Assessing the Quality of Solution Nuclear Magnetic Resonance Structures by Complete Cross-Validation

Axel T. Brünger, G. Marius Clore, Angela M. Gronenborn, Rainer Saffrich, Michael Nilges

Structure determination of macromolecules in solution by nuclear magnetic resonance (NMR) spectroscopy involves the fitting of atomic models to the observed nuclear Overhauser effect (NOE) data. Complete cross-validation has been used to define reliable and unbiased criteria for the quality of solution NMR structures. The method is based on the partitioning of NOE data into test sets and the cross-validation of statistical quantities for each of the test sets. A high correlation between cross-validated measures of fit, such as distance bound violations and NMR *R* values, and the quality of solution NMR structures was observed. Less complete data resulted in poorer satisfaction of the cross-validated measures of fit. Optimization of cross-validated measures of fit will likely produce solution NMR structures with maximal information content.

An increasing number of solution NMR structures of small proteins have been reported during the past few years (*1*), as well as structures of medium-size proteins with 130 to 180 amino acid residues (*2*). This has been made possible through the development of multidimensional correlation NMR and isotope labeling techniques, sequential crosspeak assignment strategies, and improved computational algorithms for three-dimensional structure determination (*3*). Despite this progress, no generally accepted criteria for assessment of the quality of solution NMR structures are available.

The principal source of information for solution NMR structures is the NOE, complemented by coupling constant measurements, amide hydrogen exchange data, and knowledge about covalent geometry and nonbonded interactions. The NOE is a dipolar cross-relaxation phenomenon between protons, which is a function of interproton distances and molecular motions: It exhibits an r^{-6} dependence for an isolated proton pair that is separated by a distance r. Because of indirect cross-relaxation pathways through neighboring protons ("spin diffusion") and motional effects, the quantitative interpretation of NOEs for macromolecules is difficult. Therefore, NOEs for protein structure determination are often qualitatively interpreted by assigning distance ranges of approximately 1.8 to 2.7 Å, 1.8 to 3.3 Å, and 1.8 to 5.0 Å to strong, medium, and weak NOEs, respectively (*4*, *5*). This distance-based approach suffers from information loss. The intensity-based complete matrix relaxation theory (*6*, *7*) can in principle overcome this problem. At present, this approach is not always used because accurate NOE intensity measurements can be difficult to obtain and the motional parameters required for the theory are somewhat uncertain.

Cross-validation is a statistical method that estimates the quality of the fit to the observed data without making any assumptions about the distribution of errors in the data (*8*). It is based on computer "experimentation" wherein a certain subset (the "test" set) of the data is omitted while a model is fitted against the remaining data (the "working" set). Measures of fit are then cross-validated by evaluation of the agreement between the model and the test set. Cross-validation estimates the predictability of the test set given knowledge of the working set. Successful application of cross-validation includes crystallographic structure determination where the cross-validated ("free") *R* factor is highly correlated with the model's phase accuracy (*9*).

The NOEs are much more specific than x-ray diffraction data. A single reflection in a diffraction experiment contains information relating to the whole crystal structure. In contrast, each NOE is specific to a pair of atoms. Furthermore, certain NOEs carry much more information regarding the tertiary structure than others: for example, NOEs between secondary structural elements versus intraresidue or sequential NOEs. Consequently, cross-validation with a single test set is not appropriate for solution NMR structure determination. Here we introduce a method, which we termed "complete" cross-validation, where the distance data is randomly partitioned into 10 test sets of roughly equal size (*10*). Statistical quantities are cross-validated for each of the test sets and averaged. This procedure bears some similarity to the jackknife technique (*8*) except that more than one data point is omitted at a time. Complete cross-validation can be applied at any stage of NMR structure determination, including distance geometry, distance-restrained simulated annealing (SA), and intensity-based refinement.

