216–219 *Nucleic Acids Research, 1998, Vol. 26, No. 1*

The Factor VIII Structure and Mutation Resource Site: HAMSTeRS Version 4

Geoffrey Kemball-Cook*, Edward G. D. Tuddenham and Adam I. Wacey[1]

Haemostasis Research Group, MRC Clinical Sciences Centre, Imperial College Medical School, Hammersmith Hospital, Du Cane Road, London W12 ONN, UK and [1]Computational Biology Unit, Thrombosis Research Institute, Manresa Road, London SW3 6LR, UK

Received October 6, 1997; Accepted October 8, 1997

ABSTRACT

Since 1996 the HAMSTeRS (Haemophilia A Mutation, Search, Test and Resource Site) WWW site has provided an online resource for access to data on the molecular pathology of haemophilia A, replacing previous text editions of the Haemophilia A Database published in *Nucleic Acids Research*. This report describes the continued development of the site (version 4), and in particular the expansion of factor VIII (FVIII) structure-related features. Access to the mutation database itself, both for searching the listings and for submission of new mutations, is via custom-designed forms: more powerful Boolean searches of the point mutations in the database are also available. During 1997 a total of 22 novel missense mutations were reported, increasing the total number of unique variants now described to 252 (238 in exonic sequences and 14 at intronic splice junctions). Currently, a total of 586 individual reports with associated phenotypic data are available for searching by any category including phenotype. The FVIII structure section now includes a download of a FVIII A domain homology model in Protein Data Bank format and a multiple alignment of the FVIII amino-acid sequences from four species (human, murine, porcine and canine) in addition to the virtual reality simulations, secondary structural data and FVIII animation already available. Finally, to aid navigation across this site, a clickable roadmap of the main features provides easy access to the page desired. Our intention is that continued development and updating of the site shall provide workers in the fields of molecular and structural biology with a one-stop resource site to facilitate FVIII research and education. The HAMSTeRS URL is http://europium.mrc.rpms.ac.uk

INTRODUCTION

Coagulation factor VIII (FVIII) is the essential cofactor for the activation of factor X by factor IXa (1). The FVIII gene (*F8*) contains 26 exons and spans 186 kb of DNA (2). Deleterious mutations in the FVIII gene have been demonstrated to reduce either or both activity and circulating plasma level of FVIII protein and thus cause haemophilia A (3) an X-linked bleeding disorder affecting ~1 in 5000 males (4).

As a means of both inviting data submission to centralised resources and distributing information content to users in any location, the World Wide Web (WWW) has become enormously popular and successful during the last 2–3 years. Improvements (both in the quality and availability of the hardware and software required, and in the range of services available within large areas of biological science) have been made such that many academic users now rely on the WWW both to search rapidly and effectively for relevant information and also to broadcast their own data. The main features of WWW resources relevant to their success are rapid access, interactivity and the ability to update data resources on a frequent or real-time basis. Thus, versions of the original printed haemophilia A mutation database (5,6) were updated to WWW versions (7,8) utilising these characteristics to provide an effective vehicle for dissemination and updating of mutation data: now we describe the continued expansion of the resource site dedicated to both molecular biology of the factor VIII gene and the structural basis of FVIII function.

DEVELOPMENT OF THE WWW SITE

Coding and functional system requirements for earlier releases of HAMSTeRS have been described earlier (7,8). Further modifications to the coding and design have been made with the primary goal of improving user-friendliness of the site. Multiple alignments of FVIII amino acid sequences were made using the PILEUP facility from the GCG suite of programs (9) and displayed using BOXSHADE (10): human, murine and porcine sequences have been published (11–13) while the canine FVIII sequence was kindly supplied by Dr David Lillicrap (Queen's University, Kingston, Ontario) prior to publication.

For best results it is recommended that the site is accessed through Netscape V3.0 (Netscape URL: http://home.netscape.com) or later: earlier versions or other browsers such as Internet Explorer may give unpredictable results.

