Direct observation of hierarchical protein dynamics

Józef R. Lewandowski[1]†*, Meghan E. Halse[1]‡, Martin Blackledge[2]*, Lyndon Emsley[1,3]*

[1]Université de Lyon, Institut de Sciences Analytiques (CNRS / ENS-Lyon / UCB-Lyon 1), Centre de RMN à Très Hauts Champs, 69100 Villeurbanne, France.

[2]Univ. Grenoble Alpes, Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble 38027, France. CNRS, Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble 38027, France. CEA, Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble 38027, France.

[3]Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

*Correspondence to: Józef R. Lewandowski j.r.lewandowski@warwick.ac.uk, Martin Blackledge martin.blackledge@ibs.fr, Lyndon Emsley lyndon.emsley@epfl.ch.

†Current address: Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom.

‡Current address: Department of Chemistry, University of York, York YO10 5DD, United Kingdom

Abstract:

One of the fundamental challenges of physical biology is to understand the relationship between protein dynamics and function. At physiological temperatures functional motions arise from the complex interplay of thermal motions of proteins and their environment. Here we determine the hierarchy in the protein conformational energy landscape that underlies these motions based on a series of temperature-dependent magic-angle spinning multinuclear NMR relaxation measurements in a hydrated nanocrystalline protein. The results support strong coupling between protein and solvent dynamics above 160 K, with fast solvent motions, slow protein side-chain motions, and fast protein backbone motions being activated consecutively. Low activation energy, small-amplitude local motions dominate at low temperatures, with larger-amplitude, anisotropic, and functionally relevant motions involving entire peptide units becoming dominant at temperatures above 220 K.

One Sentence Summary: The dynamic modes that relate solvent, sidechain and backbone motions in a protein are determined from NMR relaxation.

Proteins must traverse complex conformational energy landscapes in order to perform their physiological function. This is achieved through thermally activated fluctuations (*1*), and when specific molecular motions cease at low temperatures, function also ceases or becomes significantly reduced (*2, 3*). Understanding the hierarchy of these motions thus holds the key to understanding how proteins function on a molecular level at physiological temperatures. Contrary to expectations, and despite large differences in structure and function between proteins

of different families, dynamic properties appear to exhibit common general features, in particular apparent transitions between different dynamic regimes as a function of temperature. These transitions are thought to occur as a result of coupling between proteins and the surrounding solvent (*4*).

Observing and understanding protein dynamic transitions has been a focus of many fields of research over the last 40 years, including neutron scattering (*5*), Mössbauer spectroscopy (*4, 6*), Terahertz spectroscopy (*7*), dielectric spectroscopy (*4*), DSC (*8*), X-ray crystallography (*2, 9*) and molecular dynamics simulation (*10, 11*). Remarkably, this relative wealth of information has not led to a consensus picture of dynamical transitions and their origins, with different techniques detecting distinct processes, leading to apparently contradictory descriptions (*12, 13*). This may be due to the widely varying conditions required for the diverse techniques, or to the sensitivity of the different physical measurements to dynamics occurring on different timescales.

Here we measure, in a single sample, a set of thirteen different NMR observables that are sensitive to dynamics occurring on different timescales and in different parts of the system over temperatures from 105 to 280 K in the fully hydrated crystalline protein GB1, a small globular protein specifically binding to antibodies. The analysis of multiple probes that report on the different structural components of this complex system allows us to develop a complete and coherent picture of the dynamic processes across the whole temperature range, as well as to rationalize prior observations from other techniques. Our findings support strong coupling between protein and solvent dynamics above 160 K, with motions in the solvent, with protein backbone and side-chain motions being activated as temperature increases. The generally observed dynamic transitions are found to report on thermal activation of processes concerning different components of the system, with higher activation energies enabling anisotropic, functionally important backbone motions, coupled to side-chain and solvent motions, becoming dominant above 220 K.

As illustrated in Figure 1, the key to generating this complete picture is the measurement of different ^{1}H, ^{13}C and ^{15}N relaxation rates (*14*) from a single sample. This allows us to determine different dominant rotational motional modes that affect bulk solvent, hydration water, protein side chains and protein backbone in a site-specific manner. Analysis of longitudinal (R_1) and transverse relaxation ($R_{1\rho}$, R_2') rates identifies changes in picosecond-nanosecond and nanosecond-millisecond motions (see Fig. 1).

