Glycogen phosphorylase *b*

Description of the protein structure

The cover shows a colour diagram of the glycogen phosphorylase *b* dimer molecule in the T state (low affinity for substrate). One subunit is coloured cyan and the other subunit purple with regions that differ by more than 1 Å in C α positions on the transition between the T to R states (low affinity to high affinity for substrates) coloured red and yellow. The N terminal residues (10–23) and the C terminal residues (837–842) are shown in white. The view is down the 2-fold axis of the dimer. The catalytic site (at the rear of the figure) is marked by PLP (pyridoxal phosphate). The glycogen storage site and the AMP allosteric effector site are also marked. [Ribbon diagram produced with a programme of M. Carson and C. E. Bugg (J. Mol. Graphics (1986) 4, 121–122) by Dr. D. Barford on a Silicon Graphics Iris workstation. From D. Barford, S.-H. Hu and L. N. Johnson, J. Mol. Biol. (1991)]

Glycogen phosphorylase *b*

Description of the protein structure

K R Acharya, D I Stuart, K M Varvill and L N Johnson

Laboratory of Molecular Biophysics
University of Oxford

World Scientific

Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 9128

USA office: 687 Hartwell Street, Teaneck, NJ 07666

UK office: 73 Lynton Mead, Totteridge, London N20 8DH

GLYCOGEN PHOSPHORYLASE *b* : DESCRIPTION OF THE PROTEIN STRUCTURE

Copyright © 1991 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

ISBN 981-02-0540-6

981-02-0541-4 pbk

Printed in Singapore by JBW Printers and Binders Pte. Ltd.

PREFACE

Advances in protein crystallography over the last 25 years have led to the definitive description of more than 100 molecules. Knowledge of the structures of enzymes, allosteric control proteins, viruses and viral proteins, nucleic acid binding proteins, the photo-reaction centre, antibodies and antibody-antigen complexes, hormones, calcium binding proteins and other protein molecules have led to insights into the molecular basis of recognition and biological activity. This basic understanding allows the design of novel therapeutic agents to target specific molecules in the diseased state. The determination of a protein crystal structure can be a lengthy process in which the rate limiting step is the growth of suitable crystals for X-ray work. Recent achievements in NMR technology have allowed the determination of increasing numbers of important protein structures in solution but this method is restricted to proteins of less than 20,000 molecular weight. Advances in DNA technology on the other hand have allowed the rapid determination of protein sequences. The protein structure data base currently (1990) contains about 500 entries (not all of these are individual structures) while the nucleic acid data base contains over 17,000 entries (G.D. Barton, unpublished observation). Methods for the prediction of a protein structure from amino acid sequence have advanced in the last few years especially when common structural features can be recognised from similarity in protein sequences. The most productive methods have utilised rules derived from the analysis of existing protein structures. This monograph describes the determination by X-ray methods and the structural analysis of one of the largest single polypeptide chain proteins, glycogen phosphorylase (subunit molecular weight 97,440). It is shown that the large protein conforms to the rules deduced from smaller protein molecules and provides examples of almost every structural feature. The structure is analysed with special reference to the hydrogen bonding pattern of secondary structural elements, the distributions and interactions of the buried amino acids, the role of the water structure and the details of the subunit interface of the dimer. The results show some unusual secondary structural elements and environments for some side chain atoms. As methods for prediction of protein structures from amino acid sequences advance from assignments of secondary structural elements to predicted tertiary assemblies, knowledge of the preferred and exceptional detailed interactions of the amino acid side chains will need to be incorporated. Glycogen phosphorylase provides a wealth of examples of these interactions.

The crystal structure of T state glycogen phosphorylase b, crystallised in the presence of the weak activator IMP, was refined by restrained crystallographic least squares methods. The resulting fractional agreement between observed and calculated structure factor amplitudes is 0.19 and the rms deviation from ideal bond lengths is 0.02 Å.

