FEBS Letters 430 (1998) 126–129 FEBS 20322

Minireview

Insights into the structure of hepatocyte growth factor/scatter factor (HGF/SF) and implications for receptor activation

Dimitri Y. Chirgadze[1,a], Jonathan Hepple[1,b], R. Andrew Byrd[c], R. Sowdhamini[a], Tom L. Blundell[a,*], Ermanno Gherardi[b]

[a]*Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK*
[b]*Growth Factors Group, Department of Oncology, MRC Center, Hills Road, Cambridge CB2 2QH, UK*
[c]*Macromolecular NMR Section, ABL-Basic Research Program, NCI-Frederick, Cancer Research and Development Center, Frederick, MD 21702-1201, USA*

Received 4 May 1998

Abstract The modular structure of HGF/SF offers a reductionist or 'divide and rule' approach to the analysis of structure and function. Domain deletion experiments have established that the N domain, kringle 1 and kringle 2 are essential for HGF/SF activity and that truncated variants containing the N domain and kringle 1 (NK1) or kringles 1 and 2 (NK2) can exhibit partial agonistic or antagonistic activity depending on target cells. Comparative modelling has been used to predict the 3D structures of the six domains of HGF/SF. More recently, NMR methods have shown that the N domain has a novel fold, the charge distribution of which suggests a heparin binding site. Crystals of NK1 indicate the relationship of this domain to the kringle 1, offering further insights into the mechanism of domain interactions and receptor activation.

Key words: Hepatocyte growth factor/scatter factor; Protein domain; NMR; X-ray analysis; Receptor activation

1. Introduction

HGF/SF [1–5], the first member of the plasminogen related growth factor family, is composed of six domains whose boundaries are clearly defined at the sequence level. The N-terminal domain, the so-called N domain, comprising residues Glu-30-Asn-121 is homologous to the activation peptide of plasminogen. A short linker peptide connects the N domain with four copies of the kringle domain: Cys-128-Cys-206 (kringle 1), Cys-211-Cys-288 (kringle 2), Cys-305-Cys-383 (kringle 3) and Cys-391-Cys-469 (kringle 4). A linker peptide of 24 amino acids links the forth kringle to an inactive serine proteinase domain (Val-495-Ser-728). This peptide contains a half-cystine residue (Cys-487) predicted to form a disulphide bond with Cys-604 and a trypsin-like cleavage site (Gln-492-Leu-493-Arg-494 ↓ Val-495-Val-496-Asn-497), which is cleaved upon conversion of single-chain pro-HGF/SF into a two-chain molecule.

An understanding of the relationship between structure and function in HGF/SF depends on a knowledge not only of the conformations of the individual domains but also on their arrangement with respect to each other, both in solution and in complex with the HGF/SF (MET) receptor. The availability of high-resolution structures of kringle and serine proteinase domains allowed us several years ago to construct three-dimensional models of the corresponding HGF/SF domains [6]. The models of kringles 1 to 4 of HGF/SF confirmed the presence of the typical features of kringle domains, i.e. a hydrophobic core and a globular structure stabilized by the three conserved disulphide bonds. The model of the serine proteinase confirmed that the usual tertiary structure could be assumed, although critical substitutions affected two catalytic residues (His534Gln and Ser673Tyr). Substitutions were also apparent in the region corresponding to the S1 specificity pocket of active enzymes, notably in residues responsible for arginine specificity. The unavailability of structures of homologues prevented comparative modelling of the N domain. Instead, a 3D model of the 27-residue hairpin loop, which plays an important role in receptor binding and activity, was constructed. The modelling was performed assuming constraints on the conformation that occur if the four half-cystine residues (Cys-70, Cys-74, Cys-84 and Cys-96) exhibit the connectivity of the corresponding residues in the homologous plasminogen domain. The resulting structure suggested a helix-extended strand-helix motif, with a cluster of positively charged residues (Arg-73, Arg-76 and Arg-93) on the helical side of the loop and a cluster of hydrophobic residues (Leu-80, Phe-82, Ala-86, Phe-87, Val-88 and Phe-89) on the extended strand [6].

In this review we briefly examine the experimental data on domain deletion and truncated variants. We describe the novel fold of the N domain defined by NMR methods and compare this to the predicted structure. We also report the preparation and analysis of crystals of NK1, which define the relationship of this domain to kringle 1, offering further insights into the mechanism of domain interactions and receptor activation.