The observed NMR data have to be augmented with knowledge about covalent bonding geometry and nonbonded interactions. The atomic coordinates can be minimized against a cost function

$$E_{\mathrm{NMR1}} = E_{\mathrm{chemical}} + w_{\mathrm{distance}}E_{\mathrm{distance}} + w_{\mathrm{torsion}}E_{\mathrm{torsion}} \quad (1)$$

Here E_{chemical} is a geometric or empirical energy function (*11*), and E_{distance} contains NOE-derived distance information, represented as a square-well function (*5*)

$$E_{\mathrm{distance}} = \sum_{\mathrm{NOEs}} \begin{cases} (R - d_{\mathrm{upper}})^2 & d_{\mathrm{upper}} < R \\ 0 & d_{\mathrm{lower}} < R < d_{\mathrm{upper}} \\ (d_{\mathrm{lower}} - R)^2 & R < d_{\mathrm{lower}} \end{cases} \quad (2)$$

The sum is carried out over all NOEs (where R is the distance between the pair of protons between which the NOE is observed, d_{lower} is the van der Waals distance between the protons, and d_{upper} is the NOE-derived upper bound) and E_{torsion} comprises torsion angle restraints derived from coupling constants, which has a functional form identical to Eq. 2 with distances replaced by torsion angles (the weighting factors w_{distance} and w_{torsion} are used to convert E_{distance} and E_{torsion} into energies).

The NMR structures can be obtained by minimization of Eq. 1 with a variety of distance geometry and SA approaches (*3*, *12*). In general, several structures exist that agree with the NMR data and the chemical knowledge equally well. The precision of these structures can be estimated by evaluating the atomic root-mean-square (rms) differences from their mean positions (*13*). Precision must be distinguished from accuracy, the rms difference between the ensemble of structures and the true structure. Of course, the true structure is unknown unless the "observed" data was calculated from a model structure. In fact, an absolute measure for the accuracy cannot be obtained in a real case, and it has to be estimated by some statistic. Precision overestimates accuracy (*14*) for the immunoglobulin G (IgG) binding domain of streptococcal protein G (*15*) (Fig. 1). A partial data set consisting of a random selection of 12.5% of the original NOE data (*16*) produced a mean structure with a backbone precision of 1.45 Å (Fig. 1A) but a backbone accu-

A. T. Brünger, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511.
G. M. Clore and A. M. Gronenborn, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892.
R. Saffrich and M. Nilges, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Germany.

racy of only 2.8 Å (Fig. 1B). Whereas the overall topology and the local secondary structural elements are similar to the well-determined structure, the directions and positions of the secondary structural elements show large differences (Fig. 2). Complete cross-validation maintains the integrity of the mean structure for both well and poorly determined cases; only a small effect on backbone precision and accuracy is observed (compare dashed and solid lines in Fig. 1).

The NMR structure of the IgG binding domain of protein G is influenced by $w_{distance}$. The distance bound violation (Fig. 3A) decreases for larger values of $w_{distance}$, whereas the deviation of bond angles from ideality (Fig. 3C), a measure of the geometric quality of the model, increases. The accuracy (Fig. 3D) reaches a plateau at $\log(w_{distance}/w^0_{distance}) \approx 2$; that is, distance-based structure determination becomes largely independent of $w_{distance}$ once the plateau is reached. Relatively small deviations from ideal geometry are observed (Fig. 3C) because the square-well function (Eq. 2) used shows no preference for a particular distance value within the specified distance bounds.

The cross-validated distance bound violation is a measure for the quality of the NMR structures: It increases as the data set gets poorer (Fig. 3B). Cross-validated distances are less well determined in poorer data sets despite similar agreement for the distances that are included in the structure determination (compare Fig. 3, A and B). Consider, for example, a small molecule structure that can be perfectly fit to 10 distance restraints. Complete cross-validation determines the degree to which each distance can be predicted by the remaining ones. A particular distance would be poorly predicted if the remaining ones do not fully determine the structure.

