Tail Reconnection Triggering Substorm Onset

Vassilis Angelopoulos,1* James P. McFadden,2 Davin Larson,2 Charles W. Carlson,2 Stephen B. Mende,2 Harald Frey,2 Tai Phan,2 David G. Sibeck,3 Karl-Heinz Glassmeier,4 Uli Auster,3 Eric Donovan,5 Ian R. Mann,6 I. Jonathan Rae,6 Christopher T. Russell,1 Andrei Runov,1 Xu-Zhi Zhou,1 Larry Kepko7

1IGPP/ESS, UCLA, Los Angeles, CA USA. 2Space Sciences Laboratory, University of California at Berkeley, CA USA. 3Code 674, NASA/GSFC, Greenbelt, MD, USA. 4TUBS, Braunschweig, D-38106, Germany. 5Dept. of Physics and Astronomy, University of Calgary, Calgary, Canada. 6Department of Physics, University of Alberta, Edmonton, Alberta, Canada. 7Space Science Center, University of New Hampshire, Durham, NH, USA.

*To whom correspondence should be addressed. E-mail: vassilis@ucla.edu

Magnetospheric substorms explosively release solar wind energy previously stored in Earth’s magnetotail, encompassing the entire magnetosphere and producing spectacular auroral displays. It has been unclear whether a substorm is triggered by a disruption of the electrical current flowing across the near-Earth magnetotail, at ~10 R_E (R_E = Earth Radius, or 6374 km), or by the process of magnetic reconnection typically seen farther out in the magnetotail, at ~20-30 R_E. We report on simultaneous measurements in the magnetotail at multiple distances, at the time of substorm onset. Reconnection was observed at 20 R_E, at least 1.5 min before auroral intensification, at least 2 min before substorm expansion, and about 3 min before near-Earth current disruption. These results demonstrate that substorms are likely initiated by tail reconnection.

Substorms are global reconfigurations of the magnetosphere involving storage of solar wind energy in Earth’s magnetotail and its abrupt conversion to particle heating and kinetic energy (1, 2). Because phenomena related to the onset of substorms are initially localized (within 1-2 R_E) in space but expand quickly to engulf a large portion of the magnetosphere (3, 4), fortuitous (and thus unoptimized) conjunctions between single satellite missions have been unable to pinpoint the exact location of the substorm trigger in space. This has led to diverging theoretical efforts to explain the onset mechanism (5, 6). The key question is whether substorm phenomena are triggered by a near-Earth dipolarization (current disruption) process, at around 10 R_E, or by the process of magnetic reconnection at around 20-30 R_E. Both processes operate during substorms, but attempts (7, 8) to delineate the causal relationship between them and substorm onset were limited due to temporal resolution, spatial propagation effects and/or lack of simultaneous observations of the above key regions. The THEMIS mission (9, 10) was designed to address this question. The mission employs five identical satellites (hereafter termed probes) on orbits enabling recurrent probe alignments parallel to the Sun-Earth line (probes within δY_GSM±2 R_E from each other). The probes can thus monitor tail phenomena simultaneously at ~10 R_E and at ~20-30 R_E downtail, while mapping magnetically over a network of ground-based observatories (GOBs) which can determine the meridian and time of substorm onset on the ground (11). Here we present timing results from THEMIS for isolated substorms, which demonstrate that the substorm trigger mechanism is magnetic reconnection.

Substorm signatures. The ground signatures of substorms consist of a rapid auroral intensification, a breakup of auroral forms into smaller filaments, a poleward expansion and a westward surge of the most intense auroral arcs. Those are within 1-2 minutes of each other, and have often been used synonymously with substorm onset. Ground magnetic signatures of currents associated with auroral arc intensification include abrupt increases in the auroral electrojet (AE) index (12) and irregular pulsations in the 40-150 s range called Pi2s (13), or at lower periods (14), observed at high latitude (auroral) and mid-latitude (subauroral) magnetic stations. Such magnetic signatures are known to coincide with auroral intensification typically within 1-2 minutes (15) and are also used for substorm onset identification.