Hamsters is served from *europium*, a Silicon Graphics Indigo[2] workstation running IRIX 6.2 and the CERN WWW server software httpd (V3.0, obtainable as shareware from CERN at

*To whom correspondence should be addressed. Tel: +44 181 383 8253; Fax: +44 181 383 8273; Email: gkemball@hgmp.mrc.ac.uk

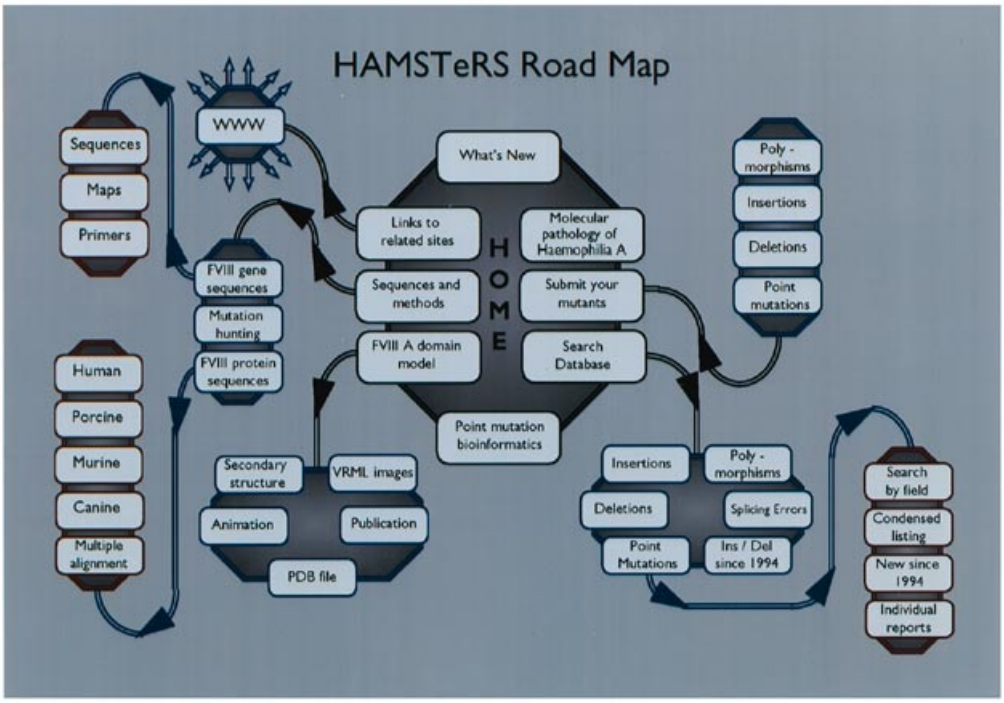

Figure 1. The new clickable Road Map allows rapid access to the individual features of the FVIII Resource Site using a single mouse click, without passing through a hierarchy of pages.

http://www.cern.ch/ExpSupport/) and served to outside users via the Hammersmith Hospital 100 MB fibre optic line.

FEATURES AVAILABLE AT THE HAMSTERS WWW SITE

Navigation

The user may click the new Road Map icon to view a graphical representation of the site layout (Fig. 1). This is a client-side 'clickable' map: clicking the mouse button over any box of interest will take the user directly to the relevant page. Alternatively all of the main functions or sections of the FVIII resource site may be accessed as before from the icons of the home page or by clicking the text links below these icons. With regard to other minor functions, the What's New page summarises recent changes and upcoming features: the Review (14) contains a concise overview of the molecular genetics of haemophilia A, linking to the updated tables including all published and unpublished mutations submitted to HAMSTeRS. The Links page has also been reorganised.

Haemophilia A Mutation Database

The layouts of the Novel Mutation Submission Forms (insertions, deletions and point mutations) have been improved, and the Point Mutation Database Search Page modified to add to the existing powerful Boolean database search (using an extensive range of search parameters such as exon and codon number, nucleotide sequence, amino acid change, clinical severity, inhibitor status and reporting group) the ability to search by range of laboratory values. During 1997 a further 22 novel point mutations, including four nonsense mutations, have been submitted to the database. We are investigating ways in which to display graphically bioinformatic analyses of the point mutations distribution and characteristics (such as bar graphs depicting the distribution of mutant reports over the 26 exons), however, the user may still browse simple tabular listings of individual reports or summaries.

Insertion and Deletion lists have been updated (accessible as HTML-formatted tables) and we have now included lists of predicted Splice Junction variants and Polymorphisms. We have continued listings of all the types of mutations reported subsequent to the last printed version in 1994 (6). Finally, a Java-coded pop-up window allows easy viewing of alphabetically-sorted journal references where required throughout the database.