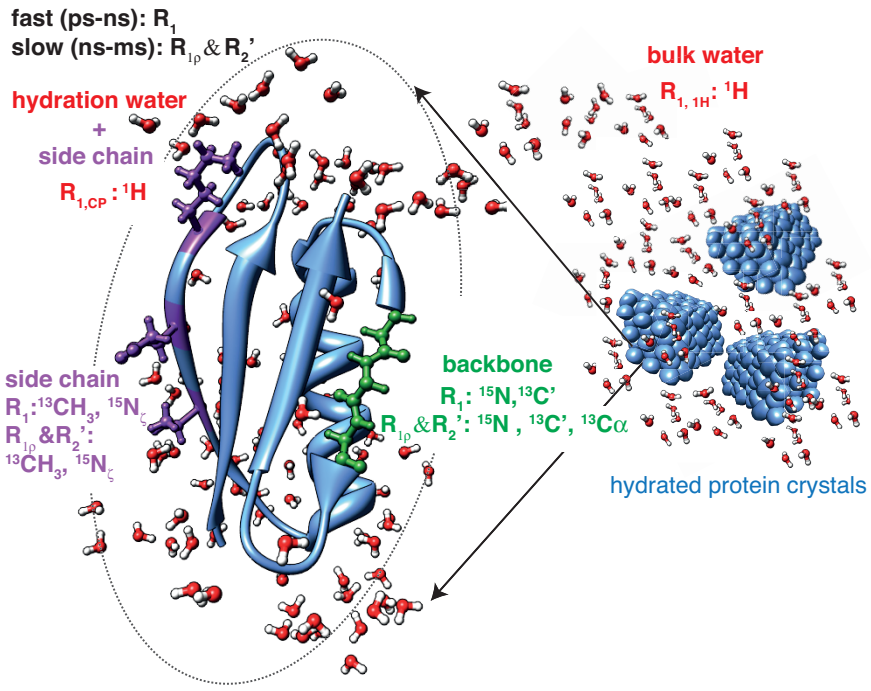

Fig. 1. Cartoon representation of the location of motions and the relaxation rates that are most sensitive to those motions. The rates in green, purple and red report on backbone, side-chain and solvent motions, respectively.

Figure 2 shows experimentally determined longitudinal relaxation rates from backbone, side chain and solvent as a function of temperature. Transverse relaxation rates are presented in Figure 3. A similar pattern is seen in all cases. As temperature increases, abrupt changes in the experimental relaxation rates are observed, suggesting that different motions dominate the measured relaxation within distinct temperature ranges. The origin of this behavior was analyzed by fitting the measured ^{1}H, ^{15}N and ^{13}C relaxation rates to a sum of Arrhenius terms (see Materials and Methods), describing the superposition of distinct motional processes with different activation energies (Figs. 2 & 3; Tables S3-4). Each individual process is fitted with three parameters: the activation energy E_a (see Fig. 2F and 3F; Table S2), the characteristic timescale, and a constant that accounts for the amplitude of the relaxation active interaction associated with this process (and a constant term accounting for coherent contributions to transverse relaxation rates). The expressions for relaxation rates are based on Redfield theory, which validity in the concerned regime has been discussed previously (*15*). Note that we assume a constant amplitude motion throughout each range, which although physically unrealistic, provides an effective amplitude reporting on a broad mean.

Remarkably, the measured rates can be accurately reproduced over the whole temperature range by invoking one to three (typically two) distinct motional processes. The apparent transitions with respect to temperature report on thermal activation of new motional modes, which rapidly become dominant for relaxation rates measured above the transition temperature.

Based on the changes in the observed trends for longitudinal and transverse relaxation rates in Fig. 2 and 3 we are able to identify three key transitions: TI at ~195 K, above which transverse relaxation reports on slow motions dominated by a mode with E_a > 20 kJ/mol, TII at ~220 K, above which longitudinal relaxation reports on fast motions dominated by a mode with E_a ~30 kJ/mol, and TIII at ~250 K, above which faster side-chain motions with E_a ~30 kJ/mol dominate longitudinal methyl relaxation.