Prior to substorm onset, during the substorm growth phase, stable arcs intensify and move equatorward while the magnetotail plasma sheet thins and the cross-tail current increases (16–19). During the expansion phase auroral arcs typically advance toward the poleward edge of the auroral oval and a current wedge develops in space, at around 10 R_E, composed of field-aligned currents into and out of the ionosphere (17). Current wedge formation is also referred to as dipolarization, because the field becomes more dipole-like, or as current disruption, because it is consistent with a
disruption (reduction) of the duskward cross-tail current (S). Further downtail, fast tailward flows threaded by southward magnetic fields, or Earthward flows threaded by northward fields are observed near substorm expansion onset, and have been interpreted as evidence for magnetic reconnection (20, 27). Substorm expansion is followed by substorm recovery, during which auroral forms remain active at the poleward boundary of the auroral oval for hours, until they eventually reduce in intensity and move equatorward, often starting another substorm sequence. Arc intensification alone does not necessarily constitute a substorm, even though it may involve another substorm sequence. Arc intensification alone does not reduce in intensity and move equatorward, often starting at the boundary of the auroral oval for hours, until they eventually been interpreted as evidence for magnetic reconnection (fields are observed near substorm expansion onset, and have magnetic fields, or Earthward flows threaded by northward flows). Further downtail, fast tailward flows threaded by southward disruption (reduction) of the duskward cross-tail current (S).

**Substorm timing on 26 February 2008.** At 4:50 UT, the THEMIS probes were aligned along the Sun-Earth line, less than 1 RE from the nominal neutral sheet (Fig. 1). A sudden increase of AE\(_{11}\) to 200nT was observed at 04:54:00 UT, indicating an isolated substorm onset (Fig. 2A). Auroral station Gillam and mid-latitude station Carson City recorded P12 pulsations (Fig. 2B, 2C). The P12 pulsation onsets were determined as the times of the first increase in signal amplitude above background, at 04:52:00 UT and 04:53:05 UT respectively. Observations from the THEMIS All-Sky Imagers (ASIs), seen in Fig. 3F, Fig. 3G, and movies S1 and S2 (23) show that a relatively stable arc extended across the sky from Gillam to Sanikiluaq at 04:50:03 UT, and that by 04:53:03 UT the arc had intensified at Gillam (see arrow in Fig. 2G). The auroral brightening region was initially ~100 km in width (note that the ASI field of view is ~800 km when mapped to 110 km altitude). Using the inflection point of the auroral intensity increase at Gillam (Fig. 3D), we determined that the auroral intensification onset was at 04:51:39 UT. As seen in Fig. 2E and in movie S2 (23) the arc intensified at 67.8° geomagnetic latitude and expanded poleward of 68.2° at 04:52:21 UT; we denote the latter as the time of substorm expansion onset. These substorm onset times are summarized in Table 1.

For several minutes prior to the abrupt brightening at 04:51:39 UT, the arc developed small (50-100 km scale) filaments that moved along the arc and/or died down at these stations. Images from a NORSTAR imager at Rabbit Lake (just to the West of Gillam), seen in movie S3 (23), also captured these transient filaments and in one instance (at 04:50 UT), the filaments were seen in conjunction with a localized transient enhancement in ULF power at that station. Since auroral intensification onset was localized within the field of view of Gillam, but the filaments were developing for several minutes and over a 2-hr magnetic longitude range, we do not associate these filaments with the sought-after, abrupt substorm trigger. These filaments, however, may be related to pre-conditioning of the magnetosphere, leading up to substorm onset.

Using the T96 magnetospheric model (24), we projected the probe locations along magnetic field lines to the ionosphere. The probe footpoints lie near the west coast of Hudson Bay, i.e., near Gillam, and with 1 hr of MLT of the meridian of auroral intensification (the substorm meridian). A current wedge analysis of the mid-latitude magnetometer data confirms that the probe footpoints were within the substorm current wedge. Therefore, the probes were well positioned meridionally to examine the relative timing of substorm signatures on the ground and in space.