The cross-validated distance bound violation also shows a correlation with the model's accuracy as a function of $w_{distance}$: It tends to decrease until it reaches an approximate plateau at $\log(w_{distance}/w^0_{distance}) = 2$ (Fig. 3B). The fluctuations of the cross-validated distance bound violation result from the use of different partitionings for complete cross-validation at each sampled value of $w_{distance}$. One can interpret these fluctuations as a measure of the standard deviation of the cross-validated distance bound violation.

Refinement based on NOE intensities can be viewed as minimization of a target function (*7*, *17*)

$$E_{NMR2} = E_{chemical} + w_{NOE}R'_{1/6} \quad (3)$$

where $R'_{1/6}$ is the weighted residual (*18*)

$$R'_{1/6} = \sum_{NOEs} (I^{1/6}_{obs} - kI^{1/6}_{calc})^2 \quad (4)$$

The scale factor k is determined by minimization of Eq. 4, and I_{obs} and I_{calc} are the observed and calculated NOE intensities. The progress of this intensity-based refinement can be monitored by the NMR $R_{1/6}$ value (*6*, *18*)

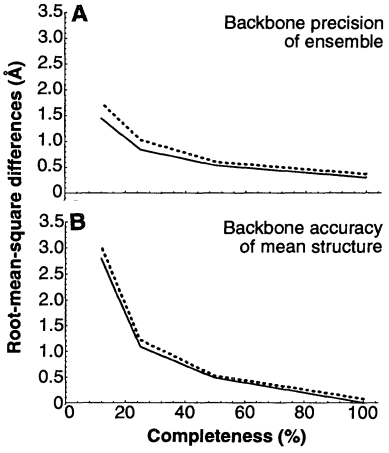

Fig. 1. (**A**) Precision and (**B**) accuracy of solution NMR structures of the 56-residue IgG binding domain of protein G (*15*) versus completeness of the distance data with (dashed lines) and without (solid lines) cross-validation. Completeness (*16*) was reduced by randomly eliminating distances from the full data set. Complete cross-validation (*10*) was used for the partial data sets; for example, for the 12.5% partial data set, the size of the test sets was 1.25% of the full data set. Ensembles of 80 structures were obtained by a hybrid distance geometry and SA protocol (*12*) starting from different initial conditions. The 12.5% partial data set corresponded to about two distances per residue and could therefore still be handled by distance geometry algorithms (*28*). Precision was determined for backbone atoms by the rms difference between the ensemble and the mean structure. Accuracy was estimated for backbone atoms by computation of the rms difference between a reference structure (mean structure of an ensemble obtained from the full data set with $w_{distance}$ = 100 kcal mol^{-1} Å^{-2}) and the mean structure. This is a reasonable assessment of accuracy as the reference structure almost perfectly fits the distance and torsion angle data. Thus, the observed data can be viewed as being calculated from the reference structure. By definition of the reference structure, the accuracy goes to zero for the full data set.

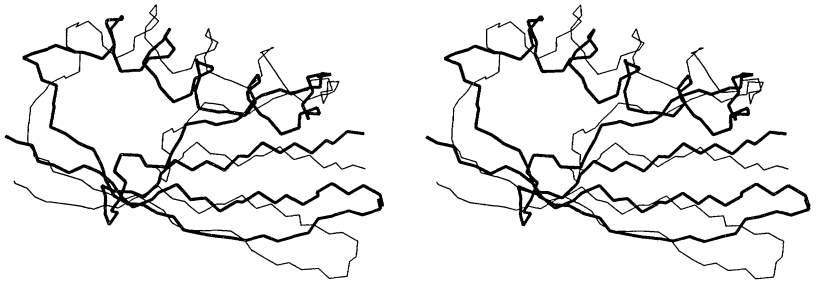

Fig. 2. Comparison of the mean structure of the IgG binding domain of protein G obtained from an ensemble of 80 structures with the full distance data set (*16*) (thick lines) and that obtained with a partial data set of only 12.5% of the distance data (thin lines). Stereo view of Cα, N, and C backbone atoms are shown.