FVIII structure pages

Users may inspect or download various representations of a homology model of the A domains of FVIII (15) based on the crystal structure of human caeruloplasmin (16), including residue-specific VRML images; the coordinates of the model in Protein Data Bank (PDB) format; secondary structural data, i.e., helix/sheet/turn/coil, amino acid sidechain area and solvent accessibility; and a 60 s Quicktime format animated sound+video file (9 MB) demonstrating the overall features of the model. The model has been useful in recent studies of FVIII variants (17): it may also be used to generate hypothetical macromolecular assemblies consistent with existing biochemical and structural data, such as a complex of the FVIII A domains with FIXa bound to a phospholipid membrane (Fig. 2): the porcine FIXaβ structure (18) has been aligned manually to the FVIII A domain model, constrained by the known interactions between (i) the FVIII A2 loop S558–Q565 and the FIXaβ serine protease domain (19),

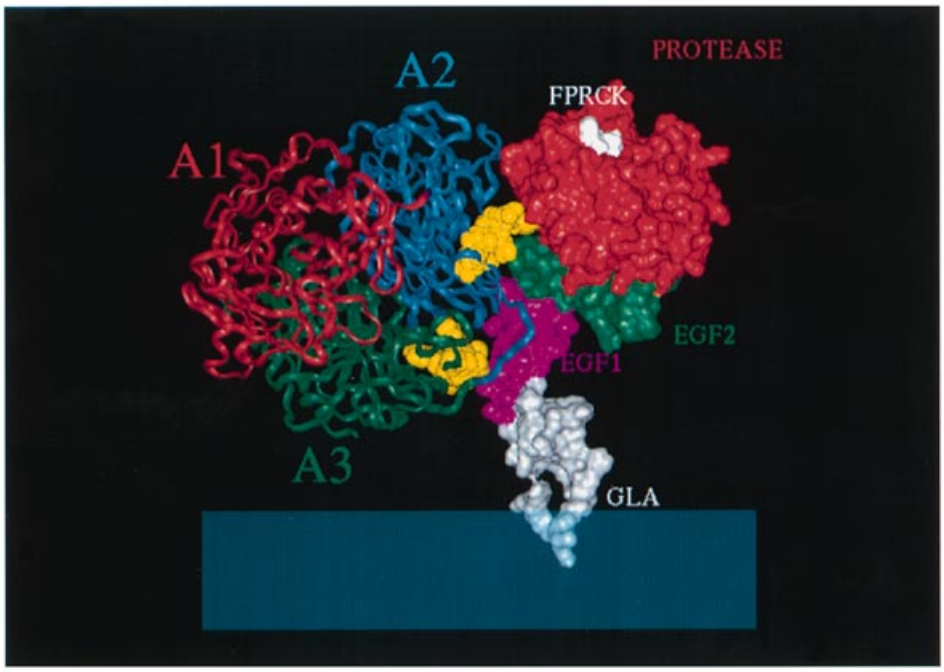

Figure 2. A hypothetical visualisation of the FX-activating complex. The FVIII A domain model (ribbon representation) has been manually aligned to the molecular surface of the porcine FIXaβ crystal structure. FIXaβ is shown anchored in a phospholipid surface (blue) via the γ-carboxyglutamic acid-rich (GLA) domain (light grey), with its serine protease domain (red) adjacent with the FVIII A2 loop (S558–Q565) and its first EGF-like domain (EGF1, magenta) aligned to the FVIII A3 loop (E1811–K1818). The FVIII loops implicated in FIXaβ binding are shown as yellow CPK spheres. The location of the FIXaβ active site is shown by its occupancy by the oligopeptide inhibitor FPRCK (white).

(ii) the FVIII A3 loop E1811–K1818 and the FIXaβ EGF-1 domain (20). The C-terminus of the A3 domain is oriented towards the phospholipid surface consistent with a specific binding interaction via the C2 domain of FVIII.

