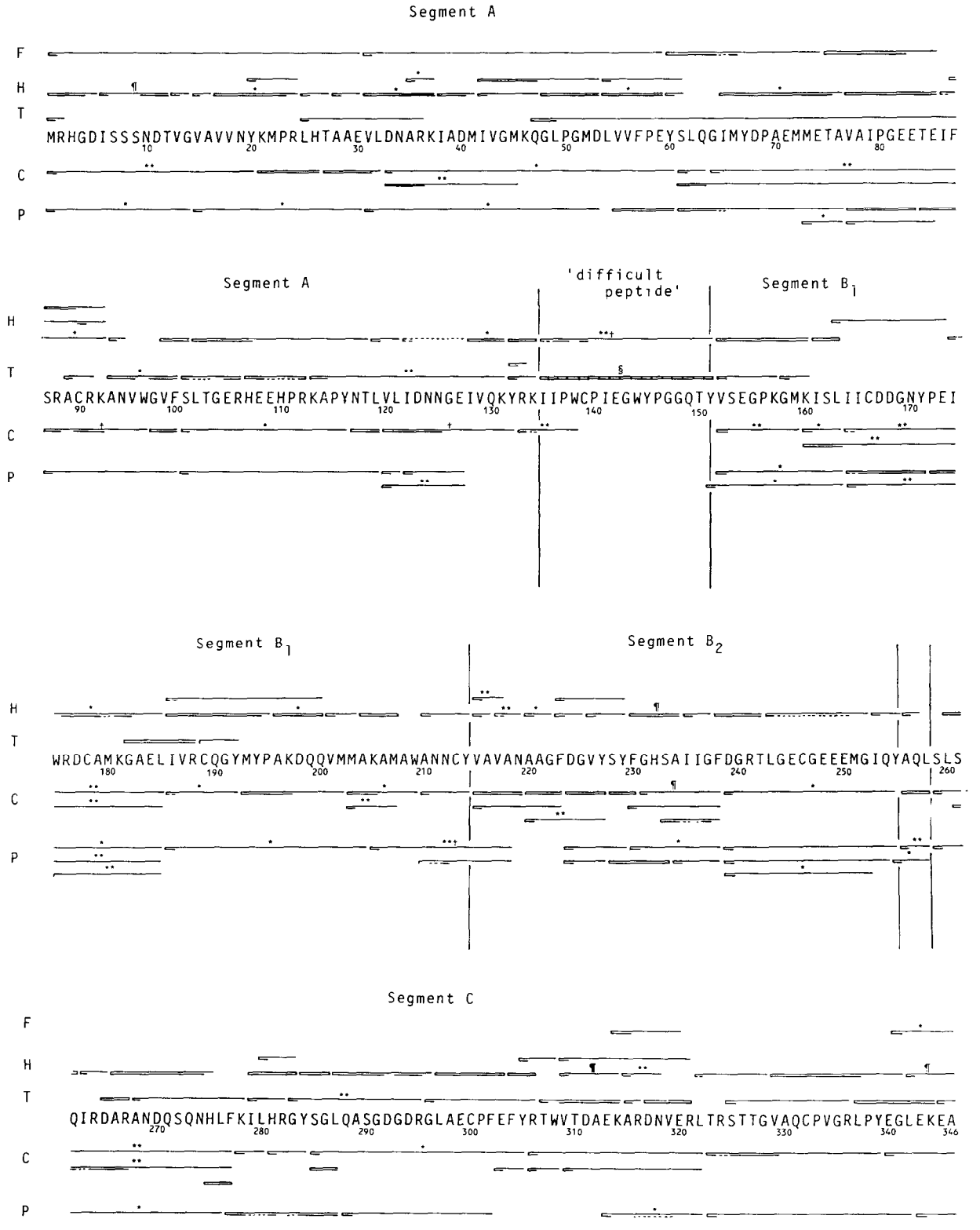
Segment A
F
H
T
MRHGDISSSNDTVGVAVVNYKMPRLHTAAEVLDNARKIADMIVGMKQGLPGMDLVVFPEYSLQGIMYDPAEMMETAVAIPGEETEIF
10 20 30 40 50 60 70 80
C
P
Segment A
'difficult peptide'
Segment B1
H
T
SRACRKANVWGVFSLTGERHEEHPRKAPYNTLVLIDNNGEIVQKYRKIIPWCPIEGWYPGGQTYVSEGPKGMKISLIICDDGNYPEI
90 100 110 120 130 140 150 160 170
C
P
Segment B1
Segment B2
H
T
WRDCAMKGAELIVRCQGYMYPAKDQQVMMAKAMAWANNCYVAVANAAGFDGVYSYFGHSAIIGFDGRTLGECGEEEMGIQYAQLSLS
180 190 200 210 220 230 240 250 260
C
P
Segment C
F
H
T
QIRDARANDQSQNHLFKILHRGYSGLQASGDGDRGLAECPFEFYRTWVTDAEKARDNVERLTRSTTGVAQCPVGRLPYEGLEKEA
270 280 290 300 310 320 330 340 346
C
P