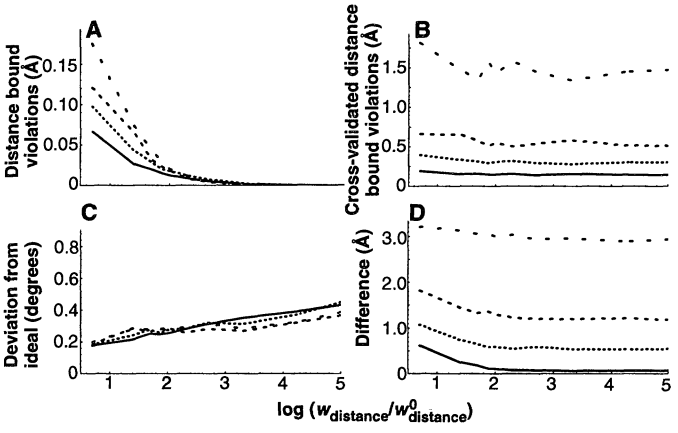

Fig. 3. Influence of the completeness of the data set and $w_{distance}$ (Eq. 1) on the quality of the NMR structures of the IgG binding domain of protein G; $w^0_{distance}$ = 1 kcal mol^{-1} Å^{-2}. Solid lines, full data set; small dashes, partial data set composed of 50% of the full set; medium dashes, 25% of the data set; and large dashes, 12.5% of the data set. Partitioning of the data sets for complete cross-validation was randomly shuffled for each value of $w_{distance}$. (**A**) Root-mean-square distance bound violations (*29*) for the partial data sets; (**B**) cross-validated (*10*) distance bound violations; (**C**) rms deviations of bond angles from values expected for ideal geometry; and (**D**) accuracy as measured by the atomic rms differences for backbone atoms between the mean structure and the reference structure used in Fig. 1. The overall correlation coefficient between the cross-validated distance bound violation (B) and the accuracy (D) for all values of $w_{distance}$ and all data sets was 0.99, whereas those for the 100, 50, 25, and 12.5% data sets were 0.90, 0.93, 0.76, and 0.66, respectively.

$$R_{1/6} = \sum_{\text{NOEs}} |I_{\text{obs}}^{1/6} - kI_{\text{calc}}^{1/6}| / \sum_{\text{NOEs}} I_{\text{obs}}^{1/6} \quad (5)$$

As in x-ray crystallography (9), the $R_{1/6}$ value for an atomic model refined against E_{NMR2} is a function of the weight w_{NOE}. Thus, the $R_{1/6}$ value is meaningless unless a rule is specified for choosing w_{NOE}.

Intensity-based refinement depends crucially on the accuracy of the observed intensities. This is illustrated for the solution structure of the squash trypsin inhibitor CMTI (*19, 20*). To obtain an objective measure of accuracy of the NMR structure, we (*21*) calculated a realistic NOE data set from an atomic model of CMTI, added a realistic amount of noise, and refined a family of four structures against this simulated data set (Fig. 4). The same NOE intensities had actually been measured previously and used for the refinement of the solution structure of CMTI (*20*). A w_{NOE} value of 2000 kcal mol^{-1} optimized the cross-validated $R_{1/6}$ value (Fig. 5). The refined structures exhibit a high backbone precision (0.2 Å) but the backbone accuracy (0.6 Å) is quite poor. The limited accuracy of the refined structures is a consequence of the noise; that is, a noise-free data set produces a backbone accuracy of 0.17 Å (not shown).

The $R_{1/6}$ value (Fig. 5A) decreases as a function of w_{NOE}, whereas the deviation of bond angles from the ideal (Fig. 5C) increases. The backbone precision assumes a minimum at $\log(w_{\text{NOE}}/w_{\text{NOE}}^0) = 4$ (Fig. 5D). One could argue that choosing w_{NOE} at this minimum produces a reasonable structure because the deviations of bond angles from ideality are about 2°, which is the observed standard deviation of bond angles in atomic-resolution crystal structures of small molecules (*22*). However, the backbone accuracy for this structure is significantly worse than that for the most accurate structure at $\log(w_{\text{NOE}}/w_{\text{NOE}}^0) = 3$; that is, the NOE data have been overfit. In contrast, the cross-validated $R_{1/6}$ value assumes a minimum near the most accurate structure. Thus, optimization of the cross-validated $R_{1/6}$ value as a function of w_{NOE} can be used in a real case in which accuracy cannot be assessed.