Phosphorylase is composed of 52% α helix and 20% β strands arranged in 2 domains with topologies similar to other α/β proteins. There is one π helix (8 residues) and two 3_{10} helices (6 and 12 residues long). These unusual conformations can be rationalised in terms of the buried environment of the π helix and the periodicity in the sequence observed for the 3_{10} regions. Phosphorylase contains examples of almost every type of turn, bend or distortion observed in protein structures including 5 Schellman αC_2 type conformations.

The non-polar core, defined by those residues whose non-polar groups are buried (solvent accessibility $< 1 \text{ \AA}^2$), is comprised of 201 out of the 821 residues located. These residues are distributed at the central regions of each domain and at the domain interface. 47% of the polar or charged residues have their polar or charged groups buried. Buried charged groups exhibit a higher proportion of charge-charge interactions (46%) than exposed charged groups (24%) but several buried charge groups are compensated by interactions solely with neutral polar groups including water and main chain N or O atoms. There are 2 buried carboxyl-carboxylate interactions which is unusual in a protein crystallised at pH 6.7.

Fifty of the 692 waters located in the structure have been identified as buried structural waters. Their role in the contribution to the stability of the structure especially in terms of compensating buried polar groups and filling interstitial cavities is discussed. There are 2 major water filled buried channels in the enzyme and one of these surrounds the phosphate of the cofactor pyridoxal phosphate. The cofactor phosphate has a buried environment. Its contacts with a lysine side chain, 2 main chain NH and a link via 2 water molecules to an acidic residue, Asp 283, provide the basis for rationalisation of the ionisation properties of the cofactor that are crucial for activity.

The subunit-subunit interface of the physiologically active dimer represents 7% of the total area of the individual subunit and comprises $2294 \text{ \AA}^2$. A possible explanation for the low or negligible activity of the monomers is presented. Lattice contacts, heavy atom sites and the location of the disordered N-terminal tail residues are also described.

The coordinates of the phosphorylase b structure have been deposited in the Brookhaven Data Base (Bernstein et al., 1977).

CONTENTS

PREFACE	v
ABBREVIATIONS USED	viii
1. INTRODUCTION	1
2. MATERIALS AND METHODS	4
2.1 Crystallisation	4
2.2 Data Collection	4
2.3 Data Processing	5
2.4 Structure Determination	7
2.5 Analysis of the Structure	12
3. RESULTS	15
3.1 Precision of the Final Model	15
3.2 Global Description of the Structure	20
3.3 Ligand Binding Sites	26
3.4 Main Chain Hydrogen Bonds	27
3.5 α Helices	28
3.5.1 Interactions of charged groups at the termini of α helices	28
3.5.2 Intra-helical main chain-side chain hydrogen bonds	31
3.5.3 Helix distortions	34
3.5.3a Mid-helix distortions	34
3.5.3b Distortions at the N termini of helices	37
3.5.3c Distortions at the C termini of helices	37
3.5.4 3_{10} Helices	39
3.5.5 π Helix	40
3.6 β Structure	43
3.7 Turns	46
3.7.1 γ turns ($O_n \dots N_{n+2}$ hydrogen bonds)	48
3.7.2 $O_n \dots N_{n+3}$ hydrogen bond turns	51
3.8 Charge-Charge Interactions: General	53
3.9 The Interior Core	59
3.9.1 Non-polar groups	59
3.9.2 Buried charged and polar groups	65
3.10 Water Structure	80
3.11 Subunit-Subunit Interactions	97
3.12 Lattice Contacts	102
3.13 Heavy Atom Sites	106
3.14 N Terminal Tail	107
4. DISCUSSION	111
REFERENCES	116

Abbreviations used:

GP_b and GP_a, Glycogen phosphorylase b and a respectively;

PLP, pyridoxal 5'-phosphate;

H2P, heptulose-2-phosphate or 8-1-C-methyl, α-D-glucose 1-phosphate;

PUG, gluconohydroximo-1,5-lactone N-phenylurethane;

EMTS, ethylmercurythiosalicylate;

rms, root mean square;

Merging R is defined: $R_m = \sum_h \sum_i |I(h) - I_i(h)| / \sum_h \sum_k I_i(h)$

where $I_i(h)$ and $I(h)$ are the i th and the mean measurement of the intensity of reflection h , respectively.