Table 1

The amino acid composition of *Pseudomonas aeruginosa* amidase

	Analyses		Sequence
	A [2]	B [4]	C
Glycine	33.9	34.9	34
Alanine	32.2	31.3	32
Valine	25.2	22.8	23
Leucine	20.8	20.1	20
Isoleucine	22.2	20.6	21
Serine	15.6	15.9	16
Threonine	12.9	12.7	13
Aspartic acid	33.3	33.7	19
Asparagine			13
Glutamic acid	36.8	41.1	26
Glutamine			15
Phenylalanine	8.8	9.0	9
Tyrosine	16.8	16.5	17
Tryptophan	6.0	*	6
Cysteine	8.7	*	9
Methionine	14.3	15.6	15
Proline	15.9	16.3	17
Lysine	15.6	14.0	15
Histidine	6.9	6.6	7
Arginine	19.8	18.9	19
Total	345.7	330.0 (+15)	346
Figure of merit	3.8%	2.5%	

The values shown in columns A and B were derived from multiple analyses of samples of hydrolysed protein, with corrections for non-hydrolysis of stable peptide bonds and for destruction during hydrolysis. They have been recalculated to fit best the length of the polypeptide chain (346 residues) deduced from the amino acid sequence. The values shown in column C are derived from the sequence given in fig.1. The figures of merit for analyses A and B are the average percentage deviation from the sequence values, for all the amino acids except cysteine and tryptophan, and are in the range achieved for careful analyses of pure proteins [19]. The best figures of merit achieved by the methods employed for analyses A and B are in the range 2–2.5% [19]

and the small peptides formed by CNBr cleavage of the protein had already been largely characterized [4]. However, the larger predicted peptides (of 96, 86 and 44 residues) could not be isolated in a satisfactorily pure state by the methods attempted.

From peptide experiments alone, we were satisfied (fig.1) about the sequence of the N-terminal segment A (residues 1–134), and of the C-terminal of segment C (residues 259–346). The segments B_1 (151–214) and B_2 (215–255) had also been deduced, although there was not good enough evidence to establish their relative order. A 'difficult peptide' (135–150) had also been recognized in several of the digests, and assigned to the A–B gap by reason of its C-terminal tyrosine residue. No satisfactory evidence for overlaps or for its complete sequence were obtained by the conventional methods used, as the combination of tryptophan, proline and cysteine residues prevented subdigestion to readily characterizable small peptides. The sequence of this region was finally determined at the amino acid level by the degradation of peptide (135–150), derived from tryptic digestion of aminoethylated protein, in a gas-phase sequencer.

As the DNA sequence results emerged, most of these hypotheses were confirmed. The 'difficult peptide' did fit between segments A and B, but our tentative ordering of B_2 before B_1 proved wrong. The link region between segments B_2 and C corresponded to the last region of DNA sequence to be rigorously determined, and the first 'translation', before both strands had been sequenced, did not show residues 258–259. However, the protein evidence, although weak, definitely existed, and the final DNA sequencing experiments confirmed the peptide abutments in this region.