Cross-validated measures of fit, such as distance bound violations or $R_{1/6}$ values, are criteria for the quality of solution NMR structures. These quantities also assess the completeness of the NOE data set; they indicate poorer fit for less complete data sets. Complete cross-validation is not restricted to well-determined systems: Tests carried out with the IgG binding domain of protein G with only two distances per residue produced relatively small fluctuations of the cross-validated distance bound violation (Fig. 3B). Furthermore, the method can be applied to proteins that exhibit little or no secondary structure, such as CMTI (Fig. 4). Complete cross-validation could be useful to assess the sensitivity of a solution NMR structure with respect to a particular distance, for example, by computation of cross-validated local R values. Finally, by comparing cross-validated measures of fit between single conformer models and multiple conformer models (*23*), one could assess if the latter models improve the information content.

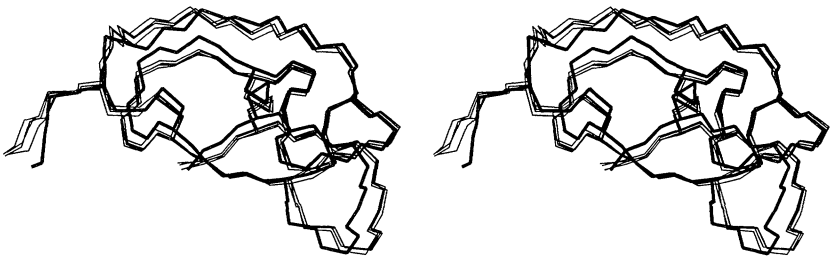

Fig. 4. Influence of noise on the NOE-intensity refined structure of the squash trypsin inhibitor CMTI (*19, 20*). Four structures with slightly different starting conformations were refined (*30*) against a noisy NOE intensity data set that was calculated from one of the initial structures ($cmti_1$) (*21*). In order to achieve convergence, a long refinement protocol was required. The protocol consisted of 20 ps of molecular dynamics at 1000 K followed by 20 steps of conjugate gradient minimization with w_{NOE} = 2000 kcal mol^{-1} (Eq. 3). This weight optimized the cross-validated $R_{1/6}$ value. The $cmti_1$ structure (thick lines) exhibits an $R_{1/6}$ value of 0.092 with deviations of bond angles from ideality of 1.67°. The average $R_{1/6}$ value of the refined structures (thin lines) is 0.077, the rms difference for backbone atoms to $cmti_1$ is 0.6 Å, and the rms difference for backbone atoms around the mean structure (precision) is 0.2 Å. Stereo view of Cα, N, C backbone and Cβ, Sγ Cys side chain atoms are shown.