2. Domain deletion analysis and truncated variants of HGF/SF

Deletion of either the hairpin loop of the N domain, kringles 1 or 2, or the serine proteinase domain (ΔH, ΔK1, ΔK2 and ΔSP) abolished biological activity, as measured by stimulation of DNA synthesis or scattering of MDCK cells. In contrast, mutants lacking kringle 3 or 4 (ΔK3 or ΔK4) showed reduced but measurable activity [7]. These observations were largely confirmed by subsequent studies in which both recep-

*Corresponding author. Fax: +44 (1223) 766 082.
E-mail: tom@cryst.bioc.cam.ac.uk

[1]These authors contributed in an equivalent way to this work.

PII: S0014-5793(98)00558-4

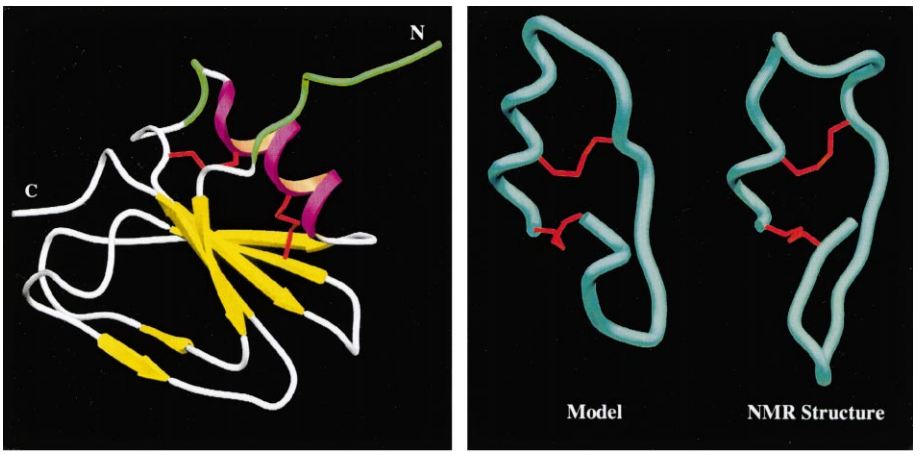


Fig. 1. a: The structure of the N domain of HGF/SF defined by NMR [18]. Helices are shown in magenta, β-strands in yellow, disulfide bridges in red and the residues that were not well defined in green. b: The hairpin loop, modelled [6] and determined by NMR [18]. The figure was produced using *SETOR* [10].

tor binding and phosphorylation as well as biological activity were investigated [8,9].

Confirmation of the critical role of the N domain, kringle 1 and kringle 2 in MET binding and activation has come from studies on two truncated forms of HGF/SF containing the N domain+kringle 1 (NK1) or the N domain+kringles 1 and 2 (NK2). NK2 was discovered as the product of an alternatively spliced form of the primary HGF/SF transcript [11], and a subsequent study indicated that NK2 had considerable antagonistic activity [12]. NK1 was first expressed as an engineered fragment and also appeared to behave as a potent HGF/SF antagonist [13]. Only recently has it become apparent that NK1 also exists in vivo as a further splice variant of the primary HGF/SF transcript [15].

Although early work suggested that both NK2 [12] and NK1 [13] behaved as receptor antagonists, other studies indicated that both proteins could act as partial agonists [14,15]. Recent experiments suggest that cell surface heparan sulphate proteoglycans (HSPGs) promote dimerization of NK1 and NK2 and induce an agonistic response in target cells [16]. Furthermore Merlino and colleagues [17] have demonstrated that mice overexpressing NK1 develop all the features of transgenic mice overexpressing full length HGF/SF, albeit in a less severe form.

3. The structure of the N domain

Three-dimensional structural information of HGF/SF N domain was recently obtained by NMR techniques [18]. The overall structure of the N domain is very well defined, although Asn-77-Gly-79 and the first seven N-terminal residues could not be observed. The central part of the structure is represented by a five-stranded antiparallel β-sheet (Fig. 1a). This sheet is flanked on one side by a two-turn alpha helix and on the other by extended loops, which transform into the β-strands, forming a small sheet. The helix and two of the strands of the central β-sheet make a hairpin-loop structure, which is stabilized by the disulphide bridges. The topology of the hairpin-loop region generally resembles the fold previously predicted by the use of disulphide restraints, although the last strand was predicted as a loose helix (Fig. 1b).

Studies of heparin binding by proteins suggest that interactions take place between basic amino acids and the sulphate and the carboxylate groups of heparin [19]. The N-terminal domain contains a large number of lysine and arginine residues, some of which are likely to be involved in heparin binding. The NMR structure suggests that a surface area, including Lys-60, Lys-62, Arg-73, Arg-76 and Lys-78, may serve as a heparin binding site. This includes Arg-73 and Arg-76, which were predicted on the basis of the model [6]. There is now biochemical evidence that these two residues are involved in the binding of HGF/SF to heparin [20].