4. CONCLUSIONS

The amino acid sequence of *Ps. aeruginosa* amidase consists of a single polypeptide chain of 346 residues, giving an M_r of 38400. The previous estimate of the M_r value, from SDS-PAGE, was 33000 [2]. There is complete agreement between the amino acid sequence evidence and that derived from the DNA sequence of the *amiE* gene [7]. The N-terminus of the protein as isolated is the initiating methionine residue.

About 85% of the sequence of the amidase from *Ps. putida* A87 has been determined, and a further 10% accounted for in well-characterized peptides. The residual 5% forms the 'difficult peptide' of fig.1. There are probably 32 differences from the *Ps. aeruginosa* sequence in the 95% of the

Table 2

Differences between the amino acid sequences of *Pseudomonas putida* and *Pseudomonas aeruginosa* amidases

Ps. putida	Residue no.	*Ps. aeruginosa*
Arg (A)	28	Ala
Asp (A)	29	Ala
Glu (A)	33	Asp
Ile (A)	41	Met
Ile (A)	56	Val
Val (A)	82	Glu
Val (C)	86	Ile
Ala (A)	88	Ser
Gln (A)	89	Arg
Ser (B)	114	Ala
Asx (D)	~147	Gly
Ala (A)	153	Ser
Asp (A)	154	Glu
Leu (B)	159	Met
Met (B)	163	Leu
Leu (A)	202	Met
Ile (B)	260	Leu
Gln (A)	268	Ala
Tyr (A)	277	Phe
Met (A)	287	Leu
His (A)	288	Gln
Leu (A)	289	Ala
Ile (A)	297	Leu
Leu (A)	310	Thr
Glu (A)	317	Asp
Ala (A)	321	Arg
Arg (B)	325	Ser
Ser (A)	327	Thr
Ser (B)	336	Arg
Asp (A)	340	Glu
Ala (B)	341	Gly
Lys (A)	342	Leu
Arg (A)	347	–

The strength of the evidence for the amino acid residue proposed for the *Ps. putida* sequence is assessed, on a scale from A to D. (A) The sequence around the residue has been positively established, and the residue occurs in one or more peptides that are longer than three residues, and which have been analysed satisfactorily for amino acid content. (B) The residue has been located in a peptide longer than three residues which has been analysed satisfactorily for amino acid content, but the sequence evidence is indirect. (C, residue 86) The residue was identified in a specific and predicted dipeptide, but there was no direct evidence except 'homology' to locate this dipeptide in the sequence. (D, residue ~147) There are suggestions from amino acid analyses of impure peptide fractions that the asparagine or aspartic acid content of the region in *Ps. putida* is higher than of the corresponding region in the *Ps. aeruginosa* protein, and that the glycine content is lower. There may well be further differences in this 'difficult peptide' (residues 135–150), but further differences in the remainder of the sequence are considered unlikely

molecule that has been characterized, and indications of some further differences in the 'difficult peptide'. The changes are listed in table 2. The amidases from *Ps. aeruginosa* and *Ps. putida* are very similar, although not identical, in substrate specificities and physicochemical properties [9], and differences are found in the enzymes isolated from different natural isolates. However, the 9% difference in sequence between homologous proteins from two very similar species of pseudomonads [8,15] is a demonstration of the large amount of divergence that takes place in prokaryotic speciation.

Studies of *Ps. aeruginosa* amidase have shown that what are probably single site mutations can result in changes in the substrate specificity of the enzyme without significant reduction in the catalytic activity [3,16]. The amino acid lesions resulting from two of these mutations have been characterized. The mutation *amiE56* results in an enzyme which, unlike the wild type, can hydrolyse acetanilide, causes a substitution of isoleucine for threonine [17], and this can now be seen to be at position 103. The mutant *amiE16* [10,16] enables the organism that contains it to use butyramide, and the lesion has been identified as a substitution of phenylalanine for serine at position 7 [18]. Despite considerable effort, the lesions caused by mutations *amiE30* and *amiE78* have not yet been located, even though about 65% of the protein from strain PAC388 (PhV1) [3,16], which contain them, has been sequenced, and most of the remainder of the molecule identified in analysed peptides.

The *Pseudomonas* amidase system retains very great potential for exploring and understanding the mechanisms of the evolution of novel enzyme activities [21]. Although the protein does not peptide map well, the lesions caused by selected mutations can now be readily located by DNA sequencing, and particular residues will be easily altered by site-directed mutagenesis. The greatest

present shortcoming of the system is the lack of knowledge about the three-dimensional structure of amidase, as unfortunately all attempts to crystallize the enzyme have so far failed.

ACKNOWLEDGEMENTS

This work was supported by grants from the MRC to R.P.A., and by a studentship from the SERC to A.D.A. We thank Mrs M. Daniel and Mrs L. McLellan for skilled and enthusiastic technical assistance, and Professor W.J. Brammar for his help and exchange of information throughout this long investigation. We are grateful to B. Dunbar and J.E. Fothergill (Aberdeen) for the automatic sequencer analysis of the 'difficult peptide'.

REFERENCES

[1] Brammar, W.J. and Clarke, P.H. (1964) J. Gen. Microbiol. 37, 307–319.
[2] Brown, P.R., Smyth, M.J., Clarke, P.H. and Rosemeyer, M.A. (1973) Eur. J. Biochem. 34, 177–187.
[3] Clarke, P.H. (1974) Symp. Soc. Gen. Microbiol. 24, 183–217.
[4] Auffret, A.D. (1976) PhD Thesis, University of Edinburgh.
[5] Drew, R.E., Clarke, P.H. and Brammar, W.J. (1980) Mol. Gen. Genet. 177, 311–320.
[6] Clarke, P.H., Drew, R.E., Turberville, C., Brammar, W.J., Ambler, R.P. and Auffret, A.D. (1980) Biosci. Rep. 1, 299–307.
[7] Brammar, W.J., Charles, I.G., Matfield, M., Cheng-Pin, L., Drew, R.E. and Clarke, P.H. (1987) FEBS Lett. 215, 291–294.
[8] Stanier, R.Y., Palleroni, N.J. and Doudoroff, M. (1966) J. Gen. Microbiol. 43, 159–271.
[9] Clarke, P.H. (1972) J. Gen. Microbiol. 71, 241–257.
[10] Brown, J.E., Brown, P.R. and Clarke, P.H. (1969) J. Gen. Microbiol. 57, 273–285.
[11] Ambler, R.P. and Wynn, M. (1973) Biochem. J. 131, 485–498.
[12] Ambler, R.P., Daniel, M., Melis, K. and Stout, C.D. (1984) Biochem. J. 222, 217–227.
[13] Ambler, R.P. and Tobari, J. (1985) Biochem. J. 232, 451–457.
[14] Ambler, R.P. (1975) Biochem. J. 151, 197–218.
[15] Champion, A.B., Barrett, E.L., Palleroni, N.J., Soderberg, K.L., Kunisawa, R., Contoupoulou, R., Wilson, A.C. and Doudoroff, M. (1980) J. Gen. Microbiol. 120, 485–511.
[16] Betz, J.L., Brown, P.R., Smyth, M.J. and Clarke, P.H. (1974) Nature 247, 261–264.
[17] Brown, P.R. and Clarke, P.H. (1972) J. Gen. Microbiol. 70, 287–298.
[18] Paterson, A. and Clarke, P.H. (1979) J. Gen. Microbiol. 114, 75–85.
[19] Ambler, R.P. (1981) in: Amino Acid Analysis in Clinical Chemistry and Medical Research (Rattenbury, J. ed.) pp.119–137, Ellis Horwood, Chichester.
[20] Auffret, A.D. and Williams, D.H. (1979) J.C.S. Chem. Commun. 16, 692–693.
[21] Mortlock, R.P. (1984) Microorganisms as Model Systems for Studying Evolution, Plenum Press, New York.