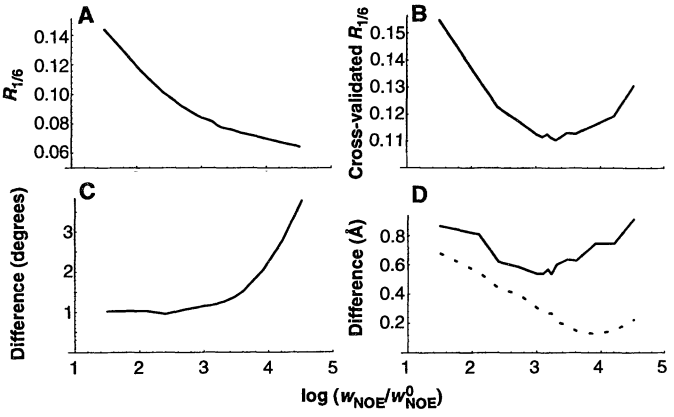

Fig. 5. Influence of the weight w_{NOE} (Eq. 3) on intensity-based refinement of CMTI for the simulated data set calculated from structure $cmti_1$ (*21*) and for four refined starting structures (compare with Fig. 4); w_{NOE}^0 = 1 kcal mol^{-1} and the NOE intensities were specified in dimensionless units. The E_{chemical} function (Eq. 3) was identical to the one described in (*12*). (**A**) Weighted R value (Eq. 5); (**B**) cross-validated weighted R value; (**C**) rms deviations of bond angles from ideality; and (**D**) accuracy as measured by the rms difference for backbone atoms between the refined structures and $cmti_1$ (solid lines) and precision as measured by the rms difference for backbone atoms between the refined structures and their mean structure (dashed lines). Cross-validation had little influence on the quantities in (A), (C), and (D). Complete cross-validation consisted of a random partitioning of the NOE data set for CMTI into 10 sets of roughly equal size. Ten structure refinements with an SA protocol (0.6 ps at 1000 K followed by 20 steps conjugate gradient minimization) were performed with one of the test sets taken out each time. The SA protocol ensured that memory toward the test set was removed. This procedure was repeated for the other three starting structures. The partitioning was randomly shuffled for each starting structure and each series of calculations for a particular value of w_{NOE}. The $R_{1/6}$ value was averaged for each value of w_{NOE}. The correlation coefficient between the cross-validated $R_{1/6}$ value (B) and the model's accuracy (D) is 0.8.

REFERENCES AND NOTES

1. W. A. Hendrickson and K. Wüthrich, *Macromolecular Structures 1991: Atomic Structures of Biological Macromolecules Reported During 1990* (Current Biology, London, 1991).
2. G. M. Clore, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **30**, 2315 (1991); L. J. Smith *et al.*, *J. Mol. Biol.* **224**, 899 (1992); R. Powers *et al.*, *Science* **256**, 1673 (1992); M. Ikura *et al.*, *ibid.*, p. 632; Y. Thériault *et al.*, *Nature* **361**, 88 (1992).
3. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986); R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon, Oxford, 1987); G. M. Clore and A. M. Gronenborn, *Annu. Rev. Biophys. Biophys. Chem.* **20**, 29 (1991); *Science* **252**, 1390 (1991); A. T. Brünger and M. Nilges, *Q. Rev. Biophys.*, in press.
4. K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* **169**, 949 (1983); M. P. Williamson, T. F. Havel, K. Wuthrich, *ibid.* **182**, 295 (1985).
5. G. M. Clore *et al.*, *EMBO J.* **5**, 2729 (1986).