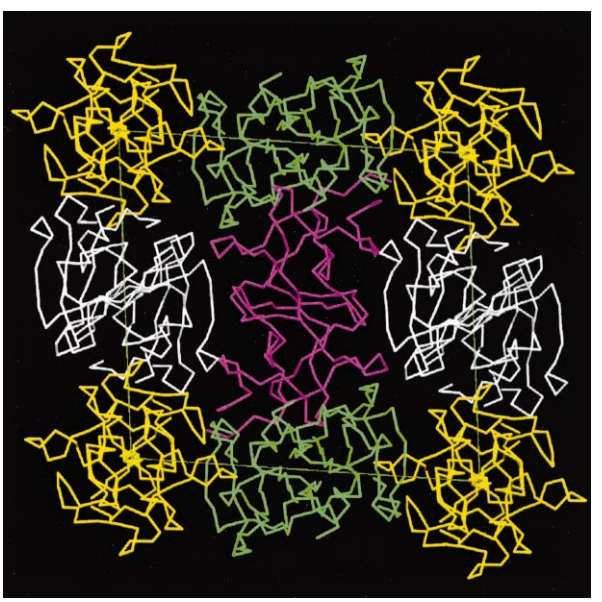


Fig. 2. The packing of NK1 in the crystal cell as defined by molecular replacement. N domains are shown in white and green, kringle domains in yellow and magenta (Chirgadze, unpublished). The figure was produced using *SETOR* [10].

4. Crystallization of NK1

We have successfully expressed NK1 and NK2 fragments of HGF/SF in the yeast *Pichia pastoris* and purified both proteins by a combination of heparin affinity and cation exchange chromatography (J. Hepple et al., to be published). Gel filtration analysis showed that yeast-derived NK1 is a monomer while NK2 is a mixture of monomeric and dimeric species. NK1 has been crystallized. The crystals belong to the $P2_1$ space group and diffract X-rays to a resolution of 2.5 Å.

In order to place the NK1 molecule in the crystal cell we have employed the molecular replacement method using the crystal structure of kringle 1 of plasminogen [21] and the solution structure of the N domain [18] as search probes. The existence of two molecules in the asymmetric unit complicated the identification of the correct molecular replacement solutions. The crystal packing of the domains is shown in Fig. 2. The structure is currently undergoing refinement and the full description of the structure determination will be published elsewhere.

5. Site-specific mutants of HGF/SF

The sequence features of HGF/SF together with the results of domain deletion experiments and the availability of 3D models of individual domains have led to several studies in which specific residues or clusters of residues have been substituted in order to clarify their role in HGF/SF activity. The main results of these studies can be summarized as follows:

(1) Although several studies have pointed to a critical role of the N domain (or its hairpin loop) in receptor binding and biological activity, a number of individual and group mutations have failed so far to identify residues critical for receptor binding within the loop itself [22]. The same study, however, established that a cluster of residues located at the C-terminus of the N domain (His-114, Glu-115 and Asp-117) is involved in both receptor binding and biological activity.

(2) Substitution of at least seven amino acids in kringle 1 (Arg-197, Glu-189, Tyr-198, Glu-195, Asp-171, Gln-173 and Ser-161) has a clear effect on receptor binding and biological activity of HGF/SF. The greatest effect is seen with substitution of Arg-197 or Glu-159 which reduce receptor binding and biological activity by >50-fold [22].

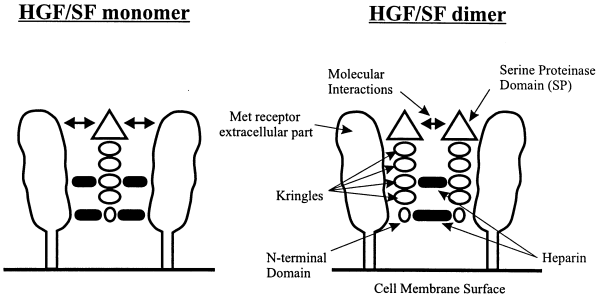
(3) Several point mutations in the serine proteinase domain of HGF/SF in which residues Gln-534 and Tyr-673 were reverted to the His and Ser residues of active serine proteinase did not affect receptor binding but markedly decreased biological activity. A similar result was obtained with the Val692Ser mutation [9].

(4) Mutation of critical residues at the trypsin-like cleavage site, located between the C-terminal kringle (kringle 4) and the serine proteinase domain, abolished conversion of the single-chain precursor form of HGF/SF (pro-HGF/SF). The resulting single-chain species lacked the biological activity of HGF/SF on target cells, although it retained receptor binding [9,23,24]. Thus, the precursor form of HGF/SF behaves like a receptor antagonist.