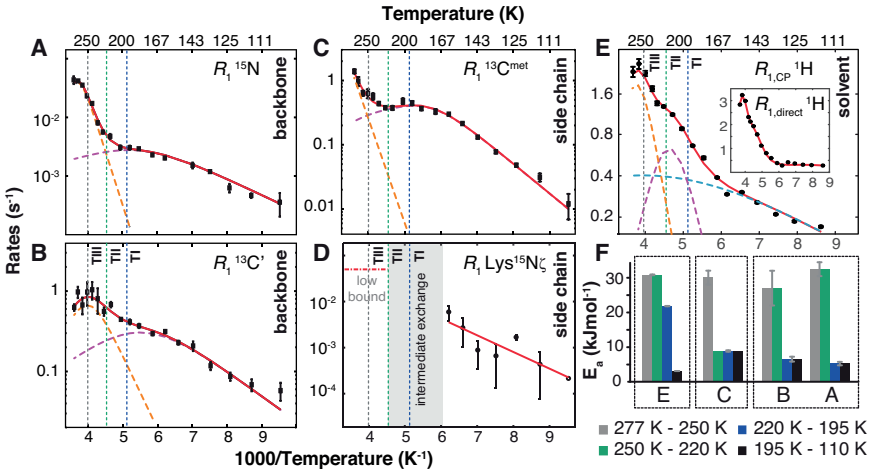

Fig. 2. Measured bulk longitudinal relaxation rates in hydrated nanocrystalline [U-^{13}C,^{15}N]GB1 as a function of temperature. Rates in (A-E) (points with error bars reflecting fitting uncertainties) are sensitive to picosecond-nanosecond motions of protein backbone (A-B), side chain (C-E) and solvent (E). Lines of best-fit over the entire temperature range are shown in red. Two components for protein motions and three components for solvent motions were justified statistically ($p < 0.05$ compared to a single (and one-two) component model). The individual components with distinct activation energies (E_a) obtained from a global fit over each type of nucleus are indicated with dashed lines. The E_a for the components dominating the relaxation rates in different temperature ranges in panels (A-E) are shown in (F). The dynamic transitions: TI, TII and TIII, identified as the temperature where a higher activation energy mode becomes dominant, are shown by vertical dotted lines.

The observation of TI-TIII dynamic boundaries is dependent on the location of the "spy" nuclei, i.e. whether specific relaxation rates probe solvent, protein side chain or protein backbone motions. ^{1}H relaxation rates report on both bulk solvent and hydration water and detect dynamical transitions for fast motions at two out of three transition temperatures (TI & TII). Hydrophobic protein side chains, observed through methyl groups, show a clear fast-motional transition at TIII. For hydrophilic protein side chains - represented by lysine side chains - 170 K and TII delineate an intermediate exchange regime, where the resonances are not observed in the spectra. ^{15}N and ^{13}C' R_1 rates detect a major transition in fast backbone dynamics at TIII. Finally, TI seems to involve slow motions of both protein side chain and backbone, but is more pronounced for the side chains.

^{15}N relaxation rates can be used to quantitatively determine motional amplitudes (see supplementary text S4), thereby providing a physical link between the activation energies and amplitudes of motions. Estimation of the median amplitude ($1\text{-}S^2$) associated with the two backbone dynamic regimes probed by ^{15}N longitudinal relaxation (Fig 2A) reveals a significantly higher amplitude (0.012) for the high-E_a component, than for the low-E_a component (0.0008). Remarkably, the amplitude of the high-E_a component is consistent with S^2 values extrapolated from temperature-dependent solution state studies of GB1 (*16, 17*). This mode can therefore be assigned to the anisotropic backbone fluctuations commonly measured at room temperature in solution and solid-state proteins.

The comparison of the observed dynamical transitions at different locations of the protein-solvent system, and the associated changes in the average activation energies for the dominant motional modes, provides (together with findings from other studies in the literature) the coherent picture for the dynamic transitions in protein behavior summarized in Figure 4.