6. T. L. James, *Curr. Opin. Struct. Biol.* **1**, 1042 (1991).
7. A. M. J. J. Bonvin, R. Boelens, R. Kaptein, in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, vol. 2, W. F. van Gunsteren, P. K. Weiner, A. J. Wilkinson, Eds. (ESCOM Science, Leiden, Netherlands, 1993).
8. M. Stone, *J. R. Stat. Soc. Ser. B* **36**, 111 (1974); B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982); B. Efron, *SIAM Rev.* **30**, 421 (1988).
9. A. T. Brunger, *Nature* **355**, 472 (1992); *Acta Crystallogr. Sect. D* **49**, 24 (1993).
10. Complete cross-validation consisted of a random partitioning of the distance data (*16*) into 10 sets of roughly equal size. Ten working sets were defined that excluded each one of the test sets from the full distance data set, one at a time. Structure determinations with a hybrid distance geometry and SA protocol (*12*) were performed with each of the working sets. A structure was accepted if the distance bound violation for the working set was less than 1 Å and the dihedral angle violation less than 10°. If one of these conditions was violated, the protocol was repeated with the same working set but different initial conditions (random number seed for distance geometry and initial velocities for SA) until the conditions were satisfied. Statistical quantities, such as distance bound violations, were averaged over the successful runs. Complete cross-validation was repeated with eight different partitionings of the data set and the statistical quantities averaged.
11. B. R. Brooks *et al.*, *J. Comput. Chem.* **4**, 187 (1983); A. T. Brunger, *Annu. Rev. Phys. Chem.* **42**, 197 (1991).
12. A hybrid distance-geometry and SA method (*24*) was used to determine conformations of the molecule consistent with both the experimental data and the chemical restraints (Eq. 1). The SA stage reduces bias attributable to insufficient sampling during the distance-geometry stage. Metric matrix distance geometry was applied to substructures consisting of Cα, Hα, N, HN, C, Cβ, and Cγ atoms (*25*). The coordinates of all atoms of a particular residue were then set to the coordinates of template structures fitted to the substructure coordinates (*24*). The resulting all-atom model was optimized by SA. The energy constants, the van der Waals radii, and the temperature were varied during SA to achieve best convergence as described in X-PLOR, protocol DGSA (*24, 26*). The geometric energy function $E_{chemical}$ (Eq. 1) used the "parallhdg.pro" parameters in X-PLOR, version 3.1 (*26*), with the following final values for the conformational energy constants: 1000 kcal mol^{-1} Å^{-2} for bond terms and 500 kcal mol^{-1} rad^{-2} for angle and improper terms. The van der Waals potential was described by a quartic repulsive potential (*24*) with the hard-sphere radii set to 0.75 times their standard values in "parallhdg.pro" and an energy constant of 4 kcal mol^{-1} Å^{-2}. The final value for the energy constant $w_{torsion}$ (Eq. 1) was set to 200 kcal mol^{-1} rad^{-2}.
13. T. F. Havel and K. Wüthrich, *J. Mol. Biol.* **182**, 281 (1985).
14. G. M. Clore, M. A. Robien, A. M. Gronenborn, *ibid.* **231**, 82 (1993).
15. A. M. Gronenborn *et al.*, *Science* **253**, 657 (1991).
16. The observed distance data set consisted of 854 distance restraints derived from NOE experiments and 68 distance restraints for 34 hydrogen bonds involving slowly exchanging backbone amide protons (*15*). The former contained 291 long-range interresidue restraints ($|i - j| > 5$ for residue numbers *i,j*) that determine the tertiary fold of the structure. In addition, 105 torsion angle restraints derived from experimental NOE and three-bond coupling constant data were included in the calculations. All torsion angle restraints were used during cross-validation of the distance data set. The influence of these torsion angle restraints on the computed statistical quantities was less than a few percent because these rather generous restraints help convergence but have little influence on the final structures.
17. P. Yip and D. A. Case, *Chem. Phys. Lett.* **161**, 50 (1989).
18. P. D. Thomas, V. J. Basus, T. L. James, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1237 (1991); C. Gonzales, J. A. C. Rullmann, A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *J. Magn. Reson.* **91**, 659 (1991); S. A. White, M. Nilges, A. Huang, A. T. Brunger, P. B. Moore, *Biochemistry* **31**, 1610 (1992); J. M. Withka, J. Srinivasan, P. H. Bolton, *J. Magn. Reson.* **98**, 611 (1992).
19. T. A. Holak, D. Gondol, J. Otlewski, T. Wilusz, *J. Mol. Biol.* **210**, 635 (1989).
20. M. Nilges, J. Habazettl, A. T. Brünger, T. A. Holak, *ibid.* **219**, 499 (1991).
21. We simulated a total of 425 NOE intensities at a mixing time of 150 ms using $\langle r^{-3}\rangle^{-1/3}$ averaging for methyl groups, uniform isotropic tumbling with a rotational correlation time of 2.3 ns, and a 4 Å cutoff (*20*) for the computation of the relaxation pathways. The structure termed "*cmti*$_1$" in (*20*) was used to calculate an intensity data set for NOEs that were actually observed for CMTI (*20*). Noise was added to the calculated data set with a uniform random number generator; 100% noise was added to the weakest NOEs, 10% noise was added to the strongest NOEs, and the noise added to the intermediate NOEs was a linear function between 100% and 10% proportional to the NOE intensity value provided by the simulation. The noise added represents a realistic estimate of the current accuracy with which NOE intensities can be observed using multidimensional homo- and hetero-nuclear spectroscopy.
22. W. A. Hendrickson, *Methods Enzymol.* **115**, 252 (1985); R. Engh and R. Huber, *Acta Crystallogr. Sect. A* **47**, 392 (1991).
23. A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *Chem. Phys. Lett.* **157**, 289 (1989); *J. Mol. Biol.* **214**, 223 (1990).
24. M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* **229**, 317 (1988).
25. T. F. Havel, I. D. Kuntz, G. M. Crippen, *Bull. Math. Biol.* **45**, 665 (1983); J. Kuszewski, M. Nilges, A. T. Brünger, *J. Biomol. NMR* **2**, 33 (1992).
26. A. T. Brünger, *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
27. P. Yip and D. A. Case, *J. Magn. Reson.* **83**, 643 (1989).
28. C. M. Oshiro, J. Thomason, I. D. Kuntz, *Biopolymers* **31**, 1049 (1991).
29. Distance bound violations are computed as rms differences between computed distances and d_{upper} (Eq. 2) for distances that are larger than d_{upper}.
30. A complete matrix relaxation technique was used to directly refine the atomic coordinates against the NOE intensity data set (*6, 20, 27*). Relaxation matrix and derivative calculations used a mixing time of 150 ms, $\langle r^{-3}\rangle^{-1/3}$ averaging for methyl groups, uniform isotropic tumbling with a rotational correlation time of 2.3 ns, and a 4 Å cutoff (*20*) for the computation of the relaxation pathways. The target function was E_{NMR2} (Eq. 3).

