FOR THE RECORD

SARS-CoV heptad repeat 2 is a trimer of parallel helices

Jessica Celigoy, Benjamin Ramirez, and Michael Caffrey*

Department of Biochemistry & Molecular Genetics, University of Illinois at Chicago, Chicago, Illinois 60607

Abstract: In severe acute respiratory syndrome coronavirus, the envelope heptad repeat 2 (HR2) plays a critical role in viral entry. Moreover, HR2 is both the target for novel antiviral therapies and, as an isolated peptide, presents a potential antiviral therapeutic. The structure of HR2, as determined by NMR spectroscopy in the presence of the co-solvent trifluoroethanol (TFE), is a trimer of parallel helices, whereas the structure of HR2, as determined by X-ray crystallography, is a tetramer of anti-parallel helices. In this work, we added a nitroxide spin label to the N-terminal region of HR2 and used paramagnetic relaxation enhancement to assess the orientation of the HR2 helices under different solution conditions. We find that the relaxation effects are consistent with an orientation corresponding to a trimer of parallel helices in both the presence and absence of TFE. This work suggests that the different orientation and oligomerization states observed by NMR and X-ray are due to the 11 additional residues present at the N-terminus of the NMR construct.

Keywords: NMR; spin label; SARS; heptad repeat; virus entry

Introduction

In severe acute respiratory syndrome coronavirus (SARS-CoV), the envelope protein, termed S or spike, plays a critical role in viral entry.[1] In the first step of entry, S binds to the target cell receptor, human angiotensin-converting enzyme 2 (hACE2), and the virus enters the cell via receptor-mediated endocytosis.[1,2] In the second step of entry, proteolysis by proteases of the cathespin family trigger conformational changes within S that bring heptad repeat domains, termed HR1 and HR2, into contact, thereby providing the driving force for fusion of the viral and endosomal membranes.[3–5] Interestingly, the HR1 and HR2 domains of other enveloped viruses including Ebola, HIV, and influenza are thought to function in a similar manner and are targets of antiviral therapies.[6] Moreover, the HR2 of the HIV envelope is a potent inhibitor of HIV entry, presumably by disrupting the intramolecular HR1–HR2 interaction during the entry process.[7] On the other hand, the HR2 of S is a relatively modest inhibitor of SARS-CoV entry.[8,9]

Structural studies of the HR1 and HR2 of SARS-CoV demonstrate that they form a "six helix bundle" that is analogous to that formed by other viral envelope proteins such as HIV/SIV gp41.[3,4,8,10–13] More recently, the SARS-CoV HR2 has been studied in isolation to lend insight into prefusion conformations of envelope.[13–16] Interestingly, the previous studies of SARS-CoV HR2 have resulted in two different structures for the prefusion state.[13,14] As shown in Figure 1, the NMR structure, in which the
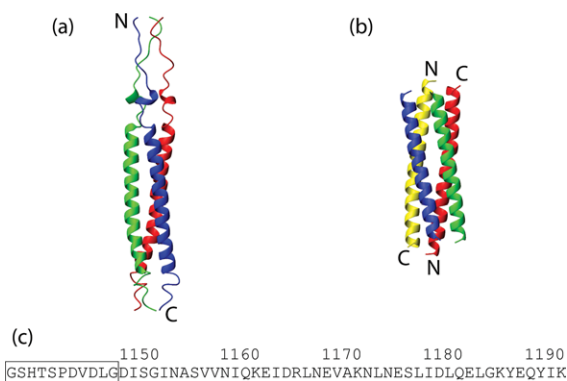
Figure 1. (a) NMR structure of SARS-CoV HR2 (residues 1141–1193 of S). (b) X-ray structure of SARS-CoV HR2 (residues 1150–1193 of S). The buffer conditions for the NMR structure were 10 mM NaPO$_4$, pH 7.0, 30% TFE.[14] The buffer conditions for the X-ray crystallization were 100 mM Tris-HCl, pH 8.5, 10 mM NiCl$_2$ with 22% PEG MME 2000.[13] (c) Amino acid sequence of SARS-CoV HR2. The additional residues present in the NMR construct are boxed. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

co-solvent trifluoroethanol (TFE) was present, indicated that the prefusion state of HR2 is a trimer of parallel helices, while the X-ray structure of HR2 is a tetramer of anti-parallel helices. In both cases, analytical ultracentrifugation confirmed the oligomeric state of the construct used. For example, the NMR construct was a trimer in the presence of 30% TFE and the X-ray construct was a tetramer under aqueous conditions.[13,14]