**Overview of tail signatures.** Probes P1 and P2 recorded a decreasing B\(_x\) between 04:45-05:01 UT (Fig. 3A and 3G), consistent with a decreasing current sheet thickness and an increasing current sheet density, as expected at substorm growth phase. The plasma sheet ion density and average energy, obtained from the data shown in Fig. 3D and 3J, were ~1 particle/cc and ~1 keV respectively. These are conditions of a cold-dense plasma sheet following prolonged intervals of northward interplanetary field (25). Probe P1 observed at ~04:54 UT tailward flows (V\(_x\)<0) accompanied by southward (B\(_y\)<0) and duskward (B\(_y\)>0) excursions of the magnetic field, followed at ~05:02 UT by Earthward flows and opposite polarity magnetic field perturbations. Probe P2 observed Earthward flows of the same nature and at approximately the same time as P1 (05:02 UT). The flow and field signatures at P1 are expected from a reconnection site first located Earthward, then retreating (or reappearing) tailward of P1, starting at ~05:01 UT. The observed B\(_x\) component variations (B\(_y\)>0 tailward of the reconnection site; B\(_x\)<0 Earthward of it) are also classical Hall signatures of reconnection (26–28).

Before the onset of the fast tailward flows, both probes observed two ion components: a 500 eV component, commensurate with the cold plasma sheet prior to the event, and a 10 keV component which appeared gradually. Both probes also observed relatively low temperature electrons (100-200 eV) of decreasing flux. Despite the plasma sheet thinning, P1 and P2 remained within the plasma sheet. As the fast tailward flows were observed at P1 the average energy of the ions and electrons increased to 10 keV and 1 keV respectively. After the tailward retreat of the reconnection site at 05:01 UT, both probes observed even hotter plasma (20 keV ions, 2 keV electrons) and crossed the neutral sheet as evidenced by the near-zero transitions of B\(_x\) at around 05:02
UT. This is evidence of plasma heating at the reconnection outflow and plasma sheet dipolarization at 05:02 UT at 22 R_E, the distance of P1.

If the flows are due to reconnection, the current sheet should resemble a slingshot-like, standing Alfvén wave(27, 28). To evaluate the shear stress balance, we examined the correlation between the measured ion flows, $\Delta V_A$, and the flows predicted from reconnection outflow, $\Delta V_A = \pm A B N_i^{1/2}$ on P1 and P2 ($V_A$: Alfvén speed, $B$: magnetic field, $N_i$: the ion density). The correlation coefficients for the tailward and Earthward flows on P1 (04:53:30-04:58:30; 05:01:00-05:13:00) and for the Earthward flows on P2 (05:01:30-05:06:00) were 0.79, 0.61, and 0.86; whereas the slopes were −0.53, 0.32, and 0.51, respectively. The slopes are likely underestimates because of temporal variations in the reconnection process and because we have not yet included energetic particles in the velocity determination.

Alternatively, the stress balance at the peak velocity near the center of the plasma sheet gives ratios: $|\Delta V_A|/|\Delta V_A| > 90\%$, 67% and 88% for the tailward and Earthward flows on P1 (04:54:30 UT, 05:02:00 UT) and the Earthward flows on P2 (05:02:30 UT) respectively. Under the caveat of the need to include the super-thermal ion corrections to the ion flow velocity, mostly required for the hotter, Earthward flows, the observations are consistent with the Alfvénic acceleration expected from reconnection.

P3 was near the neutral sheet. It observed fast (>400 km/s) Earthward flows starting at around 04:52:27 UT, followed by a transient increase in the Northward component of the magnetic field (transient dipolarization) at 04:53:05 UT. The onset of fast flows was followed by a more permanent dipolarization at 04:54:40 UT, signifying the development of a substorm current wedge in near-Earth space. The transient dipolarization at ~04:53:05 UT is interpreted as the first indication of a substorm current wedge at P3.

P5, near geosynchronous altitude, saw an energy-dispersed ion injection of the 50-200 keV ions (Fig. 3S). The flux increase at the lowest energy (highest flux) channel is an exception, as it responds to the local plasma and is correlated with convective velocity changes measured at the same time. The energetic particle dispersion (more energetic particles drifting faster than lower energy particles) is consistent with a duskward drift of those particles to the location of P5 after an injection near midnight. Such dispersed injections are classical signatures of substorms observed by geosynchronous orbit satellites (29).