31. Support by the Howard Hughes Medical Institute (A.T.B.), the National Cancer Institute's Biomedical Supercomputer Center [for computer time (A.T.B.)], and the intramural AIDS Directed Antiviral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.) is gratefully acknowledged.

2 April 1993; accepted 8 June 1993

Magnetic Field Signatures Near Galileo's Closest Approach to Gaspra

M. G. Kivelson,* L. F. Bargatze, K. K. Khurana, D. J. Southwood, R. J. Walker, P. J. Coleman, Jr.

Two large magnetic field rotations were recorded by the spacecraft Galileo 1 minute before and 2 minutes after its closest approach to the asteroid Gaspra. The timing and the geometry of the field changes suggest a connection with Gaspra, and the events can be interpreted as the result of the draping of the solar wind field around a magnetospheric obstacle. Gaspra's surface field is inferred to be within an order of magnitude of Earth's surface field, and its magnetic moment per unit mass is in the range observed for iron meteorites and highly magnetized chondrites. The location of the magnetic signatures suggests that perturbations are carried by waves in the magnetosonic-whistler mode with wavelengths between electron and ion gyro radii.

On 29 October 1991, the Galileo spacecraft made its closest approach to the asteroid 951 Gaspra. Gaspra is a small (mean radius ~7 km), oddly shaped body that orbits the sun at a mean distance of 2.2 astronomical units (AU) (*1*). It is classified as an S-type asteroid on the basis of reflectance spectra observed from Earth; the surfaces of these asteroids contain varying proportions of olivine and pyroxene and iron-nickel metal. It is a matter of current debate whether asteroids of this class are the parent bodies of chondritic meteorites or stony-iron meteorites (*1, 2*). The parents of stony-iron meteorites are metal-enriched fragments of asteroids that were chemically differentiated. As indicated by its reflectance spectra, Gaspra is unusually metal- and olivine-rich relative to the other S-class asteroids, suggesting that it is more

M. G. Kivelson and P. J. Coleman, Jr., Department of Earth and Space Sciences and the Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024.
L. F. Bargatze, K. K. Khurana, R. J. Walker, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024.
D. J. Southwood, Department of Physics, Imperial College of Science, Technology, and Medicine, London SW7 2BZ, United Kingdom.
*To whom correspondence should be addressed.