There could be several explanations for the disparity in the NMR and X-ray structures. First, the NMR structure was determined in the presence of TFE, a co-solvent used to stabilize helical structures[17] and thus the oligomerization state and relative orientation of the helices could be induced by the co-solvent. However, analytical ultracentrifugation studies of the NMR construct under aqueous conditions (i.e., in the absence of TFE) suggested that HR2 was in equilibrium between monomeric and trimeric forms.[15] Second, the determination of a protein complex, the HR2 trimer, by NMR is relatively demanding because it is based on a small number of intermolecular NOEs that could possibly be misinterpreted. Finally, the NMR construct that we used contained an 11 residue extension at the N-terminus (9 native residues and 2 residues from a cloning artifact), which may affect the orientation and oligomerization state of HR2. Clearly, it is important to resolve the orientation and oligomerization states of HR2 for a better understanding of SARS-CoV envelope-mediated viral entry and the mechanism of HR2-based inhibitory peptides. In this work, we added a nitroxide spin label to the N-terminal region of HR2 and used paramagnetic relaxation enhancement to assess the orientation of the HR2 helices under different solution conditions.

Results and Discussion

As discussed in the introduction, there are significant differences in the structure of SARS-CoV HR2 determined by NMR and X-ray crystallography (Fig. 1). To determine whether HR2 is organized in a parallel or anti-parallel manner, we mutated residue S14, which is found in the N-terminal region of HR2, to cysteine. Subsequently, we attached the nitroxide spin label MTSL to the newly introduced cysteine to form S14C-MTSL. This has the effect of introducing a paramagnetic center to this region of the protein, which has previously been shown to result in increased line-widths of resonances within 25 Å of the spin label.[18,19] If the HR2 complex is in a parallel orientation, as determined by NMR, paramagnetic relaxation enhancement would be observed for resonances at the N-terminal region (i.e., those close in sequence to residue 14). On the other hand, if the HR2 complex is in an anti-parallel orientation, as determined by X-ray crystallography, paramagnetic relaxation enhancement would be observed for resonances at the C-terminal region, as well as the N-terminal region.

In a first step, we measured the ^{15}N-edited HSQC spectrum of S14C-MTSL in the presence of 30% TFE to test the validity of the previously determined NMR structure.[14] In Figure 2(a), the relative intensities of the ^{1}H–^{15}N correlations are plotted as a function of residue number. The resonances near position 14 in sequence, the site of the MTSL, exhibit greatly reduced intensities with respect to HR2 not containing the spin label under identical experimental conditions. This observation implies that in the presence of TFE HR2 is a trimer of parallel helices, as previously determined.[14]

As mentioned in the introduction, it is possible that the co-solvent TFE significantly affects the orientation and oligomerization state of HR2. Consequently, we measured the ^{15}N-edited HSQC spectrum of S14C-MTSL under aqueous conditions (i.e., in the absence of TFE). In Figure 2(b), the relative intensities of the ^{1}H–^{15}N correlations are plotted as a function of residue number. Again the resonances near position 14 in sequence, the site of the MTSL, exhibit greatly reduced intensities with respect to HR2 not containing the spin label under identical experimental conditions. This observation is again consistent with the notion of a trimer of parallel helices, which was previously based on NMR and analytical ultracentrifugation studies in the presence and absence of TFE.[14,15] Interestingly, under the aqueous conditions the intensity of other resonances is also affected to an intermediate degree. For example, the average intensity ratio for correlations corresponding to residues 30–55 is 0.85 and 0.53 for the TFE and aqueous conditions, respectively, suggesting long-range effects under aqueous conditions.