In the region below 160 K translations and rotations of solvent molecules are arrested (*8*), protein motions are decoupled from the dynamics of the solvent, as evidenced by large differences in the activation energies for the dominant motional modes of solvent and protein (Figure 2F). In this regime, fast motions involving fluctuations of atoms in side chains in local energy wells (e.g. 3-site hops for methyl groups) are dominant (*18*), with the available thermal energy being insufficient to sample different side chain rotameric states. Backbone conformations are frozen in a quasi-continuum of non-interchanging sub-states, unambiguously identifying the origin of the broadening of protein NMR signals at low temperatures (supplementary text S11). The amplitude of motions within these states is very small, but increases gradually, concomitant with dynamics getting faster, as temperature increases from 110-160 K.

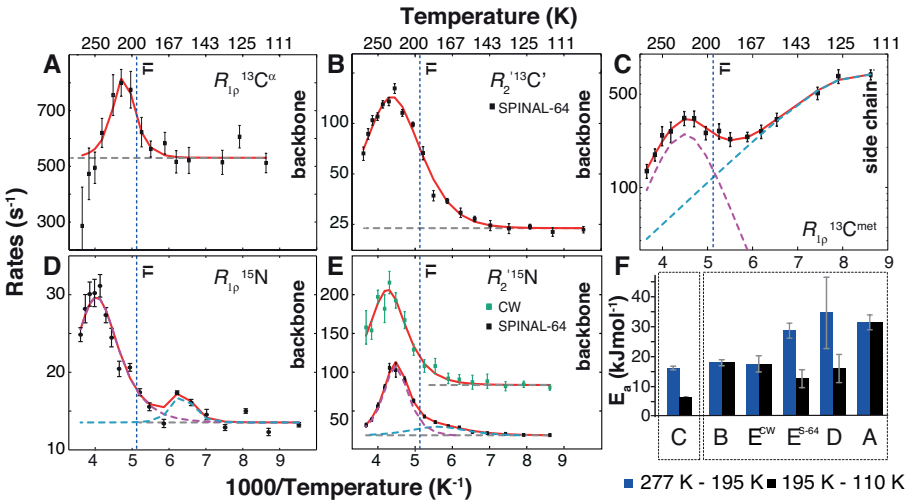

Fig. 3. Bulk transverse relaxation measurements in hydrated nanocrystalline [U-^{13}C,^{15}N]GB1 as a function of temperature. (A-E) Transverse relaxation rates (points with error bars reflecting fitting uncertainties) report on nanosecond-millisecond motions of protein backbone (A-B, D-E) and side chain (F). Lines of best-fit over the entire temperature range are shown in red. Two components were justified statistically ($p < 0.05$ compared to a single component model) for backbone ^{15}N R_2' and methyl side chain. Components with distinct average activation energies obtained from a global fit over all temperatures are indicated with dashed lines. The offset accounting for coherent contributions is shown as a horizontal gray dashed line. E_a for components dominating the relaxation rates in different temperature ranges in panels (A-E) are shown in (F). The dynamic transition TI (195 K), identified as the temperature where a higher activation energy mode becomes dominant, is shown by a vertical dotted line.

Solvent rotations, which become active at ~160 K (*8*) and dominate solvent relaxation rates above TI (~195 K), allow side chains to access different rotameric states (*19*). This has an especially dramatic effect on hydrophilic side chains (probed by lysines), which are affected by the forming and breaking of hydrogen bonds to fluctuating solvent molecules (resulting in a chemical exchange process leading to the disappearance of the lysine resonances from the spectra). The side-chain motions that become dominant above TI, and which are characterized by higher E_a values compared to the local rotations, are slow in nature and thus detected by transverse relaxation only (Fig. 3F). At TI, backbone motions with higher activation energies are also detected by transverse relaxation but the transition from low E_a to high E_a motions is less pronounced than for side chains.