FVIII sequences and methods

Mature FVIII amino-acid sequences for a total of four mammalian species (human, murine, porcine and canine) are now displayed, together with a multiple sequence alignment displaying the identity and conservation of residues between them. Figure 3 shows part of the alignment (covering ~300 amino acids from A1 residue 200 in human FVIII, across the A1–A2 acidic peptide junction and partway through the A2 domain). Human FVIII cDNA and exonic sequences, together with restriction maps and primer sequences for amplification of exons 1–26 are also provided, together with details of the various approaches to FVIII mutation screening by PCR-based methods (21).

Expansion of the URL to include FVII Mutation Database

Visitors to the HAMSTeRS URL are presented with a selection page allowing access to the FVIII Site or to the new FVII Mutation Database (McVey and Boswell, personal communication).

CONCLUSION

As high-quality structural information on proteins in the area of haemostasis becomes available, it becomes possible to make hypotheses as to the molecular cause of dysfunction in a variant protein found in a patient attending the clinic for a bleeding or thrombotic problem. Although such structural information is not yet available for FVIII or its domains, the A domain homology model available at the site (together with the spectrum of mutations listed) provides a useful starting point for such studies or hypotheses. The further development of this FVIII WWW site is intended to supply both the most up-to-date structure and mutation information together with tools and resources to assist in using them. The authors welcome suggestions which may help to realise this intention.

CITATION

Users of HAMSTeRS are asked to cite this article in their publications, including the URL, http://europium.mrc.rpms.ac.uk

DISTRIBUTION

The Haemophilia A Mutation Search, Test and Resource Site (HAMSTeRS) may be accessed via the World Wide Web at http://europium.mrc.rpms.ac.uk . For those with no Internet access, text versions of the main mutation tables may be obtained from G.K.-C.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Anne Goodeve for compiling the Polymorphisms listing, Dr David Lillicrap and colleagues for permission to include the canine FVIII amino acid sequence ahead of publication, Dr John McVey for his continued critical review of the site, our colleagues who have submitted their new results to the database; and all those who have commented

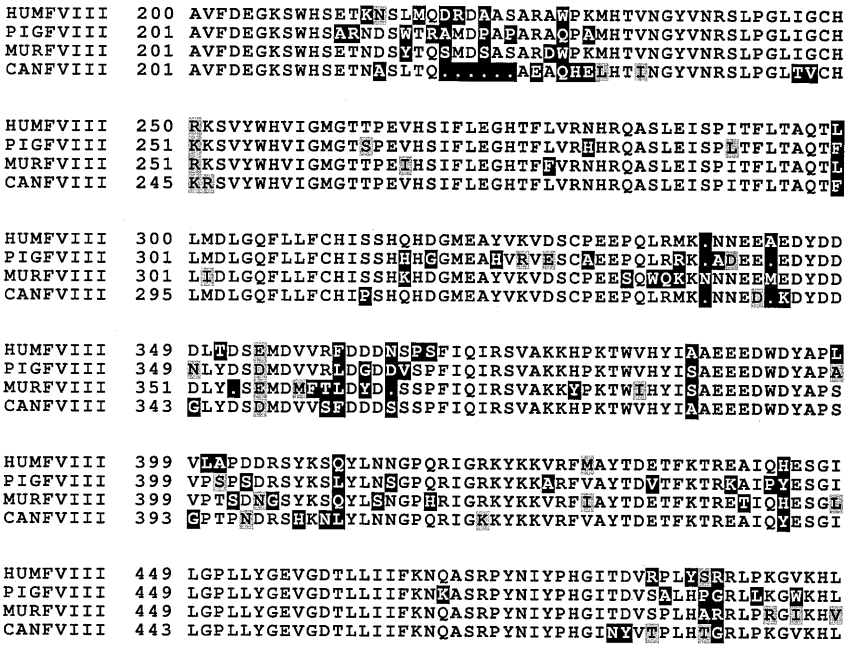

Figure 3. Multiple alignment of the mature amino-acid sequences (residues 200–500 approximately) from four mammalian species. Top to bottom: human, porcine, murine and canine FVIII. Identical residues, black text on white; similar residues, black on grey; dissimilar residues, white on black.

constructively on the use of the site. AIW is supported by the John Ellerman Foundation.