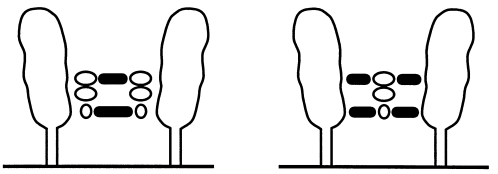
6. Models for the activation of the receptor Met protein by HGF/SF

With the exception of the insulin receptor family, all known receptors with tyrosine kinase activity undergo a transition from a monomeric to dimeric state upon binding of their cognate ligand. Therefore the ligand must either possess two binding sites or be a dimer itself. In the case of HGF/SF, the stoichiometry of ligand-receptor interaction is still unknown. A model for the formation of homodimers has been suggested by Donate and colleagues [6]. This model proposed non-covalent interactions between kringles 2, 3 and the serine proteinase domain. A lysine residue of kringle 3 in molecule 2 would interact with the lysine binding pocket of kringle 2 in molecule 1. This model has been disproved by mutational studies, which show that the lysine binding pocket of kringle 2 is not required for the biological activity of HGF/SF [21].

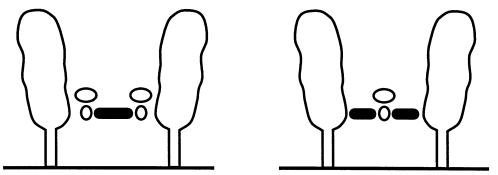
Ponting and colleagues [25] have proposed that HGF/SF may adopt similar conformations to plasminogen. In its inactive form, plasminogen adopts a closed, spiral structure, maintained by interactions between the N- and C-terminal domains while in its active form, plasminogen has an open 'horseshoe' conformation. Mutagenesis experiments have shown that cleavage of the single-chain HGF/SF into a two-chain form is required for biological activity [9]. While the single-chain form is capable of binding Met, the cleavage to a two-chain form probably results in a conformational change and a rearrangement of the relative positions of the domains. This relaxed, open conformation may allow HGF/SF to interact with and directly activate Met as a monomer or alternatively, enable HGF/SF to dimerize before receptor activation (Fig. 3). In the case of HGF/SF, the receptor-ligand



HGF/SF: AGONIST. Cleavage at Arg494-Val495 and SP residues Gln534, Tyr673 and Val692 required for activity. Heparin adds additional stability but not essential [9, 20].



NK2: PARTIAL AGONIST. Heparin required for mitogenic activity [26].



NK1: PARTIAL AGONIST. Heparin required for mitogenic activity [26].

Fig. 3. Models for Met receptor dimerization and the role of heparin.

complex is probably stabilized by interactions between the serine proteinase domain(s), as residues have been identified in this domain which are necessary for biological activity but not receptor binding [9]. Heparin and heparan sulfates appear to be essential for the partial agonist behavior of NK1 and NK2 and may do so by crosslinking the N-terminal and kringle 2 domains [26].

7. Conclusions and perspectives

Protein engineering experiments have provided valuable data on the role of HGF/SF domains in receptor binding and activation. In parallel, 3D models of the kringle and serine proteinase domains and of the hairpin loop of the N domain [6,9,22] have given a basis for the design of site-specific mutants which enable mapping of the MET and heparin binding sites of HGF/SF. With NMR and crystal structures now becoming available these studies of structure and function can be placed on a firmer basis.

There is now little doubt that the N domain and kringle 1 are directly involved in receptor binding and activation as demonstrated by domain deletion experiments, several site-specific mutants and recent experiments with the NK1 fragment in vitro [14] and in vivo [17].

The role of other domains, especially the serine proteinase domain, is more difficult to assess. Whereas it is clear that deletion of this domain leads to a marked decrease in biological activity, the mechanism by which this is achieved is less clear. Available data do not rule out that the serine proteinase domain may contain a secondary binding site for the MET receptor or alternatively, that the serine proteinase domain may act as a dimerization domain and allow an HGF/SF dimer to engage two receptor molecules through their NK1 domains. A definite answer to these questions, however, may have to wait for the structure of HGF/SF-MET complexes.

Acknowledgements: Research work was supported by the Imperial Cancer Research Fund and the Wellcome Trust.