Observations around the time of onset. The fast tailward reconnection flows ($V_z$<−100 km/s) on P1 started at ~04:52:30 UT. They were preceded by a northward convective flow ($V_z$>50 km/s) and an accompanying southward deflection of the magnetic field ($\delta B_z$<0), which we interpret as evidence of onset of reconnection inflow toward the neutral sheet at 04:50:28 UT. Ion velocity distributions (Fig. 4C) show two components: a relatively isotropic component below ~500 km/s and a duskward / tailward streaming component above 1000 km/s. These are the cold and hot ions seen earlier in the spectra of Fig. 3D. Similar behavior was found on probe P2. The anisotropy of the few keV ions intensified by 04:51:14 UT (see distribution function); even 1 keV ions (~310 km/s) exhibited pronounced duskward drift. This is consistent with a diamagnetic ion current. The inferred gradient scale is approximately the gyroradius of a 1 keV proton in the local field (~20nT) i.e., ~600 km. This is further evidence that the current sheet was thin, and the current density high.

Electron velocity space distribution functions (Fig. 4D) exhibit a bidirectional anisotropy prior to 04:50:54 UT, which intensified in the ensuing few minutes. As seen in Fig. 4E, 50-300 eV electrons were streaming toward the reconnection site (180° pitch angle, i.e., approximately Earthward), while 400-2000 eV electrons were streaming away from the reconnection site (0° pitch angle, i.e., approximately tailward). Such electron streaming is a signature of reconnection due to the Hall current system (21), suggesting that reconnection had started near the location of P1 by 04:50:28 UT.

Probe P2 was farther away from the neutral sheet than P1, as evidenced by the enhanced magnetic field (Fig. 3G) and ion energy spectra (Fig. 3I, 3J). No direct connection of field lines at P2 to the reconnection site was evident until after onset. However, similar to P1 at 04:50:28 UT, observations at P2 at 04:50:38 UT show the beginning of inflow toward the reconnection site ($V_z$>0, Fig. 4G), and the start of a positive deflection of $B_z$ along with a bipolar $B_y$. These are signatures of an Earthward flux transfer event (30, 31), signifying tail reconnection somewhere near the tailward of P2. The simultaneous deflection of $B_y$ northward at P1 and southward at P2 suggests that a reconnection topology was established at that time between the two probes, i.e., between 17 and 22 R_E, as depicted in Fig. 1.

Probe P3 observed a slow ramp-up of the Earthward flow velocity ($V_z$>50 km/s) at 04:52:27 UT (Fig. 3N; Fig. 4I), followed by a deflection of $B_z$ northward (dipolarization) and high speed flows. The slow flows seen at P3 prior to the dipolarization did not necessarily emanate directly from the reconnection region; rather they may have been nearby plasma that accelerated Earthward due to the forces from the establishment of a reconnection topology further downstream. The first signatures of dipolarization were timed at 04:53:05 UT.

Together, the observations at P1, P2 and P3 make a compelling case for onset of tail reconnection at or prior to t1=04:50:28 UT, between P1 and P2 (Table 1). Approximating the Alfvén speed near the reconnection site as
500 km/s (based on a local density measurement of 1 particle/cc and magnetic field of 20nT) we can infer the downtail location of the source, x0, and the time of reconnection onset, t0. Noting that the reconnection pulse can travel a distance of 5 Re (the P1-P2 inter-probe separation) in 60 s, we obtain \((t1-t0)+(t2-t0)=60\) s, resulting in \(t0=04:50:03\) UT and \(x0=20\) Re.

**Time history of events.** The inferred reconnection onset at 04:50:03 UT preceded the onset of auroral intensification by 96 s. Arc intensification was followed by high-latitude Pi2 onset, 20 s later. The high latitude Pi2 onset may signify the arrival of the field-aligned current pulse generated by the reconnection flows in the tail (15). It is unlikely that a shear Alfvén wave, starting at ~500 km/s, can travel from 20 Re to the ionosphere in 96 s, due to both the high density in the plasma sheet and the large distance to the source. Conversely, kinetic Alfvén waves (32) of an ion acoustic gyroradius scale may exceed the local Alfvén speed by a factor of \(\sqrt{2}\) and arrive faster. Those waves can also accelerate electrons (33), which may result in visible aurora ahead of the wave. This would explain the observations of arc intensification 20 s earlier than the Pi2 onset and would be consistent with reported observations of Alfvénic aurora (34).