At TII (~220 K), larger-amplitude, nanosecond timescale anisotropic backbone peptide motions (*20*), with significantly increased activation energies, become the dominant source of longitudinal relaxation. These motions are accompanied by the unfreezing of hydration water at

TII (~220 K), corresponding to the onset of translational diffusion of the solvent (*21*). The observed motions represent the dominant modes that are pervasive at physiological temperatures, providing access to rare conformational jumps sampling different regions of Ramachandran space and functionally important conformational rearrangements. The higher E_a associated with motions that become dominant above TII (~220 K) very likely reports on more collective modes, involving at least entire peptide planes, or possibly multiple amino acids, as was suggested in the case of myoglobin (*9*, *10*). We note that correlated motions traversing the entire β-sheet of protein GB1, culminating in larger amplitude excursions of the hydrogen bonding moieties that form the interaction interface with the physiological partner Fab, were observed with correlation times in the nanosecond to microsecond range at 298 K (*20, 22, 23*). At room temperature nanosecond motions dominate both longitudinal and transverse relaxation processes in crystalline GB1 (*22*). The very similar range of activation energies for dominant backbone and solvent motions in this range suggests strong coupling of these modes above TII (~220 K). The observed sequence of motional mode activations strongly suggests correlation between low-E_a and high-E_a motions, supporting the idea that at room temperature fast motions facilitate slow motions.

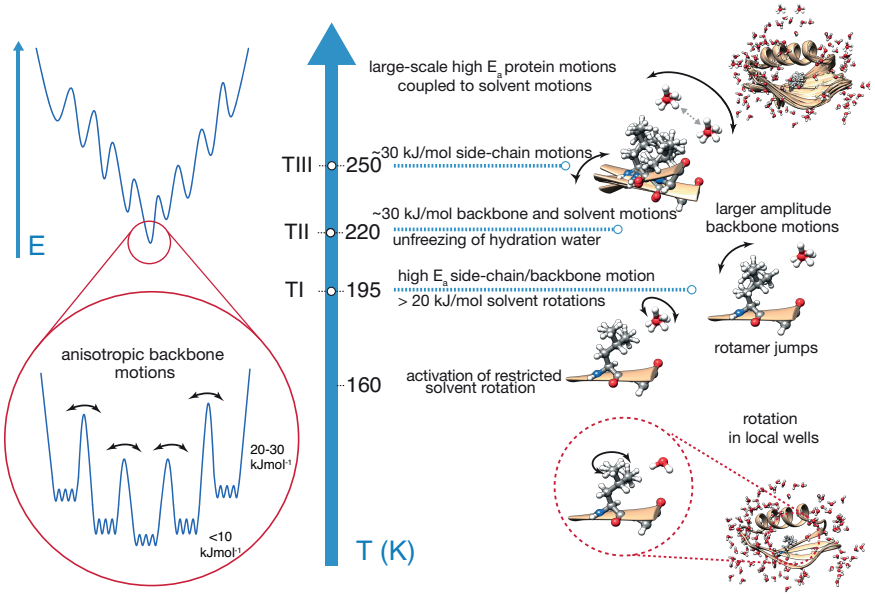

Fig. 4. Summary of hierarchical dynamic behavior of the protein-solvent system as observed by solid-state NMR in a microcrystalline globular protein GB1. The approximate temperature for the transitions between dominant dynamic modes is indicated on the blue axis. The image in the top right corner represents an ensemble extracted from a 200 ns MD simulation in a crystalline environment (*24*). The left hand panel presents a simplified representation of the link between small and larger amplitude backbone motional modes. At low temperatures the protein backbone is constrained to small amplitude modes separated by low energy barriers, within substates separated by high barriers. As temperature increases these modes are dynamically sampled, enabling larger amplitude anisotropic modes.

Finally, larger amplitude fast side-chain motions start visibly contributing to relaxation around TII and become dominant at TIII (~250 K). Notably, these motions are characterized by an E_a very similar to the E_a for the fast solvent and backbone motions in this regime (E_a ~ 30

kJ/mol) providing further evidence for strong coupling between protein and solvent motions at physiological temperatures, as well as the large scale nature of these motions extending over several moieties.

Our observations reconcile divergent interpretations from techniques that are individually sensitive to dynamic phenomena occurring on different time scales and at different locations in protein-solvent systems (supplementary text S6-10). The average activation energies for the dominant motional modes determined here compare very well to those determined from a wide variety of experimental techniques and provide insight into their physical origin. The findings from other approaches, which support our conclusions, are presented in the supporting materials.