References

[1] Nakamura, T., Nawa, K., Ichihara, A., Kaise, N. and Nishino, T. (1987) FEBS Lett. 224, 311–316.
[2] Stoker, M., Gherardi, E., Perryman, M. and Gray, J. (1987) Nature 327, 239–242.
[3] Gherardi, E., Gray, J., Stoker, M., Perryman, M. and Furlong, R. (1989) Proc. Natl. Acad. Sci. USA 86, 5844–5848.
[4] Miyazawa, K., Tsubouchi, H., Naka, D., Takahashi, K., Okigaki, M., Arakaki, N., Nakayama, H., Hirono, S., Sakiyama, O., Takahashi, K., Godha, E., Daikuhara, Y. and Kitamura, N. (1989) Biochem. biophys. Res. Commun. 163, 967–973.
[5] Nakamura, T., Nishizawa, T., Hagiya, M., Seki, T., Shimonishi, M., Sugimura, A., Tashiro, K. and Shimizu, S. (1989) Nature 342, 440–443.
[6] Donate, L., Gherardi, E., Srinivasan, N., Sowdhamini, R., Aparicio, S. and Blundell, T. (1994) Protein Sci. 3, 2378–2394.
[7] Matsumoto, K., Takehara, T., Inoue, H., Hagiya, M., Shimizu, S. and Nakamura, T. (1991) Biochem. Biophys. Res. Commun. 181, 691–699.
[8] Okigaki, M., Komada, M., Uehara, Y., Miyazawa, K. and Kitamura, N. (1992) Biochemistry 31, 9555–9561.
[9] Lokker, N.A., Mark, M.R., Luis, E.A., Bennett, G.L., Robbins, K.A., Baker, J.B. and Godowski, P.J. (1992) EMBO J. 11, 2503–2510.
[10] Evans, S.V. (1993) J. Mol. Graph. 11, 134–149.
[11] Miyazawa, K., Kitamura, A., Naka, D. and Kitamura, N. (1991) Eur. J. Biochem. 197, 15–22.
[12] Chan, A.M-L., Rubin, J.S., Bottaro, D.P., Hirschfield, D.W., Chedid, M. and Aaronson, S.A. (1991) Science 254, 1382–1385.
[13] Lokker, N.A. and Godowski, P.J. (1993) J. Biol. Chem. 268, 17145–17150.
[14] Hartmann, G., Naldini, L., Weidner, K.M., Sachs, M., Vigna, E., Comoglio, P.M. and Birchmeier, W. (1992) Proc. Natl. Acad. Sci. USA 89, 11574–11578.
[15] Cioce, V., Csaky, K.G., Chan, A.M-L., Bottaro, D.P., Taylor, W.G., Jensen, R., Aaaronson, S.A. and Rubin, J. (1996) J. Biol. Chem. 271, 13110–13115.
[16] Schwall, R.H., Chang, L.Y., Godowski, P.J., Kahn, D.W., Hillan, K.J., Bauer, K.D. and Zioncheck, T.F. (1996) J. Cell Biol. 133, 709–718.
[17] Jakubczak, J.L., LaRochelle, W.J. and Merlino, G. (1998) Mol. Cell. Biol. 18, 1275–1283.
[18] Zhou, H., Mazulla, M.J., Kaufman, J.D., Stahl, S.J., Wingfield, P.T., Rubin, J.S., Bottaro, D.P. and Byrd, R.A. (1998) Structure 6, 109–116.
[19] Jackson, R.L., Busch, S.J. and Cardin, A.D. (1991) Physiol. Rev. 71, 481–539.
[20] Hartmann, G., Prospero, T., Brinkmann, V., Ozcelik, O., Winter, G., Hepple, J., Batley, S., Bladt, F., Sachs, M., Birchmeier, C., Birchmeier, W. and Gherardi, E. (1998) Curr. Biol. 8, 125–133.
[21] Wu, T-K., Padmanabhan, K.P. and Tulinsky, A. (1994) Blood Coagul. Fibrinolysis 5, 157–166.
[22] Lokker, N.A., Presta, L.G. and Godowski, P.J. (1994) Protein Eng. 7, 895–903.
[23] Naka, D., Ishii, T., Yoshiyama, Y., Miyazawa, K., Hara, H., Hishida, T. and Kitamura, N. (1992) J. Biol. Chem. 267, 20114–20119.
[24] Gak, E., Taylor, W.G., Chan, A. and Rubin, J.S. (1992) FEBS Lett. 311, 17–21.
[25] Ponting, C.P., Marshall, J.M. and Cederholm-Williams, S.A. (1992) Blood Coagul. Fibrinolysis 3, 605–614.
[26] Sakata, S., Stahl, S.J., Taylor, W.G., Rosenberg, J.M., Sakaguchi, K., Wingfield, P.T. and Rubin, J.S. (1997) J. Biol. Chem. 272, 9457–9463.