Twenty one seconds after the high-latitude Pi2 onset, the aurorae started to expand poleward. This expansion started 6 s before the arrival of an Earthward flow perturbation at P3 and about 40 s prior to the dipolarization at P3. Therefore, the initial poleward motion of the aurora cannot be caused by the near-Earth flux pile-up of reconnection flows. It is likely associated with the change in magnetic field mapping, as reconnection at 20 Re results in the engulfment of higher magnetic latitude flux in the reconnection process. Flux pileup and current wedge formation may, however, be responsible for later stages of poleward expansion; careful modeling of the current wedge currents in realistic magnetotail fields is needed to properly address this question.

Across field lines, the reconnection process started to affect the inner magnetosphere at 11 Re (P3’s location) 144 s after onset. This time delay relative to reconnection onset is commensurate with the magnetosonic speed of the 1-4 keV plasma at the neutral sheet between P2 and P3, i.e., ~500 km/s. The first evidence of intense dipolarization, interpreted as the reconnected flux arrival at that same location, was seen 30 s after the first indication of Earthward flow and 3 minutes after reconnection onset at 20 Re. The latter is also commensurate with the simultaneously observed plasma flow speed, ~400 km/s. The onset of mid-latitude Pi2 pulsations was observed simultaneously with the dipolarization at 11 Re. Mid-latitude Pi2 onset has been interpreted previously as an integrated response to the field-aligned currents from the flows contributing to current wedge formation, but not necessarily due to the current disruption process in the near-Earth region (15). The observed Pi2 onset time at 04:53:05 UT is consistent with such an interpretation.

**Other substorm events.** In the aforementioned event, the plasma sheet was atypically cold and dense, suggesting that the slower Alfvén and magnetosonic speeds may result in longer-than-typical communication times between the various regions within the plasma sheet, as well as between the tail and the ionosphere. This may have been responsible for the easy temporal differentiation of the substorm signatures on the ground and in space observed in this event. To demonstrate that our findings are typical of other events, we examined two additional isolated substorms on the 16th and the 22nd of February, both between 04:30-05:00 UT, and reached similar conclusions (23).

In the events analyzed, the earliest indication of substorm onset followed the first evidence of tail reconnection by <96 s, and preceded the earliest indication of current disruption by >1 min. It is surprising how quickly the aurora intensifies in response to reconnection onset (<96 s). Electron acceleration by reconnection-generated kinetic Alfvén waves may explain this tight coupling between the ionosphere and the reconnecting plasma sheet. Our observations, however, raise another question: what growth-phase process preconditions and destabilizes tail reconnection during spontaneous and externally driven substorms? Since substorm arcs intensify gradually and the magnetotail thins slowly prior to onset (over several minutes), the entire magnetotail from geosynchronous altitude to 30 Re would have enough time to partake in that process.

**References and Notes**

http://www.springerlink.com/content/g139t3795g5495/fultext.pdf (2008).
Supporting Online Material

www.sciencemag.org/cgi/content/full/1160495/DC1

Materials and Methods
Figs. S1 and S2
Movies S1 to S7

14 May 2008; accepted 1 July 2008
Published online 24 July 2008; 10.1126/science.1160495
Include this information when citing this paper.

Fig. 1. Projections of THEMIS probes in X-Z_GSM plane along with representative field lines and neutral sheet location in Geocentric Solar Magnetospheric (GSM) coordinates at 04:45 UT on 26 February 2008. Times refer to the time delays in Table 1.

Fig. 2. (A): THEMIS AE index, AE_TH, computed from THEMIS GBOs (11) and select CARISMA stations; (B) and (C): Band-pass filtered (10 s – 120 s) magnetograms from Gillam (East-West component; station magnetic coordinates: 66.5N, 28.0W) and Carson City, NV (North-South component; 45.1N, 57.0W), in nT. The filter includes the Pi2 band but extends to higher frequencies to reduce aliasing; (D): Integrated auroral intensity from the northern-half of Gillam ASI’s field of view (arbitrary units). (E): Latitude of poleward-most extent of auroral luminosity at station Gillam. (F) and (G): Composite images (mosaics) from THEMIS ASIs around the location of substorm onset, over a continental outline. Stations used in both stills are: Gillam (bottom left), Rankin Inlet (top left; station magnetic coordinates: 72.7N, 25.1W), and Sanikiluaq (bottom right, 66.9N, 3.5W). Gillam had partially cloudy skies, reflecting the moon from low in the horizon and obscuring the northeast and south-west views (the latter is reflection on the dome), though this does not affect our findings. The red line is the midnight meridian. Color circles indicate ionospheric footpoints of THEMIS probes (same symbols as in Fig. 1) using the T96 mapping model (24). The arrow indicates optical onset location.