In summary, simultaneous measurement of the temperature dependence of thirteen different spin relaxation probes reporting on dynamic modes distributed throughout the solvated microcrystalline protein allows for a unified description of the essential conformational energy surface, relating the amplitude and activation of solvent, side-chain and backbone motions in a hierarchical distribution as well as unambiguous identification of NMR line broadening at cryogenic temperatures.

References and Notes:

1. H. Frauenfelder, S. G. Sligar, P. G. Wolynes, *Science* **254**, 1598-1603 (1991).
2. B. F. Rasmussen, A. M. Stock, D. Ringe, G. A. Petsko, *Nature* **357**, 423-424 (1992).
3. R. M. Daniel *et al.*, *Biophys. J.* **75**, 2504-2507 (1998).
4. H. Frauenfelder *et al.*, *P. Natl. Acad. Sci. U.S.A.* **106**, 5129-5134 (2009).
5. W. Doster, S. Cusack, W. Petry, *Nature* **337**, 754 (1989).
6. F. Parak, H. Formanek, *Acta Crystallogr. A* **27**, 573-578 (1971).
7. J. R. Knab, J. Y. Chen, Y. F. He, A. G. Markelz, *P. Ieee* **95**, 1605-1610 (2007).
8. H. Jansson, R. Bergman, J. Swenson, *J. Phys. Chem. B* **115**, 4099-4109 (2011).
9. M. Weik, J. P. Colletier, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 437-446 (2010).
10. D. Vitkup, D. Ringe, G. A. Petsko, M. Karplus, Solvent mobility and the protein 'glass' transition. *Nat Struct Biol* **7**, 34-38 (2000).
11. M. Tarek, D. J. Tobias, Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys Rev Lett* **88**, (2002).
12. W. Doster, The dynamical transition of proteins, concepts and misconceptions. *Eur. Biophys. J. Biophy.* **37**, 591-602 (2008).
13. P. W. Fenimore *et al.*, *Chem. Phys.*, (2013).
14. J. R. Lewandowski, *Acc. Chem. Res.* **46**, 2018-2027 (2013).
15. N. Giraud *et al.*, *J. Am. Chem. Soc.* **127**, 18190-18201 (2005).
16. D. Idiyatullin, I. Nesmelova, V. A. Daragan, K. H. Mayo, *J. Mol. Biol.* **325**, 149-162 (2003).
17. M. J. Seewald *et al.*, *Protein Sci.* **9**, 1177-1193 (2000).
18. E. R. Andrew, R. Gaspar, W. Vennart, *Biopolymers* **17**, 1913-1925 (1978).
19. A. L. Lee, A. J. Wand, *Nature* **411**, 501-504 (2001).
20. G. Bouvignies *et al.*, *P. Natl. Acad. Sci. U.S.A.* **102**, 13885-13890 (2005).
21. K. Wood *et al.*, *J. Am. Chem. Soc.* **130**, 4586-4587 (2008).
22. J. R. Lewandowski, H. J. J. Sass, S. Grzesiek, M. Blackledge, L. Emsley, *J. Am. Chem. Soc.* **133**, 16762-16765 (2011).
23. P. R. L. Markwick, G. Bouvignies, M. Blackledge, *J. Am. Chem. Soc.* **129**, 4724-4730 (2007).
24. L. Mollica *et al.*, *J. Phys. Chem. Lett.* **3**, 3657-3662 (2012).

Acknowledgments: We thank Stephan Grzesiek and Hans J. Sass for providing [U-^{13}C,^{15}N] GB1. We acknowledge support from Agence Nationale de la Recherche (ANR PCV 2007 Protein Motion, ComplexDynamics ANR Blanc 2012 SIMI7), the Access to Research Infrastructures Activity in the Seventh Framework Program of the EC (BioNMR 261863), and Swiss National Science Foundation Grant 31-132857. JRL acknowledges support from EU IRG

(PIRG03-GA-2008-231026) and startup funds from University of Warwick. All data used to support conclusions in this manuscript are provided in the supplementary materials.

Supplementary Materials:

Materials and Methods

Supporting text on relationship to observations from other techniques and the origins of NMR line broadening at low temperatures.

Figures S1-S6

Tables S1-S4

Raw relaxation data (zipped Excel file)

References (*25-79*)