In Figure 2(c), the residues most affected by the spin label, defined by <20% relative intensity, are
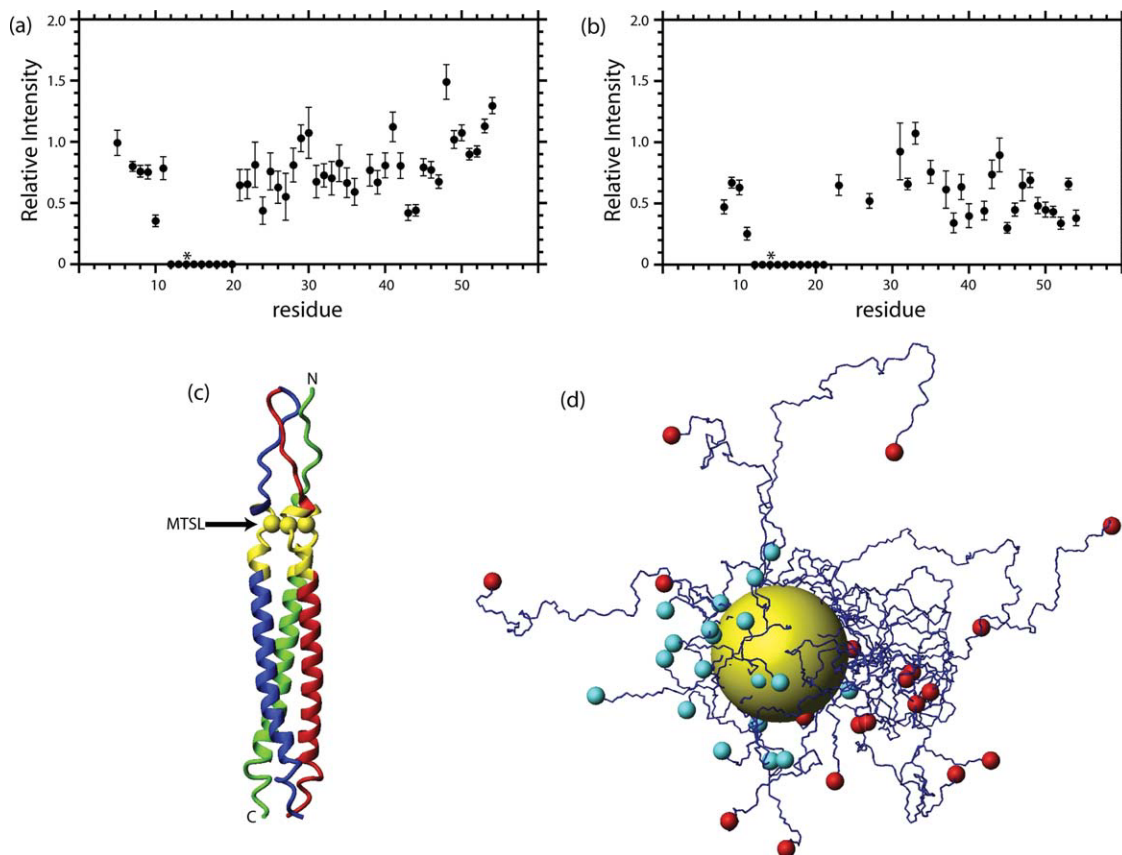
Figure 2. (a) Paramagnetic relaxation enhancement of S14C-MTSL as a function of residue number in the presence of TFE (the asterisk denotes the site of the MTSL). Experimental conditions were 1 mM S14C in 10 mM NaPO$_4$, pH 7.0, 30% TFE-d3 and 10% ^{2}H$_2$O at 25°C. (b) Paramagnetic relaxation enhancement of S14C-MTSL as a function of residue number in the absence of TFE (the asterisk denotes the site of the MTSL). Experimental conditions were 1 mM S14C in 10 mM NaPO$_4$, pH 7.0 and 10% ^{2}H$_2$O at 25°C. (c) SARS-CoV HR2 residues affected by paramagnetic relaxation enhancement. In this representation, individual monomers are colored red, green, and blue, respectively. ^{1}H–^{15}N correlations that exhibit less than 20% of the unlabeled intensity are depicted in yellow. The structure corresponds to that determined by NMR in the presence of 30% TFE.[14] (d) Model of SARS-CoV monomers in the extended coil conformation. Structures were fit to the backbone atoms of residue 14 with an RMSD of 0.003 Å. The yellow sphere corresponds to a 15 Å radius from the residue 14 Cα. The cyan and red spheres depict the N- and C-terminal residues, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