Fig. 3. Overview of magnetic field and particle data for 30 min, at 3 s resolution, around the 04:54:00 UT substorm. There are 6 panels per probe arranged from top-to-bottom for probes P1, P2, and P3, plus one (bottom) panel showing only 30-200 keV ion spectra on P5, from the SST instrument. The six panels per probe are (from top to bottom): Magnetic field measured by the FGM instrument (35); Ion flow velocity measured by the ESA instrument; Energy spectra of 0.005-2000 keV ions from the SST and ESA instruments (next two panels); Energy spectra of 0.005-2000 keV electrons from the same instruments. All energy spectrograms show omnidirectional differential energy flux (eflux) in units of eV/(cm²-s-ster-eV). The abrupt change in eflux for each species exists because the data below 25 keV and above 30 keV were obtained by two instruments (ESA and SST respectively) with different instrument geometric factors. Magnetic field and velocity are in GSM coordinates; X, Y, Z components are shown in blue, green and red, respectively.

Fig. 4. Data from P1, P2, and P3, during the first few minutes prior to substorm onset. Panels (A), (F), and (H) show the magnetic field as in Figures 2 and 3, except on P1 the resolution is now 4 samples/s. Note that on P1 and P2 the X_GSM component (but not the others) was de-trended (high-pass filtered) by subtracting a 6 min running average, to
reveal details. Panels (B), (G), and (I) show the ion velocity as in Figures 2 and 4. For dotted lines, see text. Panels (C) and (D) show the ion and electron velocity distribution functions near the spin plane in de-spun spacecraft coordinates (+X= Earthward and is to the right of the page; +Y=dawnward and is positive to the top of the page); units are in particles/(cm$^3$ km$^3$/s$^3$) and X, Y velocity planes are in km/s. Panel (E) shows the energy flux spectra of electrons along (0 Deg.), opposite (180 Deg.), and perpendicular (90 Deg.) to the magnetic field for the times indicated.
Table 1. Summary of timing results during the 26 Feb 2008 04:53:45 UT substorm onset, in order of time sequence. The last column is the time delay assuming reconnection onset at 04:50:03 UT, at 20 R_E, which was arrived at based on our interpretation of data and an estimate of an average Alfvén speed in the plasma sheet of 500 km/s.

<table>
<thead>
<tr>
<th>Event</th>
<th>Observed Time (UT)</th>
<th>Inferred delay (seconds since 04:50:03 UT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconnection onset</td>
<td>04:50:03 (inferred)</td>
<td>T_{Re}=0</td>
</tr>
<tr>
<td>Reconnection effects at P1</td>
<td>04:50:28</td>
<td>25</td>
</tr>
<tr>
<td>Reconnection effects at P2</td>
<td>04:50:38</td>
<td>35</td>
</tr>
<tr>
<td>Auroral intensification</td>
<td>04:51:39</td>
<td>T_{AI}=96</td>
</tr>
<tr>
<td>High latitude Pi2 onset</td>
<td>04:52:00</td>
<td>117</td>
</tr>
<tr>
<td>Substorm expansion onset</td>
<td>04:52:21</td>
<td>T_{EX}=138</td>
</tr>
<tr>
<td>Earthward flow onset at P3</td>
<td>04:52:27</td>
<td>144</td>
</tr>
<tr>
<td>Mid-latitude Pi2 onset</td>
<td>04:53:05</td>
<td>182</td>
</tr>
<tr>
<td>Dipolarization at P3</td>
<td>04:53:05</td>
<td>T_{CD}=182</td>
</tr>
<tr>
<td>Auroral Electroject Increase</td>
<td>04:54:00</td>
<td>237</td>
</tr>
</tbody>
</table>