REFERENCES

1 Tuddenham, E.G.D. and Cooper, D.N. (1994) *The Molecular Genetics of Haemostasis and its Inherited Disorders*. Oxford University Press, Oxford, UK.
2 Toole, J.J., Knopf, J.L, Wozney, J.M., Sultzman, L.A., Buecker, J.L., Pittman, D.D., Kaufman, R.J., Brown, E., Shoemaker, C., Orr, E.C. *et al.* (1984) *Nature*, **312**, 342–347.
3 Antonarakis, S.E., Kazazian, H.H. and Tuddenham, E.G.D. (1995) *Hum. Mutat.*, **5**, 1–22.
4 Kazazian, H.H., Tuddenham, E.G.D. and Antonarakis, S.E. (1995) In Scriver, C.R., Beaudet, A.L., Sly, W.S. and Valle, D. (eds), *The Metabolic and Molecular Bases of Inherited Disease*, 7th Edn. McGraw-Hill, New York, pp. 3241–3267.
5 Tuddenham, E.G.D., Cooper, D.N., Gitschier, J., Higuchi, M., Hoyer, L.W., Yoshioka, A., Peake, I.R., Schwaab, R., Olek, K., Kazazian, H.H. *et al.* (1991) *Nucleic Acids Res.*, **19**, 4821–4833.
6 Tuddenham, E.G.D., Schwaab, R., Seehafer, J., Millar, D.S., Gitschier, J., Higuchi, M., Bidichandani, S., Connor, J.M., Hoyer, L.W., Yoshioka, A. *et al.* (1994) *Nucleic Acids Res.*, **22**, 4851–4868.
7 Wacey, A.I., Kemball-Cook, G., Kazazian, H.H., Antonarakis, S.E., Schwaab, R., Lindley, P.F. and Tuddenham, E.G.D. (1996) *Nucleic Acids Res.*, **24**, 100–102.
8 Kemball-Cook, G. and Tuddenham, E.G.D. (1997) *Nucleic Acids Res.*, **25**, 128–132.
9 Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711
10 Boxshade Multiple Sequence Alignment display program: written by Kay Hoffman and Michael D Baron. Freely available from michael.baron@bbsrc.ac.uk.
11 Wood, W.I., Capon, D.J., Simonsen, C.C., Eaton, D.I., Gitschier, J., Keyt, B., Seeburg, P.H., Smith, D.H., Hollingshead, P., Wion, K.I. *et al.* (1984) *Nature*, **312**, 330–337.
12 Elder, B., Lakich, D. and Gitschier, J. (1993) *Genomics*, **16**, 374–379.
13 Healey, J.F., Lubin, I.M. and Lollar, P. (1997) *Blood*, **88**, 4209–4214.
14 Kemball-Cook, G. and Tuddenham, E.G.D. (1997) In Forbes, C.D., Aledort, L.M. and Madhok, R. (eds), *Textbook of Haemophilia*. Chapman and Hall, London, UK.
15 Pemberton, S., Lindley, P.F., Zaitsev, V., Card, G., Tuddenham, E.G.D. and Kemball-Cook, G. (1997) *Blood*, **89**, 2413–2421.
16 Zaitseva, I., Zaitsev, V., Card, B., Moshkov, K., Bax, B., Ralph, A. and Lindley, P.F. (1996) *J. Biol. Inorg. Chem.*, **1**, 15–23.
17 Amano, K., Sarkar, R., Pemberton, S., Kemball-Cook, G., Kazazian, H.H., Jr and Kaufman, R.J. (1998) *Blood*, in press.
18 Brandstetter, H., Bauer, M., Huber, R., Lollar, P. and Bode, W. (1995) *Proc. Nat. Acad. Sci. USA*, **92**, 9796–9800.
19 Fay, P.J., Beattie, T., Huggins, C.F. and Regan, L.M. (1994) *J. Biol. Chem.*, **269**, 20522–21527.
20 Lenting, P.J., Van de Loo, J.W.H.P., Donath, M.J.S.H., Van Mourik, J.A.and Mertens, K. (1996) *J. Biol. Chem.*, **271**, 1935–1940.
21 Michaelides, K., Schwaab, R., Lalloz, M.R.A., Schmidt, W. and Tuddenham, E.G.D. (1995) In Hames, D. and MacPherson, M (eds), *PCR II. A Practical Approach*. IRL Press, Oxford, UK, pp.255–288.