shown in yellow using the NMR structure determined in the presence of 30% TFE. Based on Cα–Cα distances, the residues most affected are <15 Å from the spin label site. Importantly, a similar effect is observed under aqueous conditions, which strongly supports the parallel orientation of SARS-CoV HR2. The differences between the NMR and X-ray structures may therefore be attributed to differences in the constructs. For example, the NMR construct contains nine additional native residues at the N-terminus with respect to the X-ray construct and it would seem that the additional residues push the equilibrium to the trimer of parallel helices. The propensity of heptad repeat regions to self-associate in different orientations and oligomerization states indicates that caution should be exercised in the interpretation of isolated domains. On the other hand, the ability of HR2 to assume different structures may have functional ramifications for the entry mechanism.[13,15,16,20]

As noted above, intermediate paramagnetic relaxation enhancement was observed for HR2 residues that are sequentially distant from the spin label under aqueous conditions. Previously, we have shown by CD, NMR, and analytical ultracentrifugation that HR2 exhibits a monomer–trimer equilibrium and that the monomer is unstructured.[15,16,20] As a consequence, the intermediate effects shown in Figure 2(b) may be attributed to the approach of distant residues to the spin label attached to residue 14. To test the validity of this notion, we determined a family of HR2 structures without using experimentally determined NOE, H-bond, and dihedral restraints. A family of 20 monomer structures is shown in Figure 2(d). In this figure, the site of the spin label is depicted by a sphere of 15 Å, the approximate radius of correlations that are completely abolished in the ^{15}N-edited HSQC spectra. The respective locations of the N- and C-terminal residues

are depicted by cyan and red spheres, respectively. From this figure it is clear that the C-terminal residues that exhibit intermediate relaxation effects may often approach the local region of the spin label. For example, from a family of 100 HR2 structures the range of Cα–Cα distances observed for residues 20, 30, 40, and 50 to residue 14 is 8–20, 16–51, 15–68, and 11–80 Å, respectively. Based on the molecular dynamics simulations, it is predicted that the closest approach of residues 30, 40, and 50 is similar and thus the relatively uniform effects on relaxation may not be surprising. Finally, we note that we have previously suggested that the flexibility of the unstructured state facilitates conformational changes that are necessary for envelope-mediated membrane fusion.[15,16,20]

Materials and Methods

Protein preparation

Cysteine was introduced at residue 14 of SARS-CoV HR2 using the site-directed mutagenesis kit from Strategene and subsequently confirmed by DNA sequencing. The mutated version of HR2, termed S14C, was subcloned into the *BamH1*/*HindIII* restriction sites of a modified pQE30-expression vector (Qiagen, Valencia, CA). The resulting construct consists of an N-terminal polyhistidine tag, followed by PG (the IgG binding domain of streptococcus protein G),[21] a TEV (tobacco etch virus) cleavage site for removal of the expression tag, and S14C. ^{15}N-labeled S14C was prepared as previously described.[14,15] Briefly, protein expression was achieved by growing *E. coli* strain SG13009 in the presence of the appropriate plasmid in 1 L of LB media supplemented with 100 μg/mL ampicillin and 50 μg/mL of kanamycin at 37°C until they reached an OD_{600} of 0.8. The cells were then pelleted, washed once with M9 salts, re-suspended in 250 mL of M9 minimal media supplemented with 1 g/L [15]NH$_4$Cl (Martek Biosciences, Columbia, MD), and set to recover in the absence of antibiotic selection. Protein expression was induced after 1 h by the addition of 0.8 mM IPTG and grown for an additional 4–5 h at 37°C. The HIS-PG-S14C fusion protein was purified from the soluble fraction using a Ni^{2+} fast-flow Sepharose column (Qiagen, Valencia, CA). The protein was then cleaved using TEV protease and run once more over the Ni^{2+} column to remove His-PG and TEV protease, which also contains a polyhistidine tag, as well as uncut HIS-PG-S14C. The flow-through fraction containing S14C was then dialyzed extensively against 10 mM NaPO$_4$, pH 7.0, and concentrated by ultrafiltration (YM3, Amicon, Billerica, MA). The purity and identity of S14C was confirmed using SDS-PAGE and MALDI-TOF mass spectrometry. The newly introduced disulfide bond was reduced using a 5× molar excess of DTT. Excess reductant was removed using a Micro-Bio Spin chromotography column (Bio-Rad, Hercules, CA), then the reduced protein was incubated with a 7× molar excess of MTSL (*S*-(2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl) methyl methanesulfonothioate) overnight at room temperature in 10 mM TRIS-HCl, pH 7.4, 0.02% sodium azide, to form S14C-MTSL. Finally, unbound MTSL was removed by dialysis against buffer containing 10 mM NaPO$_4$, pH 7.0.

NMR spectroscopy

The experimental conditions for the NMR experiments were 1 mM S14C-MTSL in 10 mM NaPO$_4$, pH 7.0, 30% TFE-d3 (Cambridge Isotope Laboratories, Andover, MA) and 10% ^{2}H$_2$O or alternatively in 10 mM NaPO$_4$, pH 7.0, 10% ^{2}H$_2$O. Spectra were recorded at 25°C on a Bruker AVANCE 800 MHz spectrometer equipped with a 5 mm TCI cryogenic triple resonance probe. Processing of the spectra was done with NMRPipe and visualized with NMRDraw.[22] The relative intensity was determined by comparing the intensity of S14C-MTSL correlations (I_{sl}) with those of the wild-type HR2 (I) under the identical experimental conditions. Errors in the relative intensity were estimated as $I_{sl}/I ((N_{Isl}/I_{sl})^2 + (N_I/I)^2)^{0.5}$, where N_{Isl} and N_I are the noise calculated by NMRDraw in the appropriate spectrum.[23]

Molecular modeling

Structures were calculated by simulated annealing in torsion angle space starting from the HR2 in an extended conformation, followed by conventional simulated annealing, using the program CNS.[24–26] A conformational database term was employed during the simulated annealing phase to favor energetically relevant dihedral angles.[27] No NOE, H-bond, or other dihedral restraints were used in the simulated annealing phase. Figures were generated using the program MOLMOL.[28]

References

1. Simmons G, Reeves J, Rennekamp A, Amberg S, Piefer A, Bates P (2004) Characterization of severe acute respiratory syndrome-associated coronavirus (SARS CoV) spike glycoprotein-mediated viral entry. Proc Natl Acad Sci USA 101:4240–4245.
2. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426:450–454.
3. Liu S, Xiao G, Chen Y, He Y, Niu J, Escalente C, Xiong H, Farmar J, Debnath A, Tien P, Jiang S (2004) Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification effusion inhibitors. Lancet 363:938–947.
4. Tripet B, Howard M, Jobling M, Holmes R, Holmes K, Hodges R (2004) Structural characterization of the SARS-coronavirus spike S fusion protein core. J Biol Chem 279:20836–20849.

5. Simmons G, Gosalia D, Rennekamp A, Reeves J, Diamond S, Bates P (2005) Inhibitors of cathespin L prevent severe acute respiratory syndrome coronavirus entry. Proc Natl Acad Sci USA 102:11876–11881.

6. Caffrey M (2011) HIV envelope: challenges and opportunities for the discovery of entry inhibitors. Trends Microbiol 9:191–197.

7. Matthews T, Salgo M, Greenberg M, Chung J, DeMasi R, Bolognesi D (2004) Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. Nat Rev Drug Discov 3:215–225.

8. Bosch B, Martina B, Van Der Zee R, Lepault J, Haijema B, Versluis C, Heck A, De Groot R, Osterhaus A, Rottier P (2004) Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc Natl Acad Sci USA 101:8455–8460.

9. Guo Y, Tisoncik J, McReynolds S, Lou G, Martinez O, Farzan M, Prabhakar B, Gallagher T, Rong L, Caffrey M (2009) Identification of a small region of SARS-CoV S protein critical for viral entry. J Mol Biol 394:600–605.

10. Supekar V, Bruckmann C, Ingallinella P, Bianchi E, Pessi A, Carfi A (2004) Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc Natl Acad Sci USA 101:17958–17963.

11. Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao G, Rao Z (2004) Crystal structure of the severe acute syndrome coronavirus spike protein fusion core. J Biol Chem 279: 49414–49419.

12. Duquerroy S, Vigouroux A, Rottier P, Rey F, Bosch B (2005) Central ions an lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. Virology 335:276–285.

13. Deng Y, Liu J, Zheng Q, Yong W, Lu M (2006) Structures and polymorhic interactions of two heptad-repeat regions of the SARS virus S2 protein. Structure 14:889–899.

14. Hakansson-McReynolds S, Jiang S, Rong L, Caffrey M (2006) The solution structure of the SARS coronavirus heptad repeat 2 in the prefusion state. J Biol Chem 281:11965–11971.

15. McReynolds S, Jiang S, Ying G, Celigoy J, Schar C, Rong L, Caffrey M (2008) Characterization of the prefusion and transition states of Severe Acute Respiratory Syndrome Coronavirus S2-HR2. Biochemistry 47: 6802–6808.

16. McReynolds S, Jiang S, Rong L, Caffrey M (2009) Dynamics of the SARS coronavirus prefusion and transition states. J Magn Reson 201:218–221.

17. Buck M (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. Q Rev Biophys 31:297–355.

18. Battiste JL, Wagner G (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. Biochemistry 39:5355–5365.

19. Park S, Caffrey M, Johnson M, Fung L (2003) Solution structural studies on human erythrocyte alpha spectrin N-terminal tetramerization domain. J Biol Chem 278: 21837–21844.

20. Celigoy J, McReynolds S, Caffrey M (2011) The SARS-CoV heptad repeat 2 exhibits pH-induced helix formation. Biochem Biophys Res Commun 412:483–486.

21. Huth J, Bewley C, Jackson B, Hinnebusch A, Clore G, Gronenborn A (1997) Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. Protein Sci 6: 2359–2364.

22. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293.

23. Jiang S, Jacobs A, Laue T, Caffrey M (2004) Solution structure of the Coxsackievirus and Adenovirus receptor domain 1. Biochemistry 43:1847–1853.

24. Stein EG, Rice LM, Brunger AT (1997) Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. J Magn Reson 124:154–164.

25. Nilges M, Clore GM, Gronenborn AM (1988) Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. FEBS Lett 229: 317–324.

26. Brunger A, Adams P, Clore G, DeLano W, Gros P, Grosse-Kunstleve R, Jiang J-S, Kuszewski J, Nilges N, Pannu N, Read R, Rice L, Simonson T, Warren G (1998) Crystallography and NMR system CNS: a new software system for macromolecular structure determination. Acta Cryst D54:905–921.

27. Kuszewski J, Gronenborn AM, Clore GM (1996) Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. Protein Sci 5:1067–1080.

28. Koradi R, Billeter, M, Wuthrich, K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14:29–32.