A novel and comprehensive synthetic approach for the elucidation of protein antigenic structures

Determination of the full antigenic profile of the α -chain of human haemoglobin

A. Latif KAZIM and M. Zouhair ATASSI*

Department of Immunology, Mayo Medical School, Rochester, MN 55901, U.S.A.

(Received 3 June 1980/Accepted 30 June 1980)

A comprehensive synthetic approach for the determination of continuous antigenic sites of proteins is presented. This approach consists of the synthesis of a series of consecutive overlapping peptides that, together, systematically represent the entire primary structure of the protein under study. Its application to the α -chain of human haemoglobin afforded, for the first time, a full profile of immunochemically active α -chain peptides and enabled the localization of all the major continuous antigenic sites of this haemoglobin subunit.

A key approach in the strategy set forth for the delineation of protein antigenic structures (Atassi, 1972) consists of immunochemical studies of overlapping peptide fragments obtained by chemical and enzymic cleavage of the parent protein. This approach, when systematically employed, can yield valuable information on the number and locations of antigenic sites, and their relative contributions to the total protein activity. In practice, however, this approach has severe shortcomings (for review see Atassi, 1977). These include the relatively limited number of reproducible cleavage procedures for proteins, the inadvertent scission within an antigenic site usually affording inactive or only slightly active peptides, the chemical modification frequently occurring at internal locations within a chemical-cleavage fragment and any traces of contamination of an isolated peptide. Furthermore, the distribution of residues whose peptide bonds are potentially susceptible to cleavage in the protein does not permit the desired overlapping peptides to be obtained. Accordingly, the strategy developed (Atassi, 1972, 1975) for the determination of protein antigenic structures employed five distinct chemical approaches, none of which is by itself capable of yielding the antigenic structure. The findings from one approach were used to confirm and/or correct those from the others. Such strategy was intricate and required considerable chemical expertise, patience and perseverance.

In order to circumvent these problems and render the determination of protein antigenic structures

more feasible within a reasonable time, we have devised a novel and comprehensive synthetic approach that is designed to yield the full antigenic profile of the protein under study. This approach consists of the direct synthesis and examination of the immunochemical activities of a series of overlapping peptides that encompass the entire primary structure of the protein, from the beginning to the end (see Fig. 1). The rationale for this approach is presented in further detail in the Discussion section. We have employed this approach to determine the antigenic sites on the α - and β -chains of human haemoglobin (A_1). In this initial report, its application to the α -chain is presented.

Experimental

Human adult haemoglobin was the major chromatographic component prepared as previously described (Atassi, 1964), and the α - and β -chains were prepared as described by Bucci & Fronticelli (1965). These preparations were homogeneous by polyacrylamide-disc-gel electrophoresis. Solid-phase peptide synthesis of the α -chain peptides (see Fig. 1 for the locations of the synthetic peptides in the α -chain) and purification of the synthetic products were performed by methods similar to those previously described (Koketsu & Atassi, 1973, 1974). After purification, each peptide migrated as a single spot on high-voltage paper electrophoresis, and on analysis yielded the expected amino acid composition. It should be noted that, to avoid difficulties associated with the synthesis of tryptophan- and cysteine-containing peptides, trypto-

* To whom correspondence should be addressed.

Fig. 1. Schematic representation of the synthetic strategy for localizing continuous antigenic sites

Here the strategy is shown applied to the α -chain of haemoglobin, depicted by the unbroken line (1–141). The synthetic overlapping peptides shown each consist of 15 residues, except peptide 131–141. Consecutive peptides overlap by five residues. Numbers refer to the locations of residues in the primary structure of the α -chain. See the text for details.

Fig. 2. Representative example of quantitative immuno-adsorbent titrations of ^{125}I -labelled immune immunoglobulin G with adsorbents of the synthetic α -chain peptide

A fixed amount of ^{125}I -labelled immune immunoglobulin G (248 000 c.p.m.) was titrated with various amounts of peptide–Sephacrose and control adsorbents as described in the text. The values shown have been corrected for non-specific binding. Haemoglobin and its α -chain (not shown) attained plateau values of 19 278 and 11 721 c.p.m. respectively. The numbers on the curves denote the following peptides: curve 1, α 81–95; curve 2, α 111–125; curve 3, α 121–135; curve 4, α 51–65; curve 5, α 41–55; curve 6, α 101–115; curve 7, α 11–25; curve 8, all other synthetic peptides.

phan-14 was replaced by phenylalanine in peptide α 1–15, and cysteine-104 by serine in peptides α 91–105 and α 101–115. Proteins and peptides were coupled to CNBr-activated Sepharose CL4B as described by March *et al.* (1974). The adsorbents for the proteins had 2–3 mg/ml packed volume and those of the peptides 0.5–1.0 mg/ml packed volume.

Antisera to haemoglobin A₁ were raised in goats as previously described for myoglobin (Atassi, 1967). The immunoglobulin G fractions of these antisera, completely free of haptoglobin, were prepared by $(\text{NH}_4)_2\text{SO}_4$ precipitation (33% saturation) followed by chromatography on DEAE-cellulose (Gray *et al.*, 1969). The pure immunoglobulin G fractions were labelled with ^{125}I by a chloramine-T procedure (Hunter & Greenwood, 1962). The specific radioactivity of the labelled immunoglobulin G was $20 \mu\text{Ci}/\mu\text{g}$.

The immunochemical activities of the protein and peptide fractions were determined by quantitative immuno-adsorbent titrations of fixed amounts of ^{125}I -labelled immune immunoglobulin G and various amounts of peptide or protein adsorbents as previously described for myoglobin antigenic sites (Twining & Atassi, 1979). Non-specific adsorption was determined by titrating equivalent volumes of uncoupled Sepharose and bovine serum albumin–Sepharose under identical conditions.

Results

Titration of ^{125}I -labelled immune immunoglobulin G with immuno-adsorbents of the synthetic peptides (see Fig. 2 for a representative titration) distinguished the peptides into two categories. With some peptides, after correction for non-specific binding, the amounts of radioactivity bound remained constant throughout the entire range of the titration. For these peptides the amounts of ^{125}I -labelled immune immunoglobulin G bound did not exceed 200 c.p.m. (Fig. 2 and Table 1). These

peptides were considered to be inactive under these experimental conditions with these antisera. Alternatively, other peptides bound increasing amounts of radioactivity that eventually levelled off at 'plateau' or maximum values. Indeed, in certain instances (Fig. 2 and Table 1) these peptides accounted for a substantial amount of the radioactivity of the α -chain itself. It should be noted that the titration curves (not shown) for haemoglobin A and its α -chain rise more sharply and attain maximal values at lower amounts of adsorbents than do those for the peptides. The reasons for the diminished affinity of an active peptide on coupling to Sepharose have been discussed (Kazim & Atassi, 1977; Twining & Atassi, 1979).

The percentage of radiolabelled antibodies bound by saturating amounts of peptide immuno-adsorbents (i.e. plateau value), relative to the amounts bound by α -chain–Sepharose in the plateau, are given in Table 1 for a goat antiserum to haemoglobin. The calculated sum of the amounts of radioactivity bound by the active peptides (Table 1) is significantly larger than the amount bound by the α -chain itself. However, these results must be considered in the context of the structures of the

Table 1. *Binding of goat anti-haemoglobin antibodies by the synthetic overlapping peptides of the α -chain*

The data here are a summary of the plateau values of the binding obtained from the titration shown in Fig. 2. In addition, the results are expressed as a percentage of the maximum binding by each peptide adsorbent relative to that of the α -chain as 100%.

Protein or peptide adsorbent	^{125}I -labelled immune immunoglobulin G bound (c.p.m.)	Relative reaction (%) (α -chain = 100%)
Haemoglobin	19 278	(164.5)
α -Chain	11 721	100
1-15	<200	<1.7
11-25	670	5.7
21-35	<200	<1.7
31-45	<200	<1.7
41-55	1455	12.4
51-65	1678	14.3
61-75	<200	<1.7
71-85	<200	<1.7
81-95	3958	33.8
91-105	<200	<1.7
101-115	1011	8.6
111-125	3303	28.2
121-125	3043	26.0
131-141	<200	<1.7

peptides and their overlaps (see Fig. 1), which were designed to detect antigenic sites located in the overlapping regions between adjacent peptides (see the Discussion section). Thus adjacent peptides that are both active may in fact carry all (or part) of the same antigenic site, and their activities therefore will reflect identical specificities.

Discussion

The elucidation of the complete antigenic structures of myoglobin (Atassi, 1975) and lysozyme (Atassi, 1978) has shown that such studies are extremely complex and require a strategy that consists of several different chemical approaches (for review see Atassi, 1977). In spite of the complexity of each chemical approach, however, the molecular details of the antigenic sites of myoglobin and lysozyme thus obtained have been immensely valuable in furthering our comprehension of immune recognition processes. The antigenic sites of myoglobin and lysozyme consist of discrete surface portions of their respective native molecules. Each site comprises five to seven amino acid residues that come into close proximity either through direct peptide linkage, such as the 'continuous' antigenic sites of myoglobin, or in a more complex manner, through polypeptide-chain folding, such as the 'discontinuous' sites of lysozyme. In view of the structural alternatives in their architecture (continuous or discontinuous sites; Atassi & Smith, 1978), the determination *ab initio* of the antigenic structures of other proteins should anticipate both

alternatives. The logical progression, therefore, is first to localize and characterize the more readily identifiable continuous antigenic sites, with attention later paid to discontinuous sites.

Any approach that could facilitate the localization of continuous antigenic sites would certainly diminish the overall effort required for complete immunochemical characterization. In examining the antigenic structure of myoglobin (Atassi, 1975), it occurred to us that it could have been arrived at more rapidly by employing a synthetic strategy, systematically making segments of the protein from the beginning to the end, each segment having an appropriate overlap both with the segment preceding it and the one following it. Among the advantages that can be realized from the synthesis of peptide fragments are high yields, predetermined structures, the absence of contamination by peptides representing extraneous parts of the protein, the sizes of the peptides can be controlled so that antigenic sites can be studied individually, and the extent of the overlaps can be designed to increase the likelihood that antigenic sites lying in the overlapping regions would be detected.

The approach that we have developed and applied to haemoglobin therefore entails the systematic synthetic dissection of the protein into a series of overlapping peptides (Fig. 1). The sizes of the peptides should at the same time be consistent with the number of residues expected for an antigenic site [e.g. five to seven residues for myoglobin (Atassi, 1975) and lysozyme sites (Atassi, 1978)] as well as having a high probability of possessing most (or all)

of a site possibly lying in the overlap regions between adjacent peptides. Peptides consisting of fifteen residues and overlapping at their N- and C-termini with adjacent peptides by five residues were considered adequate for these purposes (Fig. 1). It should be mentioned that this approach is restricted to the localization of continuous antigenic sites (e.g. myoglobin antigenic sites; Atassi, 1975). Discontinuous antigenic sites (such as those of lysozyme; Atassi, 1978) will not be detected unless a given peptide possesses residues of that site which make the major contributions to its binding energy.

The application of this strategy to the α -chain of haemoglobin required the synthesis of fourteen peptides (see Fig. 1) encompassing the entire structure of the α -chain (from residue 1 to 141). Each peptide consisted of fifteen residues except peptide α 131–141, and overlapped adjacent peptides by five residues. The immunochemical activities of these peptides were determined by immunoadsorbent titration studies with rabbit and goat antisera to native haemoglobin A. Such titrations have been found to be particularly useful for detecting relatively weak interactions such as those expected (Atassi & Saplin, 1968) between small and generally unfolded peptides and their antibodies (Twining & Atassi, 1979). These studies enabled us, for the first time, to identify the full profile of immunochemically active haemoglobin α -chain peptides (Fig. 2 and Table 1).

The results reconfirm the presence of an antigenic site on the α -chain (peptide α 11–25), which was previously localized (Kazim & Atassi, 1977) with antibodies to haemoglobin to the region α 15–23 by extrapolation of the three-dimensional location of a myoglobin antigenic site (site 1, region 15–22). The present findings have also enabled the localization of at least four additional α -chain antigenic sites occurring respectively in the regions α 41–65, α 81–95, α 101–115 and α 111–135. If adjacent peptides with similar reactivities are considered to bind to the same subpopulations of antibodies, then these studies may be further localized to overlapping regions. Of course, such assumptions cannot be made arbitrarily and must be experimentally verified by reciprocal adsorption experiments before assignment to the overlapping regions. In any event,

our results have enabled for the first time the localization of all the major continuous (represented by more than 1–2% of the antibodies) antigenic sites of the α -chain of haemoglobin. The application of this synthetic approach to the α -chain has been successful and is advocated as a convenient and comprehensive strategy for localizing antigenic sites on proteins of known primary structure.

This work was supported by a grant (AM18920) from the Institute of Arthritis and Metabolic Diseases, National Institutes of Health, U.S. Public Health Service. A. L. K. is supported by a Postdoctoral Fellowship from the Muscular Dystrophy Association.

References

- Atassi, M. Z. (1964) *Biochem. J.* **93**, 189–197
- Atassi, M. Z. (1967) *Biochem. J.* **102**, 478–487
- Atassi, M. Z. (1972) in *Specific Receptors of Antibodies, Antigens and Cells (3rd International Convocation of Immunology)* (Pressman, D., Tomasi, T. B., Grossberg, A. L. & Rose, N. R., eds.), pp. 118–136, Karger, Basle
- Atassi, M. Z. (1975) *Immunochemistry* **12**, 423–438
- Atassi, M. Z. (1977) in *Immunochemistry of Proteins* (Atassi, M. Z., ed.), vol. 1, pp. 1–69, Plenum Press, New York
- Atassi, M. Z. (1978) *Immunochemistry* **15**, 909–936
- Atassi, M. Z. & Saplin, B. J. (1968) *Biochemistry* **7**, 688–698
- Atassi, M. Z. & Smith, J. A. (1978) *Immunochemistry* **15**, 609–610
- Bucci, E. & Fronticelli, C. (1965) *J. Biol. Chem.* **240**, 551–552
- Gray, G. D., Mickelson, M. M. & Crim, J. A. (1969) *Immunochemistry* **6**, 641–644
- Hunter, W. M. & Greenwood, F. C. (1962) *Nature (London)* **194**, 495–496
- Kazim, A. L. & Atassi, M. Z. (1977) *Biochem. J.* **167**, 275–278
- Koketsu, J. & Atassi, M. Z. (1973) *Biochim. Biophys. Acta* **328**, 289–302
- Koketsu, J. & Atassi, M. Z. (1974) *Immunochemistry* **11**, 1–8
- March, S. E., Parikh, I. & Cuatrecasas, P. (1974) *Anal. Biochem.* **60**, 149–152
- Twining, S. S. & Atassi, M. Z. (1979) *J. Immunol. Methods* **